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

Founding Editors:

M. Beckmann
H.P. Künzi

Managing Editors:

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

Part II Contributed Papers

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.

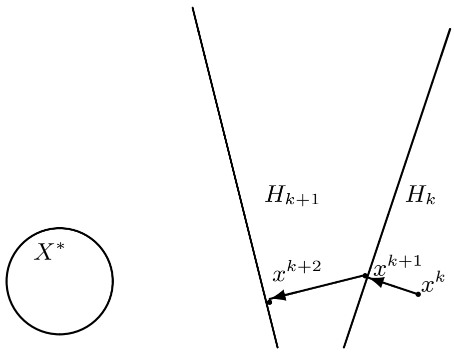

Fig. 1. The relaxation process

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

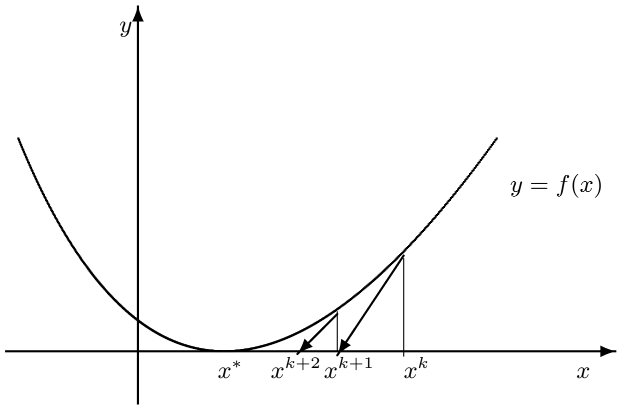

Fig. 2. The Newton method

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

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

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

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

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:

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

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

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

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 $\{T_k : Y \times Y \rightarrow \mathbb{R}^n\}$ 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 $k = 0, 1, \ldots$ 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

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

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

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

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

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

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

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

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

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

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*

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

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

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

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

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

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

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

$i := i + 1$ 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 NrS 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*

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

where $0 < B_0, B_1 < \infty$, *whenever* $0 < \varepsilon' \leq \bar{\varepsilon}$ *and* $0 < \eta' \leq \bar{\eta}$, 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,

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

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 $\Phi : X \times X \to \mathbb{R}$:

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

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*

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

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

$$\Phi(x', x'') \le 0 \Longrightarrow \Phi(x'', x') \ge 0.$$

Hence, $\Phi(x', x'') = 0$ implies $\Phi(x'', x') \ge 0$, but $\Phi(x'', x') > 0$ implies $\Phi(x', x'') < 0$, a contradiction. Thus, $\Phi(x', x'') = 0 \iff \Phi(x'', x') = 0$, 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, $H(f) := \{h \in H : h \le f\}$. 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 $c : \Omega = X \times Y \to \mathrm{I\!R} \cup \{+\infty\}$ 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 $c : \Omega = X \times X \to \mathrm{I\!R} \cup \{+\infty\}$ 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 $H = \{h = h_{u\alpha} : h_{u\alpha}(x, y) = u(x) - u(y) - \alpha,\ u \in C(X),\ \alpha \in \mathrm{I\!R}_+\}$, 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 $c : \Omega = X \times X \to \mathrm{I\!R} \cup \{+\infty\}$ of the form $c(x, y) = \varphi(x, y) - \alpha$, where $\alpha \in \mathrm{I\!R}_+, \varphi$ 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 $c(x, x) = const \leq 0 \quad \forall x \in X$.

Now suppose that $\mu_0 = \delta_\omega$ is the Dirac measure (delta function) at some point $\omega \in \Omega$, $\langle u, \delta_\omega\rangle := u(\omega) \quad$ 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.*

Proof. Taking into account that $v_I(c;\delta_\omega) = co_H(c)(\omega)$, the equivalence $(a) \Leftrightarrow (b)$ is exactly the statement of Corollary 3. The equivalence $(b) \Leftrightarrow (c)$ is a particular case of Proposition 1. □

We now consider two more general mutually dual linear programs, as follows. Suppose that E, E' is a pair of linear spaces in duality relative to a bilinear form $\langle e, e'\rangle_E$, $e \in E$, $e' \in E'$. We endow them with the corresponding weak topologies: $\sigma(E, E')$ and $\sigma(E', E)$. Given a convex cone K in E, a functional $e'_0 \in E'$, and a weakly continuous (i.e., continuous relative to the weak topology in the Banach space $C(\Omega)$ and the weak topology $\sigma(E, E')$ in E) linear map $A : E \to C(\Omega)$ such that the set $\{e \in K : Ae \le c\}$ is nonempty, one has to find the optimal values

$$v'_I(c; e'_0) := \sup\{\langle e, e'_0\rangle_E : e \in K, Ae \le c\}, \tag{13}$$

$$v'_{II}(c; e'_0) := \inf\{c(\mu) : \mu \ge 0, A^*\mu \in e'_0 - K^0\}, \tag{14}$$

where K^0 is the convex cone in E' conjugate to K,

$$K^0 := \{e' \in E' : \langle e, e'\rangle_E \le 0 \quad \text{for all } e \in K\}.$$

Clearly, both the functionals, (13) and (14), are sublinear, and $v'_I(c; e'_0) \le v'_{II}(c; e'_0)$. Similarly to Proposition 1, the next result is a particular case of Lemma 5.1 (see also Remark 1 after it) in [47].

Proposition 5. *Given* $e'_0 \in \operatorname{dom} v'_I(c;\cdot) := \{e' \in E' : v'_I(c; e') < +\infty\}$, *the following assertions are equivalent:*

(a) $v'_I(c; e'_0) = v'_{II}(c; e'_0)$;

(b) the functional $v'_{II}(c;\cdot)$ *is weakly lower semi-continuous at* e'_0.

Let us define $H := AK$; then $H^0 = (A^*)^{-1}(K^0)$.

Remark 3. Note that if

$$\operatorname{dom} v'_I(c;\cdot) \subseteq A^*C(\Omega)^*_+, \tag{15}$$

then, for every $e'_0 \in \operatorname{dom} v'_I(c;\cdot)$,

$$v'_I(c; e'_0) = v_I(c; \mu_0) \quad \text{and} \quad v'_{II}(c; e'_0) = v_{II}(c; \mu_0)$$

whenever $\mu_0 \in (A^*)^{-1}(e'_0)$. Also note that, for $e'_0 \notin \operatorname{dom} v'_I(c;\cdot)$, one has $v'_I(c; e'_0) = v'_{II}(c; e'_0) = +\infty$.

The next result follows then from Corollary 1.

Corollary 4. *Suppose the regularity assumption is satisfied. If (15) is valid, then the duality relation* $v'_I(c; e'_0) = v'_{II}(c; e'_0)$ *holds true whenever* c *is* H*-convex and* $e'_0 \in E'$. *If, in addition,* $e'_0 \in \operatorname{dom} v'_I(c;\cdot)$ *then there exists an optimal solution to program II.*

3 Abstract Convexity and the Monge - Kantorovich Problems (MKP)

In this section, we consider two variants of the Monge—Kantorovich problem (MKP), with given marginals and with a given marginal difference. Both the problems are infinite linear programs, and abstract convexity plays important role in their study. Abstract cyclic monotonicity along with optimality criteria for MKP will be studied as well.

3.1 MKP with Given Marginals

Let X and Y be compact Hausdorff topological spaces[4], σ_1 and σ_2 finite positive regular Borel measures on them, $\sigma_1 X = \sigma_2 Y$, and $c : X \times Y \to \mathbb{R} \cup \{+\infty\}$ an universally measurable function bounded from below. The natural projecting maps of $X \times Y$ onto X and Y will be denoted as π_1 and π_2, respectively.

The MKP with given marginals is to find the optimal value

$$\mathcal{C}(c; \sigma_1, \sigma_2) := \inf\{c(\mu) : \mu \geq 0, \pi_1\mu = \sigma_1, \pi_2\mu = \sigma_2\} \tag{16}$$

where

$$c(\mu) := \int_{X\times Y} c(x, y)\, \mu(d(x, y)), \tag{17}$$

$$(\pi_1\mu)B_1 = \mu\pi_1^{-1}(B_1) = \mu(B_1 \times Y) \text{ for every Borel set } B_1 \subset X,$$
$$(\pi_2\mu)B_2 = \mu\pi_2^{-1}(B_2) = \mu(X \times B_2) \text{ for every Borel set } B_2 \subset Y.$$

The dual problem is to find the optimal value

$$\mathcal{D}(c; \sigma_1, \sigma_2) := \sup\{\langle u, \sigma_1\rangle - \langle v, \sigma_2\rangle : (u, v) \in Q'(c)\}, \tag{18}$$

where

$$Q'(c) = \{(u, v) \in C(X) \times C(Y) : u(x) - v(y) \leq c(x, y), (x, y) \in X \times Y\}. \tag{19}$$

Clearly, always

$$\mathcal{D}(c; \sigma_1, \sigma_2) \leq \mathcal{D}'(c; \sigma_1, \sigma_2) \leq \mathcal{C}(c; \sigma_1, \sigma_2), \tag{20}$$

where $\mathcal{D}'(c; \sigma_1, \sigma_2)$ stands for supremum of $\int_X u(x)\, \sigma_1(dx) - \int_Y v(y)\, \sigma_2(dy)$ over all pairs of bounded Borel functions (u, v) satisfying $u(x) - v(y) \leq c(x, y)$ whenever $x \in X, y \in Y$.

Let H be as in Example 1, $E := C(X) \times C(Y)$, $E' := C(X)^* \times C(Y)^*$, $\langle e, e'\rangle_E := \langle u, \sigma_1'\rangle - \langle v, \sigma_2'\rangle$ for all $e = (u, v) \in E$, $e' = (\sigma_1', \sigma_2') \in E'$, $K := E$,

[4] For the sake of simplicity, we assume X and Y to be compact; however, the corresponding duality theorem (Theorem 1 below) holds true for any Hausdorff completely regular spaces; see [42, Theorem 1.4] and [43, Theorem 10.3]. See also [29, Theorem 1].

and $A : E \to C(X \times Y)$ is given by $Ae(x, y) := u(x) - v(y)$, $e = (u, v)$. Clearly, $H = AK$, $Q'(c) = A^{-1}(H(c))$, and

$$\mathcal{C}(c; \sigma_1, \sigma_2) = v'_{II}(c; e'_0), \quad \mathcal{D}(c; \sigma_1, \sigma_2) = v'_I(c; e'_0),$$

where $e'_0 = (\sigma_1, \sigma_2)$, $v'_I(c; e'_0)$ and $v'_{II}(c; e'_0)$ are given by (13) and (14), respectively. Note that (15) is satisfied.

Theorem 1. ([42, Theorem 1.4]). *The following statements are equivalent:*

(a) c is H-convex;

(b) c is bounded below and lsc;

(c) the duality relation $\mathcal{C}(c; \sigma_1, \sigma_2) = \mathcal{D}(c; \sigma_1, \sigma_2)$ *holds for all* $\sigma_1 \in C(X)^*_+$, $\sigma_2 \in C(Y)^*_+$.

Moreover, if these equivalent statements hold true then, for any positive measures σ_1, σ_2 *with* $\sigma_1 X = \sigma_2 Y$, *there exists an optimal solution to the MKP with marginals* σ_1, σ_2.

Proof. $(a) \Leftrightarrow (b)$ See Remark 1.

$(a) \Rightarrow (c)$ Taking into account Example 1, this follows from Corollary 4.

$(c) \Rightarrow (a)$ Since $\mu = \delta_{(x,y)}$ is the sole positive measure with marginals $\sigma_1 = \delta_x, \sigma_2 = \delta_y$, one gets $\mathcal{C}(c; \delta_x, \delta_y) = c(x, y)$. Now, taking into account Remark 3, we see that $v_I(c; \delta_{(x,y)}) = \mathcal{D}(c; \delta_x, \delta_y)$, $v_{II}(c; \delta_{(x,y)}) = \mathcal{C}(c; \delta_x, \delta_y)$; therefore, $c = c_\#$, and applying Proposition 4 completes the proof.

Finally, the latter statement of the theorem is a particular case of the latter assertion of Corollary 4. □

3.2 MKP with a Given Marginal Difference

Let X be a compact Hausdorff topological space[5], $\rho \in C(X \times X)^*$ a signed measure satisfying $\rho X := \langle \mathbf{1}_X, \rho \rangle = 0$, and $c : X \times X \to \mathbb{R} \cup \{+\infty\}$ an universally measurable function bounded from below. As before, π_1 and π_2 stand for the projecting maps of $X \times X$ onto the first and the second coordinates, respectively. The corresponding marginals of a measure $\mu \in C(X \times X)^*_+$ are designated as $\pi_1\mu$ and $\pi_2\mu$.

The MKP with a given marginal difference is to find the optimal value

$$\mathcal{A}(c; \rho) := \inf\{c(\mu) : \mu \geq 0, \ \pi_1\mu - \pi_2\mu = \rho\}, \tag{21}$$

where

$$c(\mu) := \int_{X \times X} c(x, y)\, \mu(d(x, y)). \tag{22}$$

The dual problem is to find the optimal value

[5]For the sake of simplicity, we assume X to be compact; however, the corresponding duality theorems (Theorems 2 and 3 below) hold true for more general spaces (in particular, for any Polish space); see [32, Theorems 9.2 and 9.4], [42, Theorem 1.2] and [43, Theorem 10.1 and 10.2].

$$\mathcal{B}(c;\rho) := \sup\{\langle u,\rho\rangle : u \in Q(c)\}, \tag{23}$$

where

$$Q(c) = \{u \in C(X) : u(x) - u(y) \le c(x,y) \quad \forall x,y \in X\}. \tag{24}$$

(Note that $\mathcal{A}(c;\rho) = \mathcal{B}(c;\rho) = +\infty$ when $\rho X \neq 0$.)

Suppose H is as in Example 2, $E := C(X)$, $E' := C(X)^*$, $\langle e,e'\rangle_E := \langle u,\sigma\rangle$ for all $e = u \in E$, $e' = \sigma \in E'$, $K := E$, and $A : E \to C(X \times X)$ is given by $Ae(x,y) := u(x) - u(y)$, $e = u$. Clearly, $H = AK$, $Q(c) = A^{-1}(H(c))$ (hence $H(c)$ is nonempty if and only if $Q(c)$ is such), and

$$\mathcal{A}(c;\rho) = v'_{II}(c;e'_0), \quad \mathcal{B}(c;\rho) = v'_I(c;e'_0), \tag{25}$$

where $e'_0 = \rho$, $v'_I(c;e'_0)$ and $v'_{II}(c;e'_0)$ are given by (13) and (14), respectively. Note that

$$\operatorname{dom} v'_I(c;\cdot) = \operatorname{dom} \mathcal{A}(c;\cdot) \subseteq A^*C(X\times X)^*_+ = \{\rho \in C(X)^* : \rho X = 0\}.$$

Let $U(X)$ stands for the class of all bounded universally measurable functions on X,

$$Q(c;U(X)) := \{v \in U(X) : v(x) - v(y) \le c(x,y) \quad \forall (x,y) \in X \times X\}.$$

Theorem 2. (cf. [47, Theorems 3.1, 3.2 and 4.4], [42, Theorem 1.2], [43, Theorem 10.1]). *Suppose that c is an universally measurable function vanishing on the diagonal $D = \{(x,x) : x \in X\}$ and satisfying the triangle inequality, the following statements are then equivalent:*

(a) c is H-convex relative to H from Example 2, that is $Q(c) \neq \varnothing$ and

$$c(x,y) = \sup\{u(x) - u(y) : u \in Q(c)\} \quad \text{for all } x,y \in X; \tag{26}$$

(b) c is bounded below and lsc;

(c) $Q(c) \neq \varnothing$, and the duality relation $\mathcal{A}(c;\rho) = \mathcal{B}(c;\rho)$ holds for all $\rho \in C(X)^*$;

(d) $Q(c;U(X)) \neq \varnothing$, and the duality relation $\mathcal{A}(c;\rho) = \mathcal{B}(c;\rho)$ holds for all $\rho \in C(X)^*$, $\rho X = 0$.

Moreover, if these equivalent statements hold, then, for any ρ, $\rho X = 0$, and for any positive measures σ_1, σ_2 with $\sigma_1 - \sigma_2 = \rho$, there is a measure $\mu \in C(X \times X)^*_+$ such that $\pi_1\mu = \sigma_1$, $\pi_2\mu = \sigma_2$ and $\mathcal{A}(c;\rho) = \mathcal{C}(c;\sigma_1,\sigma_2) = c(\mu)$.

Proof. $(a) \Rightarrow (b)$ and $(c) \Rightarrow (d)$ are obvious; as for $(b) \Rightarrow (a)$, see Example 2. The implication $(a) \Rightarrow (c)$ and the latter statement of the theorem follow from Corollary 4 if one takes into account Example 2 along with identities (25). The proof will be complete if we show that (d) implies (a). Suppose (26) fails; then

$$c(x_0,y_0) > \sup\{u(x_0) - u(y_0) : u \in Q(c)\} = \mathcal{B}(c;\delta_{x_0} - \delta_{y_0}) \tag{27}$$

for some $x_0, y_0 \in X$ with $x_0 \neq y_0$; hence, $\mathcal{B}(c; \delta_{x_0} - \delta_{y_0}) < +\infty$. We define the function

$$c'(x, y) := \min\{c(x, y) - v(x) + v(y), N\} + v(x) - v(y), \tag{28}$$

where $v \in Q(c; U(X))$,

$$N > \max\{0, \mathcal{B}(c; \delta_{x_0} - \delta_{y_0}) - v(x_0) + v(y_0)\}. \tag{29}$$

Clearly, it is bounded and universally measurable, and $c' \leq c$. Furthermore, c' satisfies the triangle inequality (this is easily derived from non-negativeness of $c(x, y) - v(x) + v(y)$); therefore, $w(x) := c'(x, y_0)$ belongs to $Q(c; U(X))$. Consider

$$\mathcal{B}(c; \rho; U(X)) := \sup\{\langle v', \rho\rangle : v' \in Q(c; U(X))\}$$

and note an obvious inequality

$$\mathcal{A}(c; \rho) \geq \mathcal{B}(c; \rho; U(X)) \quad \forall \rho,\ \rho X = 0. \tag{30}$$

Now, taking into account (28) - (30), one gets

$$\begin{aligned}\mathcal{A}(c; \delta_{x_0} - \delta_{y_0}) \geq \mathcal{B}(c; \delta_{x_0} - \delta_{y_0}; U(X)) &\geq w(x_0) - w(y_0)\\ = c'(x_0, y_0) &> \mathcal{B}(c; \delta_{x_0} - \delta_{y_0}),\end{aligned}$$

which contradicts the duality relation. □

The next Proposition supplements Theorem 2.

Proposition 6. . *Suppose $c : X \times X \to \mathbb{R}$ is bounded universally measurable, vanishes on the diagonal, and satisfies the triangle inequality, then $Q(c; U(X))$ is nonempty.*

Proof. Let us fix an arbitrary point $y_0 \in X$ and consider the function $v(x) = c(x, y_0)$. Clearly, it is universally measurable, real-valued and bounded, and as c satisfies the triangle inequality, one has $v \in Q(c; U(X))$. □

Remark 4. Suppose that $c : X \times X \to \mathbb{R} \cup \{+\infty\}$ satisfies the triangle inequality and vanishes on the diagonal. As follows from [47, Theorem 3.3[6]], $Q(c; U(X))$ is nonempty if c is Baire measurable or if its Lebesgue sublevel sets $L(c; \alpha) = \{(x, y) \in X \times X : c(x, y) \leq \alpha\}$, $\alpha \in \mathbb{R}$, are the results of applying the A-operation to Baire subsets of $X \times X$. (If X is metrizable, the latter means that all $L(c; \alpha)$ are analytic (Souslin).)

Now consider the case where the cost function $c : X \times X \to \mathbb{R} \cup \{+\infty\}$ vanishes on the diagonal but fails to satisfy the triangle inequality, and define the *reduced cost function* c_* associated with it as follows:

[6] See also [32, Theorem 9.2 (III)], where a more general result is proved.

$$c_*(x,y) := \inf_n \inf \left\{ \sum_{i=1}^{n+1} c(x_{i-1},x_i) : x_i \in X, x_0 = x, x_{n+1} = y \right\}$$
$$= \lim_{n\to\infty} \inf \left\{ \sum_{i=1}^{n+1} c(x_{i-1},x_i) : x_i \in X, x_0 = x, x_{n+1} = y \right\}. \qquad (31)$$

Clearly, $c_* \le c$, c_* satisfies the triangle inequality (we assume, by definition, that $+\infty + (-\infty) = +\infty$), and $H(c_*) = H(c)$; therefore, $Q(c_*) = Q(c)$, and if $Q(c)$ is nonempty, then c_* does not take the value $-\infty$ and is bounded from below. We get

$$\mathcal{B}(c;\rho) = \mathcal{B}(c_*;\rho) \le \mathcal{A}(c_*;\rho) \le \mathcal{A}(c;\rho) \quad \forall \rho \in C(X)^*. \qquad (32)$$

Proposition 7. *Suppose c_* is universally measurable. If $Q(c)$ is nonempty and $\mathcal{A}(c;\rho) = \mathcal{B}(c;\rho)$ for all $\rho \in C(X)^*$, then c_* is H-convex and $c_* = co_H(c) = c_\#$ where $c_\#$ is given by (12). In such a case,*

$$c_*(x,y) = \sup_{u\in Q(c)} (u(x) - u(y)) \quad \text{for all } x,y \in X. \qquad (33)$$

Proof. It follows from (32) that $\mathcal{A}(c_*;\rho) = \mathcal{B}(c_*;\rho)$ for all $\rho \in C(X)^*$. Note that c_* vanishes on the diagonal because c vanishes on the diagonal and $Q(c_*) = Q(c) \ne \varnothing$. Now, applying Theorem 2 yields H-convexity of c_*, and as $H(c) = H(c_*)$, one gets $c_* = co_H(c)$. Finally, the duality relation $\mathcal{A}(c;\delta_x - \delta_y) = \mathcal{B}(c;\delta_x - \delta_y)$ may be rewritten as $v_I(c;\delta_{(x,y)}) = v_{II}(c;\delta_{(x,y)})$ (see Remark 3), and applying Proposition 4 and Corollary 3 yields $c_\# = co_H(c)$. □

Remark 5. As is proved in [47, Lemma 4.2], if the Lebesgue sublevel sets of c, $L(c;\alpha) = \{(x,y) \in X \times X : c(x,y) \le \alpha\}$, $\alpha \in \mathbb{R}$, are the results of applying the A-operation to Baire subsets of $X \times X$, then Lebesgue sublevel sets of c_*, $L(c_*;\alpha)$, $\alpha \in \mathbb{R}$, have the same property hence c_* proves to be universally measurable.

The next result is a direct consequence of (32) and Theorem 2.

Proposition 8. *If c_* is H-convex and $\mathcal{A}(c;\rho) = \mathcal{A}(c_*;\rho)$, then $\mathcal{A}(c;\rho) = \mathcal{B}(c;\rho)$.*

Remark 6. As is established in [32, Theorem 9.6] and (for a metrizable case) in [47, Theorem 2.1], a *reduction theorem* is true: if the Lebesgue sublevel sets of c are the results of applying the A-operation to Baire subsets of $X \times X$, then $\mathcal{A}(c;\rho) = \mathcal{A}(c_*;\rho)$ provided that the equality holds

$$\mathcal{A}(c;\rho) = \lim_{N\to\infty} \mathcal{A}(c \wedge N;\rho), \qquad (34)$$

where $(c \wedge N)(x,y) = \min\{c(x,y), N\}$. (Note that, for a bounded c, (34) is trivial.)

Taking into account Remarks 5 and 6, the next result is derived from Propositions 7, 8 and the reduction theorem.

Theorem 3. (cf. [47, Theorems 3.1 and 3.2], [32, Theorem 9.4], [43, Theorem 10.2]). *Suppose that c is bounded from below and vanishes on the diagonal, and that its sublevel sets are the results of applying the A-operation to Baire subsets of $X \times X$. The following statements are then equivalent:*

(a) the reduced cost function c_* is H-convex, and condition (34) is satisfied whenever $\rho X = 0$;

(b) $Q(c)$ is nonempty, and the duality relation $\mathcal{A}(c;\rho) = \mathcal{B}(c;\rho)$ holds for all $\rho \in C(X)^*$.

Proof. $(a) \Rightarrow (b)$ Taking into account the reduction theorem (see Remark 6), this follows from Proposition 8.

$(b) \Rightarrow (a)$ In accordance with Remark 5, c_* is universally measurable; then, by Proposition 7, it is H-convex. It remains to show that (34) is satisfied. First, note that, being a bounded function, every $u \in Q(c)$ belongs to $Q(c \wedge N)$, where $N = N(u) > 0$ is large enough; therefore,

$$\mathcal{B}(c;\rho) = \lim_{N\uparrow\infty} \mathcal{B}(c \wedge N;\rho).$$

Now, by using the monotonicity of $\mathcal{A}(c;\rho)$ in c, one gets

$$\begin{aligned}\mathcal{A}(c;\rho) &\geq \limsup_{N\uparrow\infty} \mathcal{A}(c\wedge N;\rho) \geq \liminf_{N\uparrow\infty} \mathcal{A}(c\wedge N;\rho)\\ &\geq \lim_{N\uparrow\infty} \mathcal{B}(c\wedge N;\rho) = \mathcal{B}(c;\rho),\end{aligned}$$

which clearly implies (34). □

Corollary 5. *Suppose c is Baire measurable, bounded from below, and vanishes on the diagonal. Then c_* is H-convex if and only if $Q(c)$ is nonempty and $\mathcal{A}(c;\rho) = \mathcal{B}(c;\rho)$ for all $\rho \in C(X)^*$.*

3.3 A Connection Between Two Variants of MKP

Given compact Hausdorff topological spaces X and Y, we define $X \oplus Y$ to be the formal union $X \cup Y$ of disjoint copies of X and Y with the topology of direct sum: by definition, a set G is open in $X \oplus Y$ if $G \cap X$ is open in X and $G \cap Y$ is open in Y. Clearly, $X \oplus Y$ is compact, both X and Y are open-closed in it, and $C(X \oplus Y) = C(X) \times C(Y)$. Furthermore, $C(X \oplus Y)^* = C(X)^* \times C(Y)^*$, that is a pair $(\sigma_1, \sigma_2) \in C(X)^* \times C(Y)^*$ is identified with a measure $\hat{\sigma} \in C(X \oplus Y)^*$,

$$\hat{\sigma}B = \sigma_1(B \cap X) + \sigma_2(B \cap Y) \text{ for any Borel } B \subseteq X \oplus Y,$$

and every $\hat{\sigma} \in C(X \oplus Y)^*$ is obtained in such a way. We shall write this as $\hat{\sigma} = (\sigma_1, \sigma_2)$.

Given $\sigma_1 \in C(X)^*_+$ and $\sigma_2 \in C(Y)^*_+$, we associate them with the measures $\hat{\sigma}_1 = (\sigma_1, 0), \hat{\sigma}_2 = (0, \sigma_2) \in C(X \oplus Y)^*_+$. Similarly, every $\mu \in C(X \times Y)^*_+$ is associated with the measure $\hat{\mu} \in C((X \oplus Y) \times (X \oplus Y))^*_+$,

$$\hat{\mu}B := \mu(B \cap (X \times Y)) \text{ for any Borel } B \subseteq (X \oplus Y) \times (X \oplus Y). \tag{35}$$

Given a cost function $c : X \times Y \to \mathbb{R} \cup \{+\infty\}$, every pair $(u, v) \in Q'(c)$ is identified with a function $w \in C(X \oplus Y)$, $w|X = u$, $w|Y = v$, which belongs to $Q(\hat{c})$ for

$$\hat{c}(z, z') := \sup\{w(z) - w(z') : w = (u, v) \in Q'(c)\},\ z, z' \in X \oplus Y, \tag{36}$$

where $Q'(c) \subset C(X \oplus Y)$ is defined as in (19). Clearly, $\hat{c}$ is lsc, vanishes on the diagonal, and satisfies the triangle inequality, c majorizes the restriction of $\hat{c}$ onto $X \times Y$, and $Q(\hat{c}) = Q'(c)$. Note that if c coincides with the restriction of $\hat{c}$ onto $X \times Y$ then $\mathcal{C}(c; \sigma_1, \sigma_2) = \mathcal{C}(\hat{c}; \hat{\sigma}_1, \hat{\sigma}_2)$.

Proposition 9. (cf. [42, Theorem 1.5] and [26, Lemma 7]). *I. Given a cost function $c : X \times Y \to \mathbb{R} \cup \{+\infty\}$, the following statements are equivalent:*

(a) c is H-convex relative to H from Example 1;

(b) c is the restriction to $X \times Y$ of a function $\hat{c}$ on $(X \oplus Y) \times (X \oplus Y)$, which is H-convex relative to $H \subset C((X \oplus Y) \times (X \oplus Y))$ from Example 2.

If these equivalent statements hold, then $Q(\hat{c}) = Q'(c)$ and

$$\mathcal{A}(\hat{c}; \hat{\sigma}_1 - \hat{\sigma}_2) = \mathcal{B}(\hat{c}; \hat{\sigma}_1 - \hat{\sigma}_2) = \mathcal{C}(c; \sigma_1, \sigma_2) = \mathcal{D}(c; \sigma_1, \sigma_2) > -\infty$$

whenever $\sigma_1 \in C(X)^*_+, \sigma_2 \in C(Y)^*_+$, $\sigma_1 X = \sigma_2 Y$.

II. If $c \in C(X \times Y)$, then there is a continuous function $\hat{c}$ satisfying (b).

Proof. I. This follows easily from Theorems 1 and 2 if one takes $\hat{c}$ as given by (36).

II. Define $\hat{c}$ as follows:

$$\hat{c}(z_1, z_2) = \begin{cases} c(x, y), \text{ if } z_1 = x \in X, \ z_2 = y \in Y; \\ c_1(x_1, x_2), \text{ if } z_1 = x_1 \in X, z_2 = x_2 \in X; \\ c_2(y_1, y_2), \text{ if } z_1 = y_1 \in Y, \ z_2 = y_2 \in Y; \\ c_3(y, x), \text{ if } z_1 = y \in Y, \ z_2 = x \in X, \end{cases} \tag{37}$$

where

$$\begin{aligned} c_1(x_1, x_2) &= \max_{y \in Y} \left(c(x_1, y) - c(x_2, y)\right), \\ c_2(y_1, y_2) &= \max_{x \in X} \left(c(x, y_2) - c(x, y_1)\right), \\ c_3(y, x) &= \max_{x_1 \in X, y_1 \in Y} \left(c(x_1, y_1) - c(x_1, y) - c(x, y_1)\right). \end{aligned}$$

Clearly, $\hat{c}$ is continuous, vanishes on the diagonal, and $\hat{c}|_{X \times Y} = c$. Moreover, a direct testing shows that it satisfies the triangle inequality. Then $\hat{c}$ is H-convex with respect to H from Example 2, and the result follows. □

The next result supplements Theorem 1.

Proposition 10. (cf. [26, Theorem 5]). *Suppose* $c \in C(X \times Y)$, $\sigma_1 \in C(X)^*_+$, $\sigma_2 \in C(Y)^*_+$, *and* $\sigma_1 X = \sigma_2 Y$, *then there is an optimal solution* $(u, v) \in C(X) \times C(Y)$ *to the dual MKP, that is,* $(u, v) \in Q'(c)$ *and*

$$\int_X u(x)\,\sigma_1(dx) - \int_Y v(y)\,\sigma_2(dy) = \mathcal{D}(c; \sigma_1, \sigma_2).$$

Proof. Take a function $\hat{c} \in C((X \oplus Y) \times (X \oplus Y))$ from Proposition 9, II (see (37)) and fix arbitrarily a point $z_0 \in X \oplus Y$. Since $\hat{c}$ is continuous and vanishes on the diagonal, the set

$$Q(\hat{c}; z_0) := \{w \in Q(\hat{c}) : \, w(z_0) = 0\}$$

is compact in $C(X \oplus Y)$ and there exists a function $w_0 = (u, v) \in Q(\hat{c}; z_0)$ such that $\langle w_0, \hat{\sigma}_1 - \hat{\sigma}_2\rangle = \max\{\langle w, \hat{\sigma}_1 - \hat{\sigma}_2\rangle : \, w \in Q(\hat{c}; z_0)\}$. Now, taking into account an obvious equality $Q(\hat{c}) = Q(\hat{c}; z_0) + \mathbb{R}$ and applying Proposition 9, one gets

$$\int_X u(x)\,\sigma_1(dx) - \int_Y v(y)\,\sigma_2(dy) = \langle w_0, \hat{\sigma}_1 - \hat{\sigma}_2\rangle$$
$$= \max\{\langle w, \hat{\sigma}_1 - \hat{\sigma}_2\rangle : \, w \in Q(\hat{c})\} = \mathcal{B}(\hat{c}; \hat{\sigma}_1 - \hat{\sigma}_2) = \mathcal{D}(c; \sigma_1, \sigma_2). \qquad \square$$

3.4 Abstract Cyclic Monotonicity and Optimality Conditions for MKP

Given a set X and a subset L in $\mathbb{R}^X$, a multifunction $F : X \to L$ is called *L-cyclic monotone* if, for every cycle $x_1, \ldots, x_m, x_{m+1} = x_1$ in $\operatorname{dom} F = \{x \in X : \, F(x) \neq \varnothing\}$, the inequality holds

$$\sum_{k=1}^{m} (l_k(x_k) - l_k(x_{k+1})) \geq 0 \tag{38}$$

whenever $l_k \in F(x_k)$, $k = 1, \ldots, m$. By changing the sign of this inequality, one obtains the definition of *L-cyclic antimonotone* multifunction. Clearly, F is L-cyclic monotone if and only if $-F$ is $(-L)$-cyclic antimonotone.

We say a function $U : X \to \mathbb{R} \cup \{+\infty\}$ is *L-convex* if it is H-convex relative to

$$H := \{h_{l\alpha} : \, h_{l\alpha}(x) = l(x) - \alpha, \, (l, \alpha) \in L \times \mathbb{R}\}. \tag{39}$$

A function $V : X \to \mathbb{R} \cup \{-\infty\}$ is said to be *L-concave* if $U = -V$ is $(-L)$-convex.

Examples of L-cyclic monotone multifunctions are *L-subdifferentials* of L-convex functions, $\partial_L U : X \to L$, where

$$\partial_L U(x) := \{l \in L : \, l(z) - l(x) \leq U(z) - U(x) \quad \text{for all } z \in X\}. \tag{40}$$

Similarly, examples of L-cyclic antimonotone multifunctions are *L-superdifferentials* of L-concave functions, $\partial^L V : X \to L$, where

$$\partial^L V(x) := \{l \in L : \, l(z) - l(x) \geq V(z) - V(x) \quad \text{for all } z \in X\}. \tag{41}$$

It is obvious from the above definitions that, for every L-concave V,

$$\partial^L V = \partial_{(-L)}(-V). \tag{42}$$

Remark 7. The classic monotone (antimonotone) multifunctions can be considered as examples of L-cyclic monotone (resp., antimonotone) ones, answering the case where X is a Hausdorff locally convex space, $L = X^*$ is the dual space, and $l(x) = \langle x, l \rangle$, $x \in X$, $l \in L$. Close notions of c-monotonicity (c-antimonotonicity) and of c-subdifferentials of c-convex functions (c-superdifferentials of c-concave functions) are widespread in literature; e.g., see [12, 55, 63]. A connection between the corresponding L-concepts and c-concepts is discussed in [40].

Given a multifunction $F : X \to L$, we denote $Z = \operatorname{dom} F := \{z \in X : F(z) \neq \varnothing\}$ and consider two functions $Z \times Z \to \mathbb{R} \cup \{-\infty\}$ as follows:

$$\varphi_F(z_1, z_2) = \varphi_{F,L}(z_1, z_2) := \inf\{l(z_1) - l(z_2) : l \in F(z_1)\}, \tag{43}$$

$$\psi_F(z_1, z_2) = \psi_{F,L}(z_1, z_2) := \inf\{l(z_1) - l(z_2) : l \in F(z_2)\}. \tag{44}$$

Clearly, $\psi_{F,L}(z_1, z_2) = \varphi_{(-F),(-L)}(z_2, z_1)$.

Remark 8. Note that if $\sup\limits_{l \in L} |l(z)| < \infty$ for every $z \in Z$, then both the functions are real-valued.

Given a function $\zeta : Z \times Z \to \mathbb{R} \cup \{-\infty\}$ vanishing on the diagonal ($\zeta(z, z) = 0 \quad \forall z \in Z$), we consider the set

$$Q_0(\zeta) := \{u \in \mathbb{R}^Z : \, u(z_1) - u(z_2) \leq \zeta(z_1, z_2) \quad \forall z_1, z_2 \in Z\}. \tag{45}$$

It follows from (45) that if $Q_0(\zeta)$ is nonempty, then ζ is real-valued. Clearly, $Q_0(\zeta) = Q_0(\zeta_*)$ where ζ_* is the reduced cost function associated with ζ (for the definition of the reduced cost function, see (31)). Also, observe that if Z is a topological space and ζ is a bounded continuous function on $Z \times Z$ vanishing on the diagonal, then $Q_0(\zeta) = Q(\zeta)$. (Here, $Q(\zeta)$ is defined for a compact Z as in (26), and if Z is not compact, we define $Q(\zeta)$ to be the set of all *bounded* continuous functions u satisfying (45).)

Theorem 4. ([40, Theorem 2.1]). *A multifunction $F : X \to L$ is L-cyclic monotone if and only if $Q_0(\varphi_F)$ is nonempty.*

Theorem 5. ([40, Theorem 2.2]). *Suppose $F : X \to L$ is L-cyclic monotone. Given a function $u : Z = \operatorname{dom} F \to \mathbb{R} \cup \{+\infty\}$, the following statements are equivalent:*

(a) $u \in Q_0(\varphi_F)$;

(b) u is a restriction to Z of some L-convex function $U : X \to \mathbb{R} \cup \{+\infty\}$, and $F(z) \subseteq \partial_L U(z)$ for all $z \in Z$.

The next result extending a classical convex analysis theorem due to Rockafellar is an immediate consequence of Theorems 4 and 5.

Corollary 6. ([39], [40], [53][7]). *A multifunction $F : X \to L$ is L-cyclic monotone if and only if there is a L-convex function $U : X \to \mathbb{R} \cup \{+\infty\}$ such that $F(x) \subseteq \partial_L U(x)$ for all $x \in X$.*

Suppose $F : X \to L$ is a L-cyclic monotone multifunction. We say F is *maximal L-cyclic monotone* if $F = T$ for any L-cyclic monotone multifunction T such that $F(x) \subseteq T(x)$ whenever $x \in X$.

Theorem 6. ([40, Theorem 2.3]). *A multifunction $F : X \to L$ is maximal L-cyclic monotone if and only if $F = \partial_L U$ for all L-convex functions $U : X \to \mathbb{R} \cup \{+\infty\}$ satisfying $U|\operatorname{dom} F \in Q_0(\varphi_F)$.*

Remark 9. Theorem 6 is an abstract version of the corresponding classical result due to Rockafellar [58]. In classical setting, X is a Hausdorff locally convex space, $L = X^*$ is the conjugate space, and $l(x) = \langle x, l\rangle$. In this case, L-convex functions, their L-subdifferentials, and L-cyclic monotone multifunctions are, respectively, convex lsc functions, their subdifferentials, and classical cyclic monotone multifunction $X \to X^*$. Rockafellar's theorem says that maximal cyclic monotone multifunctions are exactly the subdifferentials of lsc convex functions, and if U_1 and U_2 are two such functions with $\partial U_1 = \partial U_2$, then $U_1 - U_2$ is a constant function. However, in general case both these assertions fail: there is a L-convex function, for which $\partial_L U$ is not maximal, and there are two L-convex functions, U_1 and U_2, such that the multifunction $F = \partial_L U_1 = \partial_L U_2$ is maximal L-cyclic monotone but the difference $U_1 - U_2$ is not constant. The corresponding counter-example can be seen in [40, Example 2.1].

Let X, Y be compact Hausdorff topological spaces. Given a cost function $c \in C(X \times Y)$, we consider the MKP with marginals σ_1 and σ_2, $\sigma_1 X = \sigma_2 Y$. Recall (see subsection 3.1), that it is to find the optimal value

$$\mathcal{C}(c; \sigma_1, \sigma_2) = \inf\{c(\mu) : \mu \in \Gamma(\sigma_1, \sigma_2)\},$$

where $c(\mu)$ is given by (17),

$$\Gamma(\sigma_1, \sigma_2) = \{\mu \in C(X \times Y)^*_+ : \pi_1\mu = \sigma_1, \pi_2\mu = \sigma_2\}.$$

We consider the set of real-valued functions on X,

$$L := \{-c(\cdot, y) : y \in \operatorname{spt}\sigma_2\}. \tag{46}$$

where the symbol spt stands for the support of the corresponding measure. Every $\mu \in \Gamma(\sigma_1, \sigma_2)$ can be associated with the multifunction $F_\mu : X \to L$,

[7]See also [4, 12, 56, 61, 63], where close abstract results related to c-cyclic monotonicity and c-subdifferentials of c-convex functions (c-cyclic antimonotonicity and c-superdifferentials of c-concave functions) may be found.

$$F_\mu(x) := \{-c(\cdot, y) : (x, y) \in \operatorname{spt}\mu\}. \tag{47}$$

(F_μ is well-defined because the projection of (a compact set) $\operatorname{spt}\mu$ onto Y is exactly $\operatorname{spt}\sigma_2$.)

Note that for $F = F_\mu$ function (43) takes the form

$$\varphi_{F_\mu}(z_1, z_2) = \inf_{y:(z_1,y)\in\operatorname{spt}\mu}(c(z_2, y) - c(z_1, y)). \tag{48}$$

Furthermore, since c is continuous and $\operatorname{spt}\mu$ is compact, infimum in (48) is attained whenever $z_1 \in Z = \operatorname{dom} F_\mu = \pi_1(\operatorname{spt}\mu) = \operatorname{spt}\sigma_1$, and the function φ_{F_μ} is continuous and vanishes on the diagonal in $Z\times Z$; therefore, $Q_0(\varphi_{F_\mu}) = Q(\varphi_{F_\mu})$.

Theorem 7. (cf. [40, Theorem 5.1] and [44, Theorem 2.1]). *Given a measure $\mu \in \Gamma(\sigma_1, \sigma_2)$, the following statements are equivalent:*

(a) μ is an optimal solution to the MKP, that is $c(\mu) = \mathcal{C}(c; \sigma_1, \sigma_2)$;

(b) the set $Q_0(\varphi_{F_\mu}) = Q(\varphi_{F_\mu})$ is nonempty;

(c) F_μ is L-cyclic monotone.

Proof. $(a) \Rightarrow (b)$ By Proposition 10, there is an optimal solution (u, v) to the dual MKP; therefore,

$$\mathcal{D}(c; \sigma_1, \sigma_2) = \int_X u(x)\,\sigma_1(dx) - \int_Y v(y)\,\sigma_2(dy), \tag{49}$$

and taking into account the duality relation $\mathcal{C}(c; \sigma_1, \sigma_2) = \mathcal{D}(c; \sigma_1, \sigma_2)$ (see Theorem 1), (49) can be rewritten as

$$c(\mu) = \int_X u(x)\,\sigma_1(dx) - \int_Y v(y)\,\sigma_2(dy). \tag{50}$$

Furthermore, since $\pi_1\mu = \sigma_1$, $\pi_2\mu = \sigma_2$, and $(u, v) \in Q'(c)$, (50) implies

$$u(x) - v(y) = c(x, y) \text{ whenever } (x, y) \in \operatorname{spt}\mu. \tag{51}$$

Note that $\pi_1(\operatorname{spt}\mu)$ is closed as the projection of a compact set; therefore, $Z = \operatorname{dom} F_\mu = \pi_1(\operatorname{spt}\mu) = \operatorname{spt}\sigma_1$, and (51) means

$$u(z) - v(y) = c(z, y) \text{ whenever } z \in Z,\ l = -c(\cdot, y) \in F_\mu(z). \tag{52}$$

Now, given any $z, z' \in Z$, and $l \in F_\mu(z) = \{-c(\cdot, y) : (z, y) \in \operatorname{spt}\mu\}$, we derive from (52) $u(z') - u(z) = u(z') - c(z, y) - v(y) \le c(z', y) - c(z, y)$, and taking infimum over all y with $(z, y) \in \operatorname{spt}\mu$, yields $u(z') - u(z) \le \varphi_{F_\mu}(z, z')$ hence $(-u) \in Q(\varphi_{F_\mu})$.

$(b) \Rightarrow (a)$ Since every measure from $\Gamma(\sigma_1, \sigma_2)$ vanishes outside the set $\operatorname{spt}\sigma_1 \times \operatorname{spt}\sigma_2$, one can consider μ as a measure on $X_\mu \times Y_\mu$ (instead of $X \times Y$), where $X_\mu = \operatorname{spt}\sigma_1$, $Y_\mu = \operatorname{spt}\sigma_2$. It suffices to show that μ is an optimal solution to the MKP on $X_\mu \times Y_\mu$.

Note that $u \in Q(\varphi_{F_\mu})$ means

$$u(z_1) - u(z_2) \le c(z_2, y) - c(z_1, y) \tag{53}$$

whenever $(z_1, y) \in \operatorname{spt}\mu$. Let us define

$$v(y) := - \inf_{z:(z,y)\in \operatorname{spt}\mu} (u(z) + c(z, y)),\ y \in Y_\mu. \tag{54}$$

Since $\operatorname{spt}\mu$ is compact and u, c are continuous, the infimum in the right-hand side of (54) is attained and v proves to be a bounded lsc function on Y_μ. Moreover, it follows from (53) that

$$-u(z) - v(y) \le c(z, y) \quad \forall (z, y) \in X_\mu \times Y_\mu \tag{55}$$

and

$$-u(z) - v(y) = c(z, y) \quad \forall (z, y) \in \operatorname{spt}\mu. \tag{56}$$

Note now that (56) implies

$$\int_{X_\mu} (-u)(x)\, \sigma_1(dx) - \int_{Y_\mu} v(y)\, \sigma_2(dy) = c(\mu). \tag{57}$$

We derive from (57) that $\mathcal{D}'(c; \sigma_1, \sigma_2) \ge c(\mu) \ge \mathcal{C}(c; \sigma_1, \sigma_2)$, and as always $\mathcal{D}'(c; \sigma_1, \sigma_2) \le \mathcal{C}(c; \sigma_1, \sigma_2)$ (see (20)), μ is optimal.

$(b) \Leftrightarrow (c)$ This is a particular case of Theorem 4. □

Remark 10. A different proof of a similar theorem is given in [40, Theorem 5.1] and [44, Theorem 2.1], where non-compact spaces are considered. A close result saying that optimality of μ and c-cyclic antimonotonicity of $\operatorname{spt}\mu$ are equivalent may be found in [12].

We now turn to the Monge problem. Recall (see Introduction) that it is to minimize the functional

$$\mathcal{F}(f) := \int_X c(x, f(x))\, \sigma_1(dx)$$

over the set $\varPhi(\sigma_1, \sigma_2)$ of measure-preserving Borel maps $f : (X, \sigma_1) \to (Y, \sigma_2)$. (A map f is called measure-preserving if $f(\sigma_1) = \sigma_2$, that is $\sigma_1 f^{-1}(B_Y) = \sigma_2 B_Y$ for every Borel set $B_Y \subseteq Y$.) Any $f \in \varPhi(\sigma_1, \sigma_2)$ is associated with a measure $\mu_f = (\mathrm{id}_X \times f)(\sigma_1) \in C(X \times Y)^*_+$, as given by

$$\int_{X\times Y} w(x, y)\, \mu_f(d(x, y)) := \int_X w(x, f(x))\, \sigma_1(dx) \quad \forall w \in C(X \times Y),$$

or, equivalently,

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

whenever $B \subseteq (X \times Y)$ is Borel. It is easily seen that $\mu_f \in \varGamma(\sigma_1, \sigma_2)$ and

$$\mathcal{F}(f) = c(\mu_f). \tag{58}$$

The measure μ_f is called a (feasible) Monge solution to MKP. It follows from (58) that if there is an optimal solution to MKP which is the Monge solution μ_f, then f is an optimal solution to the Monge problem and optimal values of both problems coincide,

$$\mathcal{C}(c; \sigma_1, \sigma_2) = \mathcal{F}(f) = \mathcal{V}(c; \sigma_1, \sigma_2), \tag{59}$$

where $\mathcal{V}(c; \sigma_1, \sigma_2) = \inf\{\mathcal{F}(f) : f \in \Phi(\sigma_1, \sigma_2)\}$. In general case, $\mathcal{C}(c; \sigma_1, \sigma_2) \leq \mathcal{V}(c; \sigma_1, \sigma_2)$, and $\Phi(\sigma_1, \sigma_2)$ can be empty; however, in some particular cases (59) holds true.

Remark 11. When X and Y are subsets in $\mathbb{R}^n$, some existence (and uniqueness) results for optimal Monge solutions based on conditions of c-cyclic monotonicity (antimonotonicity) may be found in [3], [5], [6], [12], [39], [40], [55], [65]. In most of these publications, cost functions of the form $c(x, y) = \varphi(x - y)$ are considered. (Note that, since a pioneer paper by Sudakov [64],[8] much attention is paid to cost functions $c(x, y) = \|x - y\|$ for various norms $\|\cdot\|$ in $\mathbb{R}^n$; for such cost functions the optimal solution is not unique.) Several existence and uniqueness theorems for general cost functions are established in [39, 40].

Notice that for a continuous $f \in \Phi(\sigma_1, \sigma_2)$ and $\mu = \mu_f$ one has $\operatorname{spt} \mu = \{(z, f(z)) : z \in \operatorname{spt} \sigma_1\}$; therefore, F_μ as given by (47) is single-valued, $F_\mu(x) = -c(\cdot, f(x))$, and

$$\varphi_{F_\mu}(z_1, z_2) = \varphi_f(z_1, z_2) := c(z_2, f(z_1)) - c(z_1, f(z_1)).$$

The next optimality criterion is then a direct consequence of Theorem 7.

Corollary 7. (cf. [44, Corollary 2.2]). *Suppose $f \in \Phi(\sigma_1, \sigma_2)$ is continuous, then μ_f is an optimal solution to MKP if and only if $Q(\varphi_f)$ is nonempty.*

Remark 12. If $f \in \Phi(\sigma_1, \sigma_2)$ is discontinuous, then the support of μ_f is the closure of the set $\{(z, f(z)) : z \in \operatorname{spt} \sigma_1\}$. In some cases, Corollary 7 and its generalizations following from Theorem 7 enable to find exact optimal solutions to concrete Monge problems; see [44, 45].

3.5 Some Generalizations

In this subsection we consider briefly some examples of H-convex functions similar to Example 2 and some sets of type $Q(c)$ and $Q_0(c)$ for cost functions c that can fail to vanish on the diagonal.

Given an arbitrary infinite set X, $l^\infty(X)$ and $l^\infty(X \times X)$ will denote the linear spaces of bounded real-valued functions on X and $X \times X$, respectively. They are dual Banach spaces relative to the uniform norms

[8] In spite of a gap in Sudakov's proof (see [3]), its main idea proves to be fruitful.

$\|u\|_\infty = \sup_{x\in X} |u(x)|$, $u \in l^\infty(X)$ and $\|w\|_\infty = \sup_{(x,y)\in X\times X} |w(x,y)|$, $w \in l^\infty(X\times X)$:

$$l^\infty(X) = l^1(X)^*, \quad l^\infty(X\times X) = l^1(X\times X)^*.$$

Here, $l^1(Z)$ stands for the space of real-valued functions v on Z with at most countable set $\operatorname{spt} v := \{z \in Z : v(z) \neq 0\}$ and $\|v\|_1 := \sum_{z\in \operatorname{spt} v} |v(z)| < \infty$, and the duality between $l^1(Z)$ and $l^\infty(Z)$ is given by the bilinear form

$$\langle v, u\rangle := \sum_{z\in \operatorname{spt} v} v(z)u(z),\ u \in l^\infty(Z),\ v \in l^1(Z).$$

Given a cost function $c : X \times X \to \mathbb{R} \cup \{+\infty\}$, the reduced cost function c_* is defined as follows:

$$c_*(x,y) := \min\left(c(x,y),\ \inf_n \inf_{x_1,\dots,x_n} \sum_{i=1}^{n+1} c(x_{i-1},x_i)\right), \tag{60}$$

where $x_0 = x, x_{n+1} = y$. Clearly, it turns into (31) when c vanishes on the diagonal. Also, c_* satisfies the triangle inequality $c_*(x,y) + c_*(y,z) \geq c_*(x,z)$ for all $x, y, z \in X$ if one takes, by definition, that $(+\infty) + (-\infty) = (-\infty) + (+\infty) = +\infty$.

Let us define a set

$$Q(c; l^\infty(X)) := \{u \in l^\infty(X) : u(x) - u(y) \leq c(x,y) \quad \forall x, y \in X\}. \tag{61}$$

Note that if $u \in Q(c; l^\infty(X))$, then

$$u(x_{i-1}) - u(x_i) \leq c(x_{i-1}, x_i), \quad i = 1, \dots, n+1,$$

and summing up these inequalities with $x_0 = x$, $x_{n+1} = y$ yields

$$u(x) - u(y) \leq \sum_{i=1}^{n+1}(u(x_{i-1}) - u(x_i)) \leq \sum_{i=1}^{n+1} c(x_{i-1}, x_i).$$

This implies $u \in Q(c_*; l^\infty(X))$, and as $c \geq c_*$, it follows that $Q(c; l^\infty(X)) = Q(c_*; l^\infty(X))$.

Proposition 11. (cf. [33, Lemma 2] and [37, Theorem 4.1].) *Suppose c_* is bounded from above, the following statements are then equivalent:*

(a) $Q(c; l^\infty(X)) \neq \varnothing$*;*

(b) $c_* \in l^\infty(X \times X)$*;*

(c) $c_*(x,y) > -\infty$ *for all* $x, y \in X$*;*

(d) $c_*(x,x) \geq 0$ *for all* $x \in X$*;*

(e) for all integers l and all cycles $x_0, \dots, x_{l-1}, x_l = x_0$ in X, the inequality holds $\sum_{i=1}^{l} c(x_{i-1}, x_i) \geq 0$;

(f) the function

$$\bar{c}(x,y) = \begin{cases} c_*(x,y), & \text{if } x \neq y; \\ 0, & \text{if } x = y; \end{cases} \tag{62}$$

is H-convex relative to $H := \{h_u(x,y) = u(x) - u(y) : u \in l^\infty(X)\}$.

Proof. $(a) \Rightarrow (b)$ Suppose that $u \in Q(c; l^\infty(X))$. Since $Q(c; l^\infty(X)) = Q(c_*; l^\infty(X))$, one has $u(x) - u(y) \leq c_*(x, y)$; therefore, c_* is bounded from below, and as, by hypothesis, c_* is bounded from above, $c_* \in l^\infty(X \times X)$.

$(b) \Rightarrow (a)$ Fix arbitrarily a point $x_0 \in X$ and set $u(x) := c_*(x, x_0)$. Clearly, $u \in l^\infty(X)$, and by the triangle inequality, $u(x) - u(y) = c_*(x, x_0) - c_*(y, x_0) \leq c_*(x, y)$ whenever $x, y \in X$, i.e., $u \in Q(c_*; l^\infty(X) = Q(c; l^\infty(X))$.

$(b) \Rightarrow (c)$ Obvious.

$(c) \Rightarrow (b)$ Since c_* is bounded from above, one has $c_*(x, y) < M < +\infty$ for all $(x, y) \in X \times X$. Suppose $c_* \notin l^\infty(X)$, then there are points $(x_n, y_n) \in X \times X$ such that $c_*(x_n, y_n) < -n$, and applying the triangle inequality yields $c_*(x, y) \leq c_*(x, x_n) + c_*(x_n, y_n) + c_*(y_n, y) \leq 2M - n$; therefore $c_*(x, y) = -\infty$.

$(c) \Rightarrow (d)$ It follows from the triangle inequality that $c_*(x, x) \leq 2c_*(x, x)$ whenever $x \in X$. Therefore, if $c_*(x_0, x_0) < 0$ for some $x_0 \in X$, then $c_*(x_0, x_0) = -\infty$. (Moreover, in such a case, applying again the triangle inequality yields $c_*(x, y) \leq c_*(x, x_0) + c_*(x_0, x_0) + c_*(x_0, y) = -\infty$.)

$(d) \Rightarrow (c)$ Suppose $c_*(x, y) = -\infty$ for some $(x, y) \in X \times X$, then applying the triangle inequality yields $c_*(x, x) \leq c_*(x, y) + c_*(y, x) = -\infty$.

$(d) \Leftrightarrow (e)$ Obvious.

$(b) \Rightarrow (f)$ Take a point $x_0 \in X$ and define $u_{x_0}(x) := c_*(x, x_0)$. One has $h_{u_{x_0}}(x, y) = c_*(x, x_0) - c_*(y, x_0) \leq c_*(x, y) = \bar{c}(x, y)$ for any $x \neq y$, and $h_{u_{x_0}}(x, x) = 0 = \bar{c}(x, x)$ for all $x \in X$. Thus, $h_{u_{x_0}} \in H(\bar{c})$ whenever $x_0 \in X$. Moreover, for $x_0 = y$ one gets $h_{u_y}(x, y) = \bar{c}(x, y)$, and H-convexity of $\bar{c}$ is thus established.

$(f) \Rightarrow (a)$ Obvious. □

Remark 13. It is easily seen that $Q(c; l^\infty(X)) = Q(\bar{c}; l^\infty(X))$ and $H(\bar{c}) = \{h_u : u \in Q(c; l^\infty(X))\}$.

Remark 14. It follows easily from the proof of Proposition 11 that if c_* is bounded from above then either all the statements $(a) - (f)$ hold true or $c_*(x, y) = -\infty$ whenever $(x, y) \in X \times X$.

The following proposition is established by similar arguments, and so we omit its proof.

Proposition 12. (cf. [32, 35]). *Given a function $c : X \times X \to \mathbb{R} \cup \{+\infty\}$ such that $c_*(x, y) < +\infty$ whenever $x, y \in X$, the following statements are equivalent:*

(a) $Q_0(c) := \{u \in \mathbb{R}^X : u(x) - u(y) \leq c(x, y) \quad \forall x, y \in X\} \neq \varnothing$;

(b) $c_*(x, y) > -\infty$ *for all* $x, y \in X$;

(c) $c_*(x, x) \geq 0$ *for every* $x \in X$;

(d) for all integers l *and all cycles* $x_0, \dots, x_{l-1}, x_l = x_0$ *in* X*, the inequality holds* $\sum_{i=1}^{l} c(x_{i-1}, x_i) \geq 0$.

(e) the function $\bar{c}$*, as given by (62), is* H*-convex with respect to* $H = \{h_u(x, y) = u(x) - u(y) : u \in \mathbb{R}^X\}$.

Remark 15. It is easily seen that $H(\bar{c}) = \{h_u : u \in Q_0(c)\}$.

Remark 16. If X is a domain in $\mathbb{R}^n$ and c is a smooth function vanishing on the diagonal, then either $Q_0(c)$ is empty or $Q_0(c) = \{u + const\}$ where $\nabla u(x) = -\nabla_y c(x,y)|_{y=x}$; see [33, 35, 36]. Second-order conditions (necessary ones and sufficient ones) for $Q_0(c)$ to be nonempty are given in [33, 35, 36, 46].

Let $E(X)$ be a closed linear subspace in $l^\infty(X)$ containing constant functions, separating points of X (that is, for any $x, y \in X$ there is a function $u \in E(X)$, $u(x) \neq u(y)$), and such that $u, v \in E(X)$ implies $uv \in E(X)$. Then $E(X)$ is a (commutative) Banach algebra with respect to the uniform norm $\|u\| = \sup\limits_{x\in X} |u(x)|$ and the natural (pointwise) multiplication. (Also, $E(X)$ is a Banach lattice; see [43].) As is known from theory of Banach algebras [13, 51], the set $\varkappa X$ of all non-zero multiplicative linear functionals on $E(X)$ is a weak* compact subset in $E(X)^*$, X is dense in $\varkappa X$ [9], and an isometry of Banach algebras (Gelfand's representation), $A : E(X) \to C(\varkappa X)$, $AE(X) = C(\varkappa X)$, holds as follows:

$$Au(\delta) := \langle u, \delta \rangle, \ u \in E(X), \delta \in \varkappa X.$$

Let us give three examples of Banach algebras $E(X)$. They are as follows:

1. $C^b(X)$ - the Banach algebra of bounded continuous real-valued functions on a completely regular Hausdorff topological space X. (In this case, $\varkappa X = \beta X$ is the Stone-Čech compactification of X.)

2. $U(X)$ - the Banach algebra of bounded universally measurable real-valued functions on a compact Hausdorff topological space X (we have yet met it in subsection 3.2).

3. $\mathcal{L}^\infty(\mathbb{R}^n)$ - the Banach algebra of bounded Lebesgue measurable real-valued functions on $\mathbb{R}^n$ (Lebesgue equivalent functions are not identified). This algebra will be of use in section 5.

Given a set X, a cost function $c : X \times X \to \mathbb{R} \cup \{+\infty\}$, and an algebra $E(X)$, one can define the set

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

and a class of functions on $X \times X$,

$$H := \{h_u : \ h_u(x,y) = u(x) - u(y), u \in E(X)\}.$$

Clearly, $H(c) = Q(c; E(X))$ and c is H-convex if and only if

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

whenever $x, y \in X$.

Moreover, $Q(c; E(X))$ proves to be the constraint set for an abstract (non-topological) variant of the dual MKP with a given marginal difference, and H-convexity arguments play important role in the corresponding duality results; see [37, 38] for details.

---

[9] A point $x \in X$ is identified with the functional $\delta_x \in \varkappa X$, $\langle u, \delta_x \rangle = u(x)$, $u \in E(X)$.

Similarly, given a cost function $c : X \times Y \to \mathbb{R} \cup \{+\infty\}$, one can take two algebras, $E_1(X)$ and $E_2(Y)$, and consider the set in their product,

$$Q'(c; E_1(X), E_2(Y)) := \{(u, v) : u(x) - v(y) \le c(x, y) \quad \forall x \in X, y \in Y\},$$

and a class of functions on $X \times Y$,

$$H := \{h_{uv} : h_{uv}(x, y) = u(x) - v(y),\ (u, v) \in E_1(X) \times E_2(Y)\}.$$

Clearly, $H(c) = Q'(c; E_1(X), E_2(Y))$ and c is H-convex if and only if

$$c(x, y) = \sup\{u(x) - v(y) : (u, v) \in Q'(c; E_1(X), E_2(Y))\}$$

whenever $(x, y) \in X \times Y$.

Moreover, $Q'(c; E_1(X), E_2(Y))$ is the constraint set for an abstract variant of the dual MKP with given marginals, and H-convexity arguments play important role in the corresponding duality results; see [38].

4 Applications to Mathematical Economics

In this section, we present briefly several applications to mathematical economics. In all the applications, properties of the sets $Q(c)$ and $Q_0(c)$ for various particular cost functions c are considered. The corresponding results are based on conditions for these sets to be nonempty.

4.1 Utility Theory

A *preorder* on a set X is a binary relation $\preceq$ which is reflexive ($x \preceq x$ for all $x \in X$) and transitive (for any $x, y, z \in X$, $x \preceq y$, $y \preceq z$ imply $x \preceq z$). A preorder $\preceq$ is called *total* if any two elements of X, x and y, are compatible, that is $x \preceq y$ or $y \preceq x$. A preorder $\preceq$ on a topological space X is called *closed* if its graph, $gr(\preceq) := \{(x, y) : x \preceq y\}$, is a closed subset in $X \times X$.

Any preorder $\preceq$ can be treated as a preference relation, and it determines two binary relations on X: the strict preference relation $\prec$,

$$x \prec y \Longleftrightarrow x \preceq y \text{ but not } y \preceq x,$$

and the equivalence relation $\sim$,

$$x \sim y \Longleftrightarrow x \preceq y \text{ and } y \preceq x.$$

A real-valued function u on X is said to be an *utility function* for a preorder $\preceq$ if for any $x, y \in X$ two conditions are satisfied as follows:

$$x \preceq y \Rightarrow u(x) \le u(y), \tag{63}$$

$$x \prec y \Rightarrow u(x) < u(y). \tag{64}$$

Clearly, it follows from (63) that $x \sim y \Rightarrow u(x) = u(y)$.

The pair of conditions (63),(64) is equivalent to the single condition

$$x \preceq y \Leftrightarrow u(x) \leq u(y)$$

if and only if the preorder $\preceq$ is total. (Moreover, if $\preceq$ is total, then $x \prec y \Leftrightarrow u(x) < u(y)$ and $x \sim y \Leftrightarrow u(x) = u(y)$, that is, the preference relation is completely determined by its utility function.)

One of fundamental results in the mathematical utility theory is the famous theorem due to Debreu [9, 10], which asserts the existence of a continuous utility function for every total closed preorder on a separable metrizable space. We'll give here (see also [27, 28, 32]) some extensions of that theorem to the case where the preorder is not assumed to be total. The idea of our approach is to use a specific cost function c that vanishes on the graph of the preorder and has appropriate semicontinuity properties. With help of the duality theorem (Theorem 2) we'll obtain a representation

$$gr(\preceq) = \{(x, y) : u(x) \leq u(y) \quad \forall u \in H\} \tag{65}$$

with $H \subseteq Q(c)$. Moreover, sometimes it is possible to choose a countable $H = \{u_k : k = 1, 2, \dots\}$, and in such a case

$$u_0(x) = \sum_{k=1}^{\infty} 2^{-k} \frac{u_k(x)}{1 + |u_k(x)|}$$

proves to be a continuous utility function for $\preceq$.

Theorem 8. ([22, 27]). *Let $\preceq$ be a closed preorder on a compact metrizable space X. Then $gr(\preceq)$ has a representation (65) with a countable H; hence there is a continuous utility function for $\preceq$.*

Proof.[10] Consider on $X \times X$ the cost function

$$c(x, y) = \begin{cases} 0, & \text{if } x \preceq y; \\ +\infty, & \text{otherwise.} \end{cases}$$

It satisfies the triangle inequality and vanishes on the diagonal because $\preceq$ is transitive and reflexive. Also, it is lsc because $\preceq$ is closed. It follows from Theorem 2 that $Q(c)$ is nonempty and

$$c(x, y) = \sup_{u \in Q(c)} (u(x) - u(y));$$

therefore,

$$gr(\preceq) = \{(x, y) : u(x) \leq u(y) \quad \forall u \in Q(c)\}.$$

Since $C(X)$ is separable, one can choose a dense countable subset H in $Q(c)$. Then (65) holds with that H, and the result follows. □

The next result is derived from Theorem 8.

[10] This proof follows [27]; a proof in [22] is different.

Corollary 8.. ([28, 32]). *Theorem 8 is extended to X being a separable metrizable locally compact space.*

Theorem 9. ([31]). *Let $\preceq$ be a preorder on a separable metrizable space X, the following statements are then equivalent:*

(a) a representation (65) holds with a countable family $H \subset C^b(X)$;

(b) $\preceq$ is a restriction to X of a closed preorder $\preceq_1$ on X_1, where X_1 is a metrizable compactification of X.

If these equivalent statements hold true, then there is a continuous utility function for $\preceq$.

We consider now the following question. Given a closed preorder $\preceq_\omega$ depending on a parameter ω, when is there a *continuous utility*, i.e. a jointly continuous real-valued function $u(\omega, x)$ such that, for every ω, $u(\omega, \cdot)$ is a utility function for $\preceq_\omega$? This question arises in various parts of mathematical economics. In case of total preorders $\preceq_\omega$, some *sufficient* conditions for the existence of a continuous utility were obtained in [8, 48, 50, 52]. The corresponding existence results are rather special consequences of the following general theorem.

Theorem 10. ([28, 32]). *Suppose that Ω and X are metrizable topological spaces, and X, in addition, is separable locally compact. Suppose also that for every $\omega \in \Omega$ a preorder $\preceq_\omega$ is given on X, and that the set $\{(\omega, x, y) : x \preceq_\omega y\}$ is closed in $\Omega \times X \times X$. Then there exists a continuous utility $u : \Omega \times X \to [0, 1]$.*

Proof (the case where Ω is separable locally compact).[11] Let us define a preorder $\preceq$ on $\Omega \times X$,

$$(\omega_1, x_1) \preceq (\omega_2, x_2) \Longleftrightarrow \omega_1 = \omega_2, x_1 \preceq_{\omega_1} x_2.$$

It is obviously closed, and as $\Omega \times X$ is separable locally compact, the result follows from Corollary 8. □

Remark 17. Observe that if all $\preceq_\omega$ are total then the condition that the set $\{(\omega, x, y) : x \preceq_\omega y\}$ is closed in $\Omega \times X \times X$ is *necessary* (as well as sufficient) for the existence of a continuous utility $u : \Omega \times X \to [0, 1]$.

Let $\mathcal{P}$ denote the set of all closed preorders on X. By identifying a preorder $\preceq \in \mathcal{P}$ with its graph in $X \times X$, we consider in $\mathcal{P}$ the topology t which is induced by the exponential topology on the space of closed subsets in the one-point compactification of $X \times X$ (for the definition and properties of the exponential topology, see [23]). Obviously, $(\mathcal{P}, t)$ is a metrizable space. The next result is obtained by applying Theorem 10 to $\Omega = (\mathcal{P}, t)$.

Corollary 9. (Universal Utility Theorem [28, 32]). *There exists a continuous function $u : (\mathcal{P}, t) \times X \to [0, 1]$ such that $u(\preceq, \cdot)$ is a utility function for $\preceq$ whenever $\preceq \in \mathcal{P}$.*

[11] For the sake of simplicity, we restrict ourselves to the case where Ω is separable and locally compact. In the general case, the proof makes substantial use of a version of Michael's continuous selection theorem in a locally convex Fréchet space.

4.2 Demand Analysis

Given a *price set* $P \subseteq \operatorname{int}\mathbb{R}^n_+$, we mean by a *demand function* any map $f : P \to \operatorname{int}\mathbb{R}^n_+$. We will say that an utility function $U : \mathbb{R}^n_+ \to \mathbb{R}_+$ *rationalizes* f if, for every $p \in P$,

$$q \in \mathbb{R}^n_+,\ p \cdot q \le p \cdot f(p) \Longrightarrow U(f(p)) \ge U(q). \tag{66}$$

Theorem 11. (cf. [46, Corollary 3]). *Given a function $f : P \to \operatorname{int}\mathbb{R}^n_+$, the following statements are equivalent:*

(a) there is a positive homogeneous utility function $U : \mathbb{R}^n_+ \to \mathbb{R}_+$, which is strictly positive on $f(P)$ and rationalizes f;

(b) there is a positive homogeneous continuous concave utility function $U : \mathbb{R}^n_+ \to \mathbb{R}_+$, which is strictly positive on $f(P)$ and rationalizes f;

(c) for a cost function ξ on $P \times P$, as given by

$$\xi(p, p') := \ln(p' \cdot f(p)) - \ln(p' \cdot f(p')),$$

the set $Q_0(\xi)$ is nonempty;

(d) for every cycle $p^1, \dots, p^l, p^{l+1} = p^1$ in P, the inequality holds true

$$\prod_{k=1}^{l} p^{k+1} \cdot f(p^k) \ge \prod_{k=1}^{l} p^k \cdot f(p^k);$$

(e) there is a strictly positive solution to the system

$$u(p) \ge \frac{p \cdot f(p)}{p \cdot f(p')} u(p') \quad \textit{for all} \quad p, p' \in P. \tag{67}$$

Proof. $(b) \Rightarrow (a)$ Obvious.

$(a) \Rightarrow (e)$ Define $u(p) := U(f(p))$, $p \in P$. Since for every $q \in \mathbb{R}^n_+$,

$$p \cdot \frac{p \cdot f(p)}{p \cdot q} q = p \cdot f(p),$$

it follows from (66) that

$$\frac{p \cdot f(p)}{p \cdot q} U(q) = U\left(\frac{p \cdot f(p)}{p \cdot q} q\right) \le U(f(p)),$$

which implies (67) for $q = f(p')$.

$(e) \Rightarrow (c)$ Since a solution $u(p)$ to (67) is strictly positive, it follows that $v(p) := \ln u(p)$ makes sense and belongs to $Q_0(\xi)$.

$(c) \Leftrightarrow (d)$ This is an easy consequence of Proposition 12.

$(c) \Rightarrow (e)$ Suppose $v \in Q_0(\xi)$, then $u(p) = e^{v(p)}$ is strictly positive and satisfies (67).

$(e) \Rightarrow (b)$ Let us define

$$U(q) := \inf_{p' \in P} \frac{u(p')}{p' \cdot f(p')} p' \cdot q.$$

It follows easily from (67) that $U(f(p)) = u(p)$ whenever $p \in P$, hence U is strictly positive on $f(P)$. If now $p \cdot q \leq p \cdot f(p)$, then

$$U(q) \leq \frac{u(p)}{p \cdot f(p)} p \cdot q \leq u(p) = U(f(p)),$$

that is U rationalizes f. Since U is clearly upper semi-continuous concave (hence continuous; see [57, Theorem 10.2]) and positive homogeneous, the implication is completely established. □

Remark 18. Statement (d) can be considered as a particular (strengthened) version of the strong revealed preference axiom, and (e) generalizes the corresponding variant of the Afriat—Varian theory (see [1, 2, 66, 67]) to the case of infinite set of ‘observed data’. Further results on conditions for rationalizing demand functions by concave utility functions with nice additional properties in terms of non-emptiness of sets $Q_0(\varphi)$ for various price sets P and some specific cost functions φ on $P \times P$ may be found in [46].

4.3 Dynamics Models

In this subsection (see also [35, 36, 37]), we consider an abstract dynamic optimization problem resembling, in some respects, models of economic system development.

Suppose X is an arbitrary set and $a : X \to X$ is a multifunction with nonempty values. Its graph, $gr(a) = \{(x, y) : y \in a(x)\}$, may be considered as a continual net with vertices $x \in X$ and arcs $(x, y) \in gr(a)$, respectively. A finite sequence of elements of X, $\chi = (\chi(t))_{t=0}^{T}$ (where $T = T(\chi) < +\infty$ depends on χ), satisfying

$$\chi(t) \in a(\chi(t-1)), \quad t = 1, \dots, T,$$

is called a (finite) *trajectory*. We assume that the *connectivity hypothesis* is satisfied: for any $x, y \in X$, there is a trajectory χ that starts at x ($\chi(0) = x$) and finishes at y ($\chi(T) = y$).

Given a terminal function $l : X \times X \to \mathbb{R} \cup \{+\infty\}$ with $\operatorname{dom} l \neq \varnothing$ and a cost function $c : X \times X \to \mathbb{R} \cup \{+\infty\}$ with $\operatorname{dom} c = gr(a)$, the payment for moving along the trajectory χ equals

$$g(\chi) := l(\chi(0), \chi(T)) + \sum_{t=1}^{T} c(\chi(t-1), \chi(t)).$$

The problem is to minimize $g(\chi)$ over the set τ of all trajectories.

Observe that the connectivity hypothesis can be rewritten as

$$c_*(x,y) < +\infty \quad \text{for all} \quad x, y \in X,$$

and the optimality of a trajectory $\bar{\chi}$ means exactly

$$g(\bar{\chi}) = \min\{l(x,y) + c_*(x,y) : x, y \in X\}. \tag{68}$$

Theorem 12. ([35, Theorem 7.1]). *Suppose X is a compact topological space, both functions on $X \times X$, l and c, are lsc, and $c(x,y) > 0$ for all $x, y \in X$. Then there exists an optimal trajectory.*

An important particular case of the above problem is to minimize the functional

$$g_1(\chi) := \sum_{t=1}^{T(\chi)} c(\chi(t-1), \chi(t))$$

over the set $\tau(X_1, X_2)$ of trajectories that start in X_1 and finish in X_2 (i.e., $\chi(0) \in X_1$, $\chi(T) \in X_2$), where X_1 and X_2 are given subsets of X. This problem is reduced to minimizing $g(\chi)$ over τ if one takes l to be the indicator function of $X_1 \times X_2$ (i.e., $l(x,y) = 0$ for $(x,y) \in X_1 \times X_2$ and $l(x,y) = +\infty$ otherwise).

The next result is a direct consequence of Theorem 12.

Corollary 10. ([35, Corollary 7.1]). *Let X and c be as in Theorem 12, and suppose that X_1 and X_2 are closed in X. Then there exists a trajectory $\bar{\chi} \in \tau(X_1, X_2)$ minimizing g_1 over $\tau(X_1, X_2)$.*

We now return to the general (non-topological) version of the problem.

Theorem 13. ([35, Theorem 7.2]). *A trajectory $\bar{\chi} = (\bar{\chi}(t))_{t=0}^{T}$ is optimal in τ if and only if: (a) the equality holds*

$$l(\bar{\chi}(0), \bar{\chi}(T)) + c_*(\bar{\chi}(0), \bar{\chi}(T)) = \min\{l(x,y) + c_*(x,y) : x, y \in X\}, \tag{69}$$

and (b) there is a function $u \in Q_0(c)$ satisfying

$$u(\bar{\chi}(t-1)) - u(\bar{\chi}(t)) = c(\bar{\chi}(t-1), \bar{\chi}(t)),\ t = 1, \ldots, T. \tag{70}$$

An infinite sequence of elements of X, $\chi = (\chi(t))_{t=0}^{\infty}$, satisfying

$$\chi(t) \in a(\chi(t-1)), \quad t = 1, 2, \ldots,$$

is called an *infinite trajectory*. Say an infinite trajectory $\chi = (\chi(t))_{t=0}^{\infty}$ is *efficient* if there exists $T_1 = T_1(\chi) < +\infty$ such that, for every $T \geq T_1$, the finite trajectory $\chi^T := (\chi(t))_{t=0}^{T}$ is optimal in τ.

The next result is derived from Theorem 13 with help of the Banach limit technique; see [35, Theorem 7.4] for details.

Theorem 14. *An infinite trajectory $\chi = (\chi(t))_{t=0}^{\infty}$ is efficient if and only if: (a) (69) holds for all $T \geq T_1$ and (b) there is a function $u \in Q_0(c)$ satisfying (70) for all t.*

4.4 Economics of Corruption

Following [7] (see also [36, 56]), we briefly outline here some kind of principal-agents models relating to economics of corruption and dealing with distorting substantial economic information. Suppose there is a population of agents, each of them is characterized by his state (a variable of economic information), which is an element of some set X, and there is yet one agent called the principal (State, monopoly, social planner, insurance company and so on). The principal pays to an agent some amount of money $u(x)$ which depends on information x about agent's state. It is assumed that the actual state of the agent, y, cannot be observed directly by the principal; therefore, agents have a possibility to misrepresent at some cost[12] the relevant information to the principal. Thus, we assume that an agent can at the cost $c(x, y)$ to misrepresent his real state y into the state x without being detected. In such a case, his income equals $u(x) - c(x, y)$. The cost function c may take the value $+\infty$, which occurs when x is too far from y for falsifying y into x be possible without being detected. Also, it is assumed that $c(y, y) = 0$; therefore, if an agent gives true information to the principal, then his income equals the payoff $u(y)$. If now there is an $x \in X$ such that $u(x) - c(x, y) > u(y)$, then, for an agent with the actual state y, it proves to be profitable to falsify his state information. Say, in a model of collusion with a third party, an agent with the actual state y and a supervisor may agree to report the state x maximizing their total income $u(x) - c(x, y)$ and then to share between them the surplus $u(x) - c(x, y) - u(y) > 0$. Similar situations arise in other models (insurance fraud, corruption in taxation); see [7] for details.

Thus, given a cost function $c : X \times X \to \mathbb{R}_+ \cup \{+\infty\}$ vanishing on the diagonal, a question arises, whether the payoff function $u : X \to \mathbb{R}$ is *non-manipulable* or *collusion-proof* in the sense that it is in the interest of each agent to be honest. The answer is affirmative if and only if $u \in Q_0(c)$.

5 An Application to Approximation Theory

In this section, we deal with some best approximation problems.

Let us consider a linear subspace in $l^\infty(X \times X)$,

$$H_0 := \{u(x) - u(y) : \ u \in l^\infty(X)\}. \tag{71}$$

Given a function $f \in l^\infty(X \times X)$, the problem is to find the value

$$m(f; H_0) := \min_{h \in H_0} \|f - h\|_\infty = \min_{u \in l^\infty(X)} \sup_{x,y \in X} |f(x, y) - u(x) + u(y)|. \tag{72}$$

Note that the minimum in (72) is attained at some $h = h_u \in H_0$, $h_u(x, y) = u(x) - u(y)$, because closed balls in the dual Banach space $l^\infty(X)$

[12] For instance, by colluding with a third party (expert, supervisor, tax officer).

are weak* compact and the functional on $l^\infty(X)$, $u \mapsto \sup_{x,y\in X} |f(x,y) - u(x) + u(y)|$ is weak* lsc. Moreover, (72) can be rewritten as

$$m(f; H_0) = \min\{\alpha > 0 :\ Q(c+\alpha; l^\infty) \neq \varnothing\}, \tag{73}$$

where

$$c(x,y) := \min(f(x,y), -f(y,x)), \quad x, y \in X, \tag{74}$$

and there exists a function u in $Q(c + m(f; H_0); l^\infty)$.

A topological analog of this problem is as follows. Given a completely regular Hausdorff topological space X, a subspace $\mathcal{H}_0$ in $C^b(X \times X)$,

$$\mathcal{H}_0 := \{u(x) - u(y) :\ u \in C^b(X)\}, \tag{75}$$

and a function $f \in C^b(X \times X)$, one has to find the value

$$m(f; \mathcal{H}_0) := \inf_{h\in\mathcal{H}_0} \|f - h\| = \inf_{u\in C^b(X)} \sup_{x,y\in X} |f(x,y) - u(x) + u(y)|. \tag{76}$$

Recall (see subsection 3.5) that, for every topological space X, $C^b(X)$ denotes the space of bounded continuous real-valued functions on it with the uniform norm $\|u\| = \sup_{x\in X} |u(x)|$. Clearly, (76) is equivalent to

$$m(f; \mathcal{H}_0) = \inf\{\alpha > 0 :\ Q(c+\alpha; C^b(X)) \neq \varnothing\}, \tag{77}$$

where c is given by (74).

Theorem 15. ([37, Theorem 5.1]). *For every $f \in l^\infty(X \times X)$,*

$$m(f; H_0) = -\inf \frac{1}{n} \sum_{i=1}^{n} c(x_{i-1}, x_i), \tag{78}$$

and if X is a compact space, then for every $f \in C(X \times X)$,

$$m(f; \mathcal{H}_0) = -\inf \frac{1}{n} \sum_{i=1}^{n} c(x_{i-1}, x_i). \tag{79}$$

Here, both infima, in (78) and (79), are taken over all integers n and all cycles $x_0, \dots, x_{n-1}, x_n = x_0$ in X.

Proof. As follows from (73), for every $\alpha > m(f; H_0)$ there is a function $u \in Q(c+\alpha; l^\infty)$. Then $u(x_{i-1}) - u(x_i) \le c(x_{i-1}, x_i)$, $i = 1, \dots, n$, and summing up these inequalities yields

$$0 = \sum_{i=1}^{n} (u(x_{i-1}) - u(x_i)) \le \sum_{i=1}^{n} (c(x_{i-1}, x_i) + \alpha) = \sum_{i=1}^{n} c(x_{i-1}, x_i) + n\alpha$$

(this follows also from implication $(a) \Rightarrow (e)$ of Proposition 11); therefore, $\alpha \ge -\inf \frac{1}{n} \sum\limits_{i=1}^{n} c(x_{i-1}, x_i)$, and

$$m(f;H_0) = \inf\{\alpha : Q(c+\alpha;l^\infty) \neq \varnothing\} \geq -\inf \frac{1}{n}\sum_{i=1}^{n} c(x_{i-1},x_i). \tag{80}$$

Suppose now that $\alpha < m(f;H_0)$. Then $Q(c+\alpha;l^\infty) = \varnothing$, and taking into account Remark 14, one has $(c+\alpha)_* \equiv -\infty$; therefore, there is a cycle $x_0,\dots,x_{n-1},x_n = x_0$ such that $\sum_{i=1}^{n} c(x_{i-1},x_i) + n\alpha < 0$. One obtains

$$\alpha < -\frac{1}{n}\sum_{i=1}^{n} c(x_{i-1},x_i) \leq -\inf \frac{1}{n}\sum_{i=1}^{n} c(x_{i-1},x_i),$$

and as this holds true whenever $\alpha < m(f;H_0)$, one gets

$$m(f;H_0) \leq -\inf \frac{1}{n}\sum_{i=1}^{n} c(x_{i-1},x_i), \tag{81}$$

and (78) follows from (80),(81).

The proof of (79) is similar if one replaces $Q(c+\alpha;l^\infty)$ with $Q(c+\alpha)$ and takes into account that for every $c \in C(X\times X)$ and every $\alpha \in \mathbb{R}$ either $(c+\alpha)_* \in C(X\times X)$ or $(c+\alpha)_* \equiv -\infty$ (see [47, Lemma 2.4], where a more general result is established). □

Corollary 11. *If X is a compact topological space and $f \in C(X\times X)$, then $m(f;H_0) = m(f;\mathcal{H}_0)$.*

Remark 19. If X is a *non-compact* completely regular Hausdorff topological space, one can pass to its Stone-Čech compactification $X' = \beta X$. Taking into account the natural linear isometry $C^b(X) = C(X')$, the next result is an easy consequence of Theorem 15.

Corollary 12. *Theorem 15 is extended to X being any completely regular Hausdorff topological space provided that $f \in C(X'\times X')$, $C(X)$ in (75) is replaced with $C^b(X)$, and* max *in (76) is replaced with* sup.

Note that $C(\beta X\times\beta X)$ can be considered as the closure in $C^b(X\times X)$ of the subspace of finite sums $f(x,y) = \sum_1^n a_k(x)b_k(y)$, $a_k, b_k \in C^b(X)$, $k = 1,\dots,n$.

Say $u \in C^b(X)$ is an *exact solution* to the approximation problem if the infimum in the right-hand side of (76) is attained at it, that is $m(f;\mathcal{H}_0) = \sup_{x,y\in X} |f(x,y) - u(x) + u(y)|$. It follows from (77) that $u \in C^b(X)$ is an exact solution if and only if it belongs to $Q(c + m(f;\mathcal{H}_0); C^b(X))$; therefore, exact solutions exist if and only if $Q(c + m(f;\mathcal{H}_0); C^b(X))$ is nonempty.

Let $C^{n,\infty}$ be the linear space of bounded infinitely differentiable real-valued functions on $\mathbb{R}^n$, $\mathcal{H}_0^\infty$ a subspace in $C^{2n,\infty}$,

$$\mathcal{H}_0^\infty := \{u(x) - u(y) : u \in C^{n,\infty}\}.$$

Theorem 16. *Suppose* $f(x,y) = g(x-y)$, *where* $g \in C^b(\mathrm{I\!R}^n)$. *Then*

$$m(f;\mathcal{H}_0) = m(f;\mathcal{H}_0^\infty) := \inf_{h\in\mathcal{H}_0^\infty} \|f-h\|,$$

and there is a function $u \in C^{n,\infty}$, *which is an exact solution to the approximation problem:*

$$m(f;\mathcal{H}_0) = m(f;\mathcal{H}_0^\infty) = \|f-h_u\|, \quad h_u(x,y) = u(x) - u(y). \tag{82}$$

To prove the theorem, some facts from the lifting theory [16][13] will be needed. The main of them is the existence of a strong lifting of $\mathcal{L}^\infty(\mathrm{I\!R}^n)$. Recall (see subsection 3.5), that $\mathcal{L}^\infty(\mathrm{I\!R}^n)$ is the space (Banach algebra and Banach lattice) of bounded Lebesgue measurable real-valued functions on $\mathrm{I\!R}^n$ with the uniform norm on it, $\|u\| = \sup_{x\in\mathrm{I\!R}^n}|u(x)|$, $u \in \mathcal{L}^\infty(\mathrm{I\!R}^n)$. A homomorphism of Banach algebras (i.e., a multiplicative linear operator) $\rho : \mathcal{L}^\infty(\mathrm{I\!R}^n) \to \mathcal{L}^\infty(\mathrm{I\!R}^n)$ is said to be a *strong lifting* of $\mathcal{L}^\infty(\mathrm{I\!R}^n)$ if four conditions are satisfied as follows:

1. ρ is a projector, that is $\rho^2 = \rho$;
2. for every $u \in \mathcal{L}^\infty(\mathrm{I\!R}^n)$, $\rho(u) = u$ almost everywhere (a.e.), that is the set $\{x \in \mathrm{I\!R}^n : \rho(u)(x) \neq u(x)\}$ is Lebesgue negligible;
3. for every $u \in \mathcal{L}^\infty(\mathrm{I\!R}^n)$, $u = 0$ a.e. implies $\rho(u) \equiv 0$;
4. $\rho(u) \equiv u$ whenever $u \in C^b(\mathrm{I\!R}^n)$.

It follows from these conditions along with linearity and multiplicativity of ρ that ρ is also a homomorphism of Banach lattices, i.e. $\rho(u \vee v) = \rho(u) \vee \rho(v)$ and $\rho(u \wedge v) = \rho(u) \wedge \rho(v)$ whenever $u, v \in \mathcal{L}^\infty(\mathrm{I\!R}^n)$. Furthermore, $u \geq v$ a.e. implies $\rho(u)(x) \geq \rho(v)(x)$ for all $x \in \mathrm{I\!R}^n$.

The Lebesgue space $L^\infty(\mathrm{I\!R}^n)$ is a Banach algebra and a Banach lattice, and the operator $\pi : \mathcal{L}^\infty(\mathrm{I\!R}^n) \to L^\infty(\mathrm{I\!R}^n)$ mapping every function $u \in \mathcal{L}^\infty(\mathrm{I\!R}^n)$ into its Lebesgue equivalence class is a homomorphism both of Banach algebras and of Banach lattices. Thus, π maps $\mathcal{L}^\infty(\mathrm{I\!R}^n)$ onto the factor space $L^\infty(\mathrm{I\!R}^n) = \mathcal{L}^\infty(\mathrm{I\!R}^n)/\mathcal{N}_0$ where $\mathcal{N}_0$ is the subspace in $\mathcal{L}^\infty(\mathrm{I\!R}^n)$ consisting of Lebesgue negligible functions, and the standard norm in $L^\infty(\mathrm{I\!R}^n)$ is precisely the factor-norm with respect to π. Since $\rho(u) = \rho(v)$ whenever $u - v \in \mathcal{N}_0$, ρ generates a homomorphism of Banach algebras (and of Banach lattices) $\rho' : L^\infty(\mathrm{I\!R}^n) \to \mathcal{L}^\infty(\mathrm{I\!R}^n)$ (a strong lifting of $L^\infty(\mathrm{I\!R}^n)$) such that $\pi \circ \rho' = \mathrm{id}_{L^\infty(\mathrm{I\!R}^n)}$ and $\rho' \circ \pi = \rho$.

Proof of Theorem 16. It follows from (77) that for every k there is a function $u_k \in Q(c+m(f;\mathcal{H}_0)+\frac{1}{k}; C^b(\mathrm{I\!R}^n))$. Fix an arbitrary point x_0 in $\mathrm{I\!R}^n$ and assume without loss of generality that $u_k(x_0) = 0$. Then, for all $x \in \mathrm{I\!R}^n$, one has

$$-c(x_0,x) - m(f;\mathcal{H}_0) - 1 \leq u_k(x) \leq c(x,x_0) + m(f;\mathcal{H}_0) + 1, \quad k = 1,2,\ldots;$$

[13] See also [30, 43], where connections between the lifting theory and abstract convexity are given.

therefore, the sequence (u_k) is bounded in $C^b(\mathbb{R}^n)$. Now, taking into account that $C^b(\mathbb{R}^n)$ is a closed linear subspace in $L^\infty(\mathbb{R}^n)$ and that $L^\infty(\mathbb{R}^n) = L^1(\mathbb{R}^n)^*$ is a dual Banach space, (u_k) is bounded hence weak* precompact in $L^\infty(\mathbb{R}^n)$. We shall assume by passing, if needed, to a subsequence[14] that the sequence (u_k) converges weakly* to an element of $L^\infty(\mathbb{R}^n)$. In other words, there exists a function $v \in \mathcal{L}^\infty(\mathbb{R}^n)$ such that u_k converges weakly* to $\pi(v)$. It follows that the sequence $(u_k(x) - u_k(y)) \subset C(\mathbb{R}^{2n}) \subset L^\infty(\mathbb{R}^{2n})$ converges weakly* in $L^\infty(\mathbb{R}^{2n})$ to the element of $L^\infty(\mathbb{R}^{2n})$, which is the Lebesgue equivalence class of the function $v(x)-v(y)$. Now, as $u_k(x)-u_k(y) \le c(x,y) + m(f;\mathcal{H}_0) + \frac{1}{k}$, and the positive cone $L^\infty_+(\mathbb{R}^{2n})$ is weakly* closed, one gets

$$v(x) - v(y) \le c(x,y) + m(f;\mathcal{H}_0) \quad \text{a.e. in } \mathbb{R}^{2n}. \tag{83}$$

Let us define

$$N(y) := \{x \in \mathbb{R}^n : v(x) - v(y) > c(x,y) + m(f;\mathcal{H}_0)\}, \quad y \in \mathbb{R}^n.$$

It follows from (83) that the set

$$N := \{y \in \mathbb{R}^n : N(y) \text{ is not Lebesgue negligible}\}$$

is Lebesgue negligible. Consider y as a parameter and observe that, for every $y \notin N$, the inequality

$$v(x) - v(y) \le c(x,y) + m(f;\mathcal{H}_0)$$

holds true for almost all $x \in \mathbb{R}^n$. Applying a strong lifting ρ to both sides of that inequality yields

$$\rho(v)(x) - v(y) \le c(x,y) + m(f;\mathcal{H}_0) \tag{84}$$

for all $x \in \mathbb{R}^n$ and all $y \notin N$. Now, considering x as a parameter and applying ρ to both sides of (84) yields

$$\rho(v)(x) - \rho(v)(y) \le c(x,y) + m(f;\mathcal{H}_0) \quad \forall x, y \in \mathbb{R}^n, \tag{85}$$

that is $\rho(v) \in Q(c + m(f;\mathcal{H}_0); \mathcal{L}^\infty(\mathbb{R}^n))$.

We define u to be the convolution of $\rho(v)$ and η,

$$u(x) = (\rho(v) * \eta)(x) := \int_{\mathbb{R}^n} \rho(v)(x - z)\eta(z)\, dz, \tag{86}$$

where $\eta(z) := \pi^{-n/2} e^{-z \cdot z} = \pi^{-n/2} e^{-(z_1^2 + \cdots + z_n^2)}$, $z = (z_1, \dots, z_n) \in \mathbb{R}^n$. Since $\eta \in C^{n,\infty}$ and $\int_{\mathbb{R}^n} \rho(v)(x - z)\eta(z)\, dz = \int_{\mathbb{R}^n} \rho(v)(z)\eta(x - z)\, dz$, (86) implies $u \in C^{n,\infty}$.

[14] Since $L^1(\mathbb{R}^n)$ is separable, the restriction of the weak* topology to any bounded subset of $L^\infty(\mathbb{R}^n)$ is metrizable.

Now, taking into account the form of the function f, one has $c(x, y) = \min(g(x-y), -g(y-x))$; hence, $c(x-z, y-z) = c(x, y)$, and (85) implies

$$\rho(v)(x-z) - \rho(v)(y-z) \leq c(x, y) + m(f; \mathcal{H}_0) \quad \forall x, y \in \mathbb{R}^n. \tag{87}$$

Multiplying (87) by $\eta(z)$, integrating the obtained inequality by dz, and taking into account that $\int_{\mathbb{R}^n} \eta(z)\, dz = 1$, one gets $u(x) - u(y) \leq c(x, y) + m(f; \mathcal{H}_0)$ for all $x, y \in \mathbb{R}^n$. Thus, $u \in Q_0(c + m(f; \mathcal{H}_0)) \cap C^{n,\infty}$, and the result follows. □

References

1. Afriat S.N., The construction of utility functions from expenditure data. - *Intern. Econ. Rev.* 1967, **8**, 67–77.
2. Afriat S.N., On a system of inequalities on demand analysis: an extension of the classical method. - *Intern. Econ. Rev.* 1973, **14**, 460–472.
3. Ambrosio L., Lecture notes on optimal transport problems. In: Colli P. and Rodrigues J.F. (eds) *Mathematical Aspects of Evolving Interfaces*, Lecture Notes in Math. **1812**, 2003, pp.1–52.
4. Balder E.J., An extension of duality-stability relations to non-convex optimization problems. - *SIAM J. Contr. Optim.* 1977, **15**, 329–343.
5. Brenier Y., Polar factorization and monotone rearrangement of vector-valued functions. - Comm. Pure Appl. Math. 1991, **44**, 375–417.
6. Cafarelli L., Feldman M., McCann R.J., Constructing optimal maps for Monge's transport problem as a limit of strictly convex costs. - *J. Amer. Math. Soc.* 2002, **15**, 1–26.
7. Carlier G., Ekeland I., Levin V.L., Shananin A.A., A system of inequalities arising in mathematical economics and connected with the Monge - Kantorovich problem. - *CAHIERS du CEREMADE* 2002, Ceremade - UMR 7534 - Univ. Paris Dauphine, No 0213, 3 Mai 2002, 13 pp (http://www.ceremade.dauphine.fr/preprint/CMD/2002-13.pdf).
8. Chichilnisky G., Continuous representation of preferences. - *Rev. Econ. Studies* 1980, **47**, 959–963.
9. Debreu G., Representation of a preference ordering by a numerical function. In: R.Thrall et al. (eds) *Decision Processes*, Wiley, N.Y. 1954, pp.159–165.
10. Debreu G., Continuity properties of Paretian utility. - *Intern. Econ. Rev.* 1964, **5**, 285–293.
11. Dunford N., Schwartz J.T. *Linear Operators. Part I: General Theory*, Intersciense Publishers, N.Y., London 1958.
12. Gangbo W., McCann R.J., The geometry of optimal transportation. - *Acta Math.* 1996, **177**, 113–161.
13. Gelfand I.M., Raikov D.A., Shilov G.E. *Commutative Normed Rings*, Chelsea, N.Y. 1964.
14. Hadjisavvas N., Komlosi S., Schaible S. (eds) *Handbook of Generalized Convexity and Generalized Monotonicity*, Springer, N.Y. 2005.
15. Ioffe A.D., Abstract convexity and non-smooth analysis. - *Adv. Math. Econ.* 2001, **3**, 45–61.

16. Ionescu Tulcéa A., Ionescu Tulcéa C. *Topics in the Theory of Lifting*, Springer, N.Y. 1969.
17. Kantorovich L.V., On mass transfer. - *Dokl. Acad. Nauk SSSR* 1942, **37** (7-8), 199–201 (in Russian).
18. Kantorovich L.V., On a problem of Monge. - *Uspekhi Matem. Nauk* 1948, **3**, no.2, 225–226 (in Russian).
19. Kantorovich L.V., Akilov G.P. *Functional Analysis*, Nauka, Moscow 1984 (in Russian).
20. Kantorovich L.V., Rubinshtein G.S., On a function space and some extremal problems. - *Dokl. Acad. Nauk SSSR* 1957, **115**, 1058–1061 (in Russian).
21. Kantorovich L.V., Rubinshtein G.S., On a space of completely additive functions. - *Vestnik Leningrad Univ.* 1958, **13**(7), 52–59 (in Russian).
22. Kiruta A.Ya., Rubinov A.M., Yanovskaya E.B. *Optimal Choice of Distributions in Complex Social and Economics Problems*, Nauka, Leningrad 1980 (in Russian).
23. Kuratowski K. *Topology, vol.1*, Academic Press, N.Y.; PWN, Warsaw 1966.
24. Levin V.L., Duality and approximation in the mass transfer problem. In: B.S.Mityagin (ed) *Mathematical Economics and Functional Analysis*, Nauka, Moscow 1974, pp.94–108 (in Russian).
25. Levin V.L., On the problem of mass transfer. - *Soviet Math. Dokl.* 1975, **16**, 1349–1353 (English translation).
26. Levin V.L., The Monge—Kantorovich mass transfer problem. In: B.A.Efimov (ed) *Methods of Functional Analysis in Mathematical Economics*, Nauka, Moscow 1978, pp.23–55 (in Russian).
27. Levin V.L., Some applications of duality for the problem of translocation of masses with a lower semicontinuous cost function. Closed preferences and Choquet theory. - *Soviet Math. Dokl.* 1981, **24**, no. 2, 262–267 (English translation).
28. Levin V.L., A continuous utility theorem for closed preorders on a σ-compact metrizable space. - *Soviet Math. Dokl.* 1983, **28**, no. 3, 715–718 (English translation).
29. Levin V.L., The problem of mass transfer in a topological space, and probability measures having given marginal measures on product of two spaces. - *Soviet Math. Dokl.* 1984, **29**, 638–643 (English translation).
30. Levin V.L. *Convex Analysis in Spaces of Measurable Functions and Its Applications in Mathematics and Economics*, Nauka, Moscow 1985 (in Russian).
31. Levin V.L., Extremal problems with probability measures, functionally closed preorders and strong stochastic dominance. In: V.I.Arkin, A,Shiraev, R.Wets (eds) *Stochastic Optimization, LN in Control and Inform. Sci.* **81**, Springer, Berlin 1986, pp.435–447.
32. Levin V.L., General Monge—Kantorovich problem and its applications in measure theory and mathematical economics. In: L.J. Leifman (ed) *Functional Analysis, Optimization, and Mathematical Economics (A collection of papers dedicated to memory of L.V. Kantorovich)*, Oxford University Press, N.Y., Oxford 1990, pp.141–176.
33. Levin V.L., A formula for the optimal value of the Monge—Kantorovich problem with a smooth cost function and a characterization of cyclically monotone maps. - *Mat. Sbornik* 1990, **181**, no.12, 1694–1709 (in Russian); English translation in: *Math. USSR-Sb.* 1992, **71**, 533–548.
34. Levin V.L., Some applications of set-valued mappings in mathematical economics. - *J. of Math. Economics* 1991, **20**, 69–87.

35. Levin V.L., A superlinear multifunction arising in connection with mass transfer problems. - *Set-Valued Analysis* 1996, **4**, 41–65.
36. Levin V.L., Reduced cost functions and their applications. - *J. of Math. Economics* 1997, **28**, 155–186.
37. Levin V.L., Topics in the duality theory for mass transfer problem. In: V. Beneš, J. Štěpán (eds) *Distributions with Given Marginals and Moment Problems*, Kluwer, Dordrecht 1997, pp.243–252.
38. Levin V.L., On duality theory for non-topological variants of the mass transfer problem. - *Sbornik: Mathematics* 1997, **188**, no.4, 571–602 (English translation).
39. Levin V.L., Existence and uniqueness of a measure-preserving optimal mapping in a general Monge—Kantorovich problem. - *Funct. Anal. Appl.* 1998, **32**, 205–208 (English translation).
40. Levin V.L., Abstract cyclical monotonicity and Monge solutions for the general Monge—Kantorovich problem. - *Set-Valued Analysis* 1999, **7**, 7–32.
41. Levin V.L., Semiconic duality in convex analysis. - *Trans. Moscow Math. Soc.* 2000, **61**, 197–238 (English translation).
42. Levin V.L., The Monge—Kantorovich problems and stochastic preference relations. - *Adv. Math. Econ.* 2001, **3**, 97–124.
43. Levin V.L., Abstract convexity in measure theory and in convex analysis. - *J. Math. Sciences* 2003, **116**, issue 4, 3432–3467; (http://www.kluweronline.com/issn/1072-3374).
44. Levin V.L., Optimal solutions of the Monge problem. - *Adv. Math. Econ.* 2004, **6**, 85–122.
45. Levin V.L., Optimality conditions and exact solutions to the two-dimensional Monge—Kantorovich problem. In: A.M.Vershik (ed) *Representation Theory, Dynamical Systems. XI* (special issue) - *Zapiski nauchnyh seminarov POMI (Notes of scientific seminars of St.-Petersburg's branch of the Steklov Institute Math.)* 2004, **312**, 150–164 (in Russian);(http://www.pdmi.ras.ru/znsl/index.html); English translation to appear in *J. Math. Sciences.*
46. Levin V.L., A method in demand analysis connected with the Monge—Kantorovich problem. - *Adv. Math. Econ.* 2005, **7**, 47–93.
47. Levin V.L., Milyutin A.A., The problem of mass transfer with a discontinuous cost function and a mass statement of the duality problem for convex extremal problems. - *Russian Math. Surveys* 1979, **34**, 1–78 (English translation).
48. Mas-Colell A., On the continuous representation of preorders. - *Intern. Econ. Rev.* 1977, **18**, 509–513.
49. Monge G., Mémoire sur la théorie des déblais et de remblais. In: *Histoire de l'Académie Royale des Sciences de Paris, avec les Mémoires de Mathématique et de Physique pour la même année*, 1781, pp.666–704.
50. Mount K.R., Reiter S., Construction of a continuous utility function for a class of preferences. - *J. of Math. Economics* 1976, **3**, 227–245.
51. Naimark M.A. *Normed Algebras*, Wolters-Noordhoff, Groningen 1972.
52. Neuefeind W., On continuous utility. - *J. of Economic Theory* 1972, **5**, 174–176.
53. Pallaschke D., Rolewicz S. *Foundations of Mathematical Optimization (Convex analysis without linearity)*, Kluwer, Dordrecht 1997.
54. Penot J.P., The bearing of duality on microeconomics. - *Adv. Math. Econ.* 2005, **7**, 113–139.
55. Rachev S.T., Rüschendorf L. *Mass Transportation Problems*, Vol.1:Theory, Vol.2:Applications. Springer, Berlin 1998.

56. Rochet J.-C., A necessary and sufficient condition for rationalizability in a quasi-linear context. - *J. of Math. Economics* 1987, **16**, 191–200.
57. Rockafellar R.T. *Convex Analysis*, Princeton Univ. Press, Princeton 1970.
58. Rockafellar R.T., On the maximal monotonicity of the subdifferential mappings. - *Pacific J. Math.* 1970, **33**, 209–216.
59. Rubinov A.M. *Abstract Convexity and Global Optimization*, Kluwer, Dordrecht 2000.
60. Rubinov A.M., Abstract convexity: examples and applications. - *Optimization* 2000, **47**, 1–33.
61. Rüschendorf L., On c-optimal random variables. - *Statist. Probab. Letters* 1996, **27**, 267–270.
62. Singer I. *Abstract Convex Analysis*, Wiley-Interscience Publication, N.Y. 1997.
63. Smith C., Knott M., On Höffding-Fréchet bounds and cyclic monotone relations. - *J. Multivar. Anal.* 1992, **40**, 328–334.
64. Sudakov V.N., Geometric problems in the theory of infinite-dimensional probability distributions. - *Proc. Steklov Inst. Math.* 1979, **141**, 1–178; translated from the Russian: *Trudy Mat. Inst. Steklov* 1976, **141**.
65. Trudinger N.S., Wang X.J., On the Monge mass transfer problem. - *Calc. Var. PDE* 2001, **13**, 19–31.
66. Varian H.R., The nonparametric approach to demand analysis. - *Econometrica* 1982, **50**, 945–973.
67. Varian H.R., Non-parametric tests of consumer behaviour. - *Rev. Econ. Studies* 1983, **V(1)**, no.160, 99–110.

Optimality Conditions and Duality for Multiobjective Programming Involving (C, α, ρ, d) type-I Functions *

Dehui Yuan[1], Altannar Chinchuluun[2], Xiaoling Liu[1], and Panos M. Pardalos[2]

[1] Hanshan Teachers' College, Department of Mathematics and Computer Science, Chaozhou, Guangdong Province, China lxlydh0125@sina.com.cn, ydhlxl@sohu.com

[2] Department of Industrial and Systems Engineering, University of Florida, Gainesville, FL32611, USA altannar@ufl.edu, pardalos@ufl.edu

Summary. In this chapter, we present a unified formulation of generalized convex functions. Based on these concepts, sufficient optimality conditions for a nondifferentiable multiobjective programming problem are presented. We also introduce a general Mond-Weir type dual problem of the problem and establish weak duality theorem under generalized convexity assumptions. Strong duality result is derived using a constraint qualification for nondifferentiable multiobjective programming problems.

Key words: Multiobjective programming problem, (C, α, ρ, d)-type I function, Optimality conditions, Duality.

1 Introduction

Convexity plays an important role in the design and analysis of successful algorithms for solving optimization problems. However, the convexity assumption must be weakened in order to tackle different real-world optimization problems. Therefore, several classes of generalized convex functions have been introduced in the literature and corresponding optimality conditions and duality theorems for mathematical programming problems involving these generalized convexities have been derived. In 1981, Hanson introduced the concept of invexity in [10]. Optimality conditions and duality for different mathematical programming problems with invex functions have also been obtained

*This research of the first two authors is partially supported in part by NSF, Air Force, and CRDF grants. The research of the third author is partially supported by National Natural Science Foundation of China under Project 10201017.

by other researchers. For example, Bector and Bhatia [4] studied minimax programming problems and relaxed the convexity assumptions in the sufficient optimality in Schmitendorf [26] using invexity. Jeyakumar and Mond [12] introduced the concept of v-invexity, which can be seen as an extension of invexity, and derived optimality conditions and duality theorems for multiobjective programming problems involving the generalized convexity. Some other extensions of these generalized convexities can be found in [13], [5] and [22]. Other classes of generalized convex functions were defined in [27, 28, 11, 24, 27, 28, 5, 9, 18, 25, 6, 1, 32].

Liang et al. [14], [15] and [16] introduced a unified formulation of generalized convexity so called (F, α, ρ, d)-convexity. Recently, Yuan et al. [33] defined (C, α, ρ, d)-convexity, which is a generalization of (F, α, ρ, d)-convexity, and established optimality conditions and duality results for nondifferentiable minimax fractional programming problems involving the generalized convexity. Chinchuluun et al. [7] also considered nondifferentiable multiobjective fractional programming problems under (C, α, ρ, d)-convexity assumptions.

On the other hand, Hanson and Mond [11] defined two new classes of functions called type I and type II functions.

Based on type I functions and (F, α, ρ, d)-convexity, Hachimi and Aghezzaf [9] defined (F, α, ρ, d)-type I functions for differentiable multiobjective programming problems and derived sufficient optimality conditions and duality theorems.

In this chapter, motivated by [9], [11] and [33], we introduce (C, α, ρ, d)-type I functions. Based on the new concept of generalized convexity, we establish optimality conditions and duality theorems for the following nondifferentiable multiobjective programming problem:

$$\text{(VOP)} \qquad \min \; f(x) = (f_1(x), \cdots, f_l(x))$$
$$\text{s.t. } x \in S = \{x \in \mathbb{R}^n |\; g(x) = (g_1(x), \cdots, g_q(x)) \leqq 0\},$$

where $f_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, 2, \ldots, l$, and $g_j : \mathbb{R}^n \to \mathbb{R}$, $j = 1, 2, \ldots, q$, are Lipschitz functions on $\mathbb{R}^n$.

Throughout this chapter, we use the following notations. Let $L = \{1, \ldots, l\}$ and $Q = \{1, \ldots, q\}$ be index sets for objective and constraint functions, respectively. For $x_0 \in S$, the index set of the equality constraints is denoted by $I(x_0) = \{j | g_j(x_0) = 0\}$. If x and $y \in \mathbb{R}^n$, then

$$x \leqq y \Leftrightarrow x_i \leq y_i, i = 1, \ldots, n;$$
$$x \leqslant y \Leftrightarrow x \leqq y \text{ and } x \neq y;$$
$$x < y \Leftrightarrow x_i < y_i, i = 1, \ldots, n.$$

We denote the Clarke generalized directional derivative of f at x in the direction y and Clarke generalized gradient of f at x by $f^\circ(x; y) = (f_1^\circ(x; y), \ldots, f_l^\circ(x; y))$ and $\partial^\circ f(x) = (\partial^\circ f_1(x), \ldots, \partial^\circ f_l(x))$, respectively [8].

Definition 1. *We say that $x_0 \in S$ is an (a weak) efficient solution for problem (VOP) if and only if there exists no $x \in S$ such that $f(x) \leqslant (<) f(x_0)$.*

This chapter is organized as follows. In the next section, we introduce a unified formulation of generalized convexity. Sufficient optimality conditions for the multiobjective programming problem involving the new generalized convexity are established in Section 3. In Section 4, we extend a constraint qualification in [23] in terms of Hadamard type derivatives, relaxing some assumptions. In the last section, we present the general mixed Mond-Weir dual program for (VOP) and derive weak and strong duality results.

2 Definitions

Convexity plays a central role in mathematical programming. In addition, several problems with nonconvex functions still have properties similar to convex problems. By defining more general classes of functions, we are able to understand the structures of more general optimization problems.

In this section we introduce a unified formulation of generalized convex functions, which are extensions of (F, ρ, α, d) type-I functions presented in [9] and (C, ρ, α, d)-convex functions presented in [33].

Let $C : X \times X \times \mathbb{R}^n \to \mathbb{R}$ be convex with respect to the third argument such that $C_{(x,x_0)}(0) = 0$ for any $(x, x_0) \in S \times S$. Let $\rho = (\rho^1, \rho^2)$, where $\rho^1 = (\rho^1_1, \ldots, \rho^1_l) \in \mathbb{R}^l$, $\rho^2 = (\rho^2_1, \ldots, \rho^2_q) \in \mathbb{R}^q$. Let $\alpha = (\alpha^1, \alpha^2)$, where $\alpha^1 = (\alpha^1_1, \ldots, \alpha^1_l)$, $\alpha^2 = (\alpha^2_1, \ldots, \alpha^2_q)$, and $\alpha^i_j(\cdot, \cdot) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+ \backslash \{0\}$, $i = 1, 2$, $j \in L$ or Q. $d = (d^1, d^2)$ is a vector function, where $d^1 = (d^1_1, \ldots, d^1_l)$, $d^2 = (d^2_1, \ldots, d^2_q)$, and $d^i_j(\cdot, \cdot)$ is pseudometric on $\mathbb{R}^n$, $i = 1, 2$, $j \in L$ or Q. We assume that, for any $a, b, c \in \mathbb{R}^s$, the symbol $\frac{ab}{c}$ denotes $\left(\frac{a_1 b_1}{c_1}, \ldots, \frac{a_s b_s}{c_s}\right)$, and the symbol $\frac{a+b}{c}$ denotes $\left(\frac{a_1+b_1}{c_1}, \ldots, \frac{a_s+b_s}{c_s}\right)$. If $\xi = (\xi_1, \ldots, \xi_l) \in \partial^\circ \varphi(x_0)$, then $C_{(x,x_0)}(\xi)$ denotes the vector $(C_{(x,x_0)}(\xi_1), \ldots, C_{(x,x_0)}(\xi_l))$. We are now ready to present the new classes of functions.

(φ, ψ) is (C, α, ρ, d)-*type I* at x_0, if for all $x \in S$ we have

$$\frac{\varphi(x) - \varphi(x_0)}{\alpha^1(x, x_0)} \geqq C_{(x,x_0)}(\xi) + \frac{\rho^1 d^1(x, x_0)}{\alpha^1(x, x_0)}, \forall \xi \in \partial^\circ \varphi(x_0)$$
$$\frac{-\psi(x_0)}{\alpha^2(x, x_0)} \geqq C_{(x,x_0)}(\eta) + \frac{\rho^2 d^2(x, x_0)}{\alpha^2(x, x_0)}, \forall \eta \in \partial^\circ \psi(x_0)$$

(φ, ψ) is *pseudoquasi* (*strictly pseudoquasi*) (C, α, ρ, d)-*type I* at x_0, if for all $x \in S$ we have

$$\varphi(x) < (\leqq) \varphi(x_0) \quad \Rightarrow \quad C_{(x,x_0)}(\xi) + \frac{\rho^1 d^1(x, x_0)}{\alpha^1(x, x_0)} < 0, \forall \xi \in \partial^\circ \varphi(x_0) \quad (1)$$
$$-\psi(x_0) \leqq 0 \quad \Rightarrow \quad C_{(x,x_0)}(\eta) + \frac{\rho^2 d^2(x, x_0)}{\alpha^2(x, x_0)} \leqq 0, \forall \eta \in \partial^\circ \psi(x_0)$$

(φ, ψ) is *weak strictly-pseudoquasi* (C, α, ρ, d)-*type I* at x_0, if for all $x \in S$ we have

$$\varphi(x) \leqslant \varphi(x_0) \quad \Rightarrow \quad C_{(x,x_0)}(\xi) + \frac{\rho^1 d^1(x,x_0)}{\alpha^1(x,x_0)} < 0, \forall \xi \in \partial^\circ \varphi(x_0)$$

$$-\psi(x_0) \leqq 0 \quad \Rightarrow \quad C_{(x,x_0)}(\eta) + \frac{\rho^2 d^2(x,x_0)}{\alpha^2(x,x_0)} \leqq 0, \forall \eta \in \partial^\circ \psi(x_0)$$

(φ, ψ) is *strong pseudoquasi(weak pseudoquasi)* (C, α, ρ, d)*-type I* at x_0, if for all $x \in S$ we have

$$\varphi(x) \leqslant (<) \varphi(x_0) \quad \Rightarrow \quad C_{(x,x_0)}(\xi) + \frac{\rho^1 d^1(x,x_0)}{\alpha^1(x,x_0)} \leqslant 0, \forall \xi \in \partial^\circ \varphi(x_0) \quad (2)$$

$$-\psi(x_0) \leqq 0 \quad \Rightarrow \quad C_{(x,x_0)}(\eta) + \frac{\rho^2 d^2(x,x_0)}{\alpha^2(x,x_0)} \leqq 0, \forall \eta \in \partial^\circ \psi(x_0)$$

(φ, ψ) is *weak quasi-strictly-pseudo* (C, α, ρ, d)*-type I* at x_0, if for all $x \in S$ we have

$$\varphi(x) \leqslant \varphi(x_0) \quad \Rightarrow \quad C_{(x,x_0)}(\xi) + \frac{\rho^1 d^1(x,x_0)}{\alpha^1(x,x_0)} \leqq 0, \forall \xi \in \partial^\circ \varphi(x_0)$$

$$-\psi(x_0) \leqq 0 \quad \Rightarrow \quad C_{(x,x_0)}(\eta) + \frac{\rho^2 d^2(x,x_0)}{\alpha^2(x,x_0)} \leqslant 0, \forall \eta \in \partial^\circ \psi(x_0)$$

We note that we can derive many different classes of generalized convex functions by changing the inequalities of these conditions.

3 Sufficient Optimality

Aghezzaf and Hachimi [1, 9] considered multiobjective programming problems with (F,ρ)-convex functions and (F,α, ρ,d)-type I functions, and established a number of sufficient optimality conditions. We adapt these results to the classes of generalized (C,α, ρ,d)-type I functions.

Theorem 1. *Assume that there exist a feasible solution x_0 for (VOP) and vectors $\bar{u} = (\bar{u}_1, \ldots, \bar{u}_l) \in \mathbb{R}^l$ and $\bar{v} = (\bar{v}_1, \ldots, \bar{v}_q) \in \mathbb{R}^q$ such that*

$$0 \in \bar{u}^T \partial^\circ f(x_0) + \bar{v}^T \partial^\circ g(x_0), \quad (3)$$

$$\bar{v}^T g(x_0) = 0, \quad (4)$$

$$\bar{u} > 0, \bar{v} \geqq 0. \quad (5)$$

If (f, g_I) is strong pseudoquasi (C, α, ρ, d)-type I at x_0, and

$$\bar{u}^T \frac{\rho^1 d^1(x,x_0)}{\alpha^1(x,x_0)} + \bar{v}_I^T \frac{\rho_I^2 d_I^2(x,x_0)}{\alpha_I^2(x,x_0)} \geq 0, \quad (6)$$

then x_0 is an efficient solution of (VOP).

Proof. Suppose to the contrary that x_0 is not an efficient solution of (VOP). Then there exists a feasible solution x such that

$$f(x) \leqslant f(x_0) \text{ and } g_I(x_0) = 0.$$

Hence,

$$f(x) \leqslant f(x_0) \text{ and } -g_I(x_0) \leqq 0.$$

Since (f, g_I) is strong pseudoquasi (C, α, ρ, d)-type I at x_0, we can write

$$C_{(x,x_0)}(\xi) + \frac{\rho^1 d^1(x,x_0)}{\alpha^1(x,x_0)} \leqslant 0, \forall\ \xi \in \partial^\circ f(x_0),$$
$$C_{(x,x_0)}(\eta_I) + \frac{\rho_I^2 d_I^2(x,x_0)}{\alpha_I^2(x,x_0)} \leqq 0, \forall\ \eta_I \in \partial^\circ g_I(x_0).$$

Let us denote $\tau = \sum_{i=1}^{l} \bar{u}_i + \sum_{j\in I} \bar{v}_j$. Multiplying the above inequalities with $\frac{1}{\tau}\bar{u}$ and $\frac{1}{\tau}\bar{v}_I$, respectively, and using the convexity assumption of C, we have

$$C_{(x,x_0)}\left(\frac{1}{\tau}\bar{u}^T\xi + \frac{1}{\tau}\bar{v}_I^T\eta_I\right) + \frac{1}{\tau}\bar{u}^T\frac{\rho^1 d^1(x,x_0)}{\alpha^1(x,x_0)} + \frac{1}{\tau}\bar{v}_I^T\frac{\rho_I^2 d_I^2(x,x_0)}{\alpha_I^2(x,x_0)} < 0,$$
$$\forall \xi \in \partial^\circ f(x_0), \eta_I \in \partial^\circ g_I(x_0),$$

since $\bar{u} > 0$. From the last inequality, using (3) and (4), we have

$$\bar{u}^T\frac{\rho^1 d^1(x,x_0)}{\alpha^1(x,x_0)} + \bar{v}^T\frac{\rho^2 d^2(x,x_0)}{\alpha^2(x,x_0)} < 0,$$

which contradicts (6). □

The next theorems will be presented without proofs since they can be proven using the similar argument as in the proof of Theorem (1).

We can weaken the strict inequality requirement that $\bar{u} > 0$ in the above theorem but we require different convexity conditions on (f, g_I). This adjustment is given by the following theorem.

Theorem 2. *Assume that there exist a feasible solution x_0 for (VOP) and vectors $\bar{u} \in \mathbb{R}^l$ and $\bar{v} \in \mathbb{R}^q$ such that*

$$0 \in \bar{u}^T\partial^\circ f(x_0) + \bar{v}^T\partial^\circ g(x_0), \tag{7}$$
$$\bar{v}^T g(x_0) = 0, \tag{8}$$
$$\bar{u} \geqslant 0, \bar{v} \geqq 0.$$

If (f, g_I) is weak strictly-pseudoquasi (C, α, ρ, d)-type I at x_0, and

$$\bar{u}^T\frac{\rho^1 d^1(x,x_0)}{\alpha^1(x,x_0)} + \bar{v}_I^T\frac{\rho_I^2 d_I^2(x,x_0)}{\alpha_I^2(x,x_0)} \geq 0, \tag{9}$$

then x_0 is an efficient solution of (VOP).

Since an efficient solution is also weak efficient, the above formulated theorems are still valid for weak efficiency, however, we can weaken the convexity assumptions for weak efficient solutions. Therefore, the following theorems can be formulated.

Theorem 3. *Assume that there exist a feasible solution x_0 for (VOP) and vectors $\bar{u} \in \mathbb{R}^l$ and $\bar{v} \in \mathbb{R}^q$ such that the triplet $(x_0, \bar{u}, \bar{v})$ satisfies (3), (4) and (5). If $(f,\, g_I)$ is weak pseudoquasi $(C,\, \alpha,\, \rho,\, d)$-type I at x_0, and*

$$\bar{u}^T \frac{\rho^1 d^1(x, x_0)}{\alpha^1(x, x_0)} + \bar{v}_I^T \frac{\rho_I^2 d_I^2(x, x_0)}{\alpha_I^2(x, x_0)} \geqslant 0,$$

then x_0 is a weak efficient solution of (VOP).

Theorem 4. *Assume that there exist a feasible solution x_0 for (VOP) and vectors $\bar{u} \in \mathbb{R}^l$ and $\bar{v} \in \mathbb{R}^q$ such that the triplet $(x_0, \bar{u}, \bar{v})$ satisfies (7), (8) and (9). If $(f,\, g_I)$ is pseudoquasi $(C,\, \alpha,\, \rho,\, d)$-type I at x_0 with*

$$\sum_{i=1}^{l} \bar{u}_i \rho_i^1 \frac{d_i^1(x, x_0)}{\alpha_i^1(x, x_0)} + \sum_{j \in I} \bar{v}_j \rho_j^2 \frac{d_j^2(x, x_0)}{\alpha_j^2(x, x_0)} \geqslant 0,$$

then x_0 is a weak efficient solution for (VOP).

4 A Constraint Qualification

For some necessary optimality conditions of multiobjective programming problems, constraint qualifications are used in order to avoid the situation where some of the Lagrange multipliers vanish [17, 23]. In this section, we weaken assumptions of constraint qualification in Preda [23] in terms of Hadamard type derivatives, relaxing some assumptions. The Hadamard derivative of f at x_0 in the direction $v \in \mathbb{R}^n$ is defined by

$$df(x_0, v) = \lim_{(t,u)\to(0^+,v)} \frac{f(x_0 + tu) - f(x_0)}{t}.$$

f is said to be Hadamard differentiable at x_0 if $df(x_0, v)$ exists for all $v \in \mathbb{R}^n$. Obviously, $df(x_0, 0) = 0$.

Following Preda and Chitescu [23] we use the following notations. The tangent cone to a nonempty set W at point $x \in \text{cl}W$ is defined by

$$T(W; x) = \{v \in \mathbb{R}^n \mid \exists\{x^m\} \subset W :\ x = \lim_{m\to\infty} x^m,\ \exists\{t^m\}, t^m > 0 :\ v = \lim_{m\to\infty} t^m(x^m - x)\},$$

where $\text{cl}W$ is the closure of W.

Let x_0 be a feasible solution of Problem (VOP). For each $i \in L$, let $L^i = L\backslash\{i\}$, and let the nonempty sets $W^i(x_0)$ and $W(x_0)$ be defined as follows: $W(x_0) = \{x \in S | f(x) \leq f(x_0)\}$, $W^i(x_0) = \{x \in S | f_k(x) \leq f_k(x_0),\ \text{for}\ k \in L^i\}(l > 1)$, and $W^i(x_0) = W(x_0)(l = 1)$. Then, we give the following definition.

Definition 2. *The almost linearizing cone to $W(x_0)$ at x_0 is defined by*

$$H\left(W(x_0); x_0\right) = \{v \in \mathbb{R}^n | df_i(x_0, v) \leq 0,\ i \in L, \mathit{and}\ dg_j(x_0, v) \leq 0,\ j \in I(x_0)\}$$

Proposition 1. *If $df_i(x_0, \cdot)$ $i \in L$, and $dg_j(x_0, \cdot)$ $j \in I(x_0)$ are convex functions on $\mathbb{R}^n$, then $H\left(W(x_0); x_0\right)$ is a closed convex cone.*

Proof. The proof is very similar to that of Proposition 3.1 in [23]. So we omit this. □

The following lemma illustrates the relationship between the tangent cones $T\left(W^i(x_0); x_0\right)$ and the almost linearizing cone $H\left(W(x_0); x_0\right)$.

Lemma 1. *Let x_0 be a feasible solution of Problem* (VOP). *If $df_i(x_0, \cdot)$ $i \in L$, and $dg_j(x_0, \cdot)$ $j \in I(x_0)(\neq \varnothing)$ are convex functions on $\mathbb{R}^n$, then*

$$\bigcap_{i \in L} \text{clco}\ T\left(W^i(x_0); x_0\right) \subseteq H\left(W(x_0); x_0\right) \tag{10}$$

Proof. Here, we give a proof for only part $l > 1$ since the proof for part $l = 1$ is similar. For $i \in L$, let us define

$$H\left(W^i(x_0); x_0\right) = \{v \in \mathbb{R}^n |\ df_k(x_0, v) \leq 0,\ k \in L^i, \text{and}$$
$$dg_j(x_0, v) \leq 0,\ j \in I(x_0)\}$$

According to Proposition 1, $H\left(W^i(x_0); x_0\right)$ is closed and convex for all $i \in L$. We know that

$$\bigcap_{i \in L} H\left(W^i(x_0); x_0\right) \subseteq H\left(W(x_0); x_0\right)$$

Next, we show that, for every $i \in L$,

$$T\left(W^i(x_0); x_0\right) \subseteq H\left(W^i(x_0); x_0\right). \tag{11}$$

Let $i \in L$ and $v \in T\left(W^i(x_0); x_0\right)$. If $v = 0$, it is obvious that $v = 0 \in H\left(W^i(x_0); x_0\right)$. Now, we assume $v \neq 0$. Therefore, we have a sequence $\{x^m\} \subseteq W^i(x_0)$ and a sequence $\{t^m\} \subseteq \mathbb{R}$, with $t^m > 0$, such that

$$\lim_{m \to \infty} x^m = x_0,\ \lim_{m \to \infty} t^m(x^m - x_0) = v.$$

Let us take $v^m = t^m(x^m - x_0)$. Then, $\frac{v^m}{t^m} \to 0$ as $m \to \infty$. Since $v^m \to v$ and $v \neq 0$, for any positive real number ε, there exists a positive integer number

N such that $v^m \in B(v, \varepsilon)$ for all $m > N$. Therefore $\|v\| - \varepsilon \le \|v^m\| \le \|v\| + \varepsilon$ for all $m > N$. Hence, for all $m > N$, we have

$$\frac{\|v\| - \varepsilon}{t^m} \le \frac{\|v^m\|}{t^m} \to 0.$$

Since ε is an arbitrary positive number, selecting ε as a sufficiently small number, we can deduce that $\frac{1}{t^m} \to 0$. Then for all $j \in I(x_0)$ and for all sufficiently large m, we have

$$g_j\left(x_0 + \frac{1}{t^m} v^m\right) = g_j(x^m) \le 0 = g_j(x_0),\ j \in I(x_0), \tag{12}$$

$$f_k\left(x_0 + \frac{1}{t^m} v^m\right) = f_k(x^m) \le f_k(x_0),\ k \in L^i. \tag{13}$$

By definition of Hadamard derivative, we have

$$dg_j(x_0, v) \le 0,\ j \in I(x_0), \tag{14}$$

$$df_k(x_0, v) \le 0,\ k \in L^i. \tag{15}$$

This shows $v \in H\left(W^i(x_0); x_0\right)$ or (11) is true. Hence, due to the fact that every $H\left(W^i(x_0); x_0\right)$ is convex and closed, one obtains

$$\text{clco}\ T\left(W^i(x_0); x_0\right) \subseteq H\left(W^i(x_0); x_0\right),\ \forall\ i \in L.$$

Thus (10) holds. □

Definition 3. *We say that Problem* (VOP) *satisfies the generalized Guignard constraint qualification* (GGCQ) *at x_0 if*

$$\bigcap_{i \in L} \text{clco}\ T\left(W^i(x_0); x_0\right) \supseteq H\left(W(x_0); x_0\right). \tag{16}$$

holds.

Theorem 5. *Let $x_0 \in S$ be an efficient solution of Problem* (VOP). *Suppose that $l > 1$, and*
(A1) *constraint qualification* (GGCQ) *holds at x_0;*
(A2) *there exists $i \in L$ such that $df_i(x_0, \cdot)$ is a concave function on $\mathbb{R}^n$*
(A3) *$df_k(x_0, \cdot), k \in L^i$ and $dg_j(x_0, \cdot), j \in I(x_0)$ are convex function on $\mathbb{R}^n$.*
Then the system

$$df_k(x_0, v) \le 0, k \in L^i \tag{17}$$

$$df_i(x_0, v) < 0 \tag{18}$$

$$dg_j(x_0, v) \le 0, j \in I(x_0) \tag{19}$$

has no solution $v \in \mathbb{R}^n$.

Proof. Suppose to the contrary that there exists $v \in \mathbb{R}^n$ such that (17)–(19) hold. Obviously, $v \neq 0$. Thus, we have $0 \neq v \in H\left(W(x_0); x_0\right)$. Using Assumption (A1), we have $v \in \operatorname{clco} T\left(W^i(x_0); x_0\right)$. Therefore, there exists a sequence $\{v_s\} \subseteq T\left(W^i(x_0); x_0\right)$ such that

$$\lim_{s \to \infty} v_s = v \tag{20}$$

For any v_s, $s = 1, 2, \ldots$, there exist numbers $k_s, \lambda_{sr} \geq 0$, and $v_{sr} \in T\left(W^i(x_0); x_0\right), r = 1, 2, \ldots, k_s$, such that

$$\sum_{r=1}^{k_s} \lambda_{sr} = 1, \quad \sum_{r=1}^{k_s} \lambda_{sr} v_{sr} = v_s \tag{21}$$

Since $v_{sr} \in T\left(W^i(x_0); x_0\right)$, by definition, there exist sequences $\{x_{sr}^m\} \subseteq W^i(x_0)$ and $\{t_{sr}^m\} \subseteq \mathbb{R}$, $t_{sr}^m > 0$ for all n, such that, for any s and r,

$$\lim_{m \to \infty} x_{sr}^m = x_0, \lim_{m \to \infty} t_{sr}^m (x_{sr}^m - x_0) = v_{sr} \tag{22}$$

Let us denote $v_{sr}^m = t_{sr}^m (x_{sr}^m - x_0)$. Similarly to the corresponding part of the proof in Lemma 1, we know that $\frac{1}{t_{sr}^m} \to 0$ as $m \to \infty$. Then for any sufficiently large m, we have

$$g_j\left(x_0 + \frac{1}{t_{sr}^m} v_{sr}^m\right) = g_j(x_{sr}^m) \leq 0 = g_j(x_0), \ j \in I(x_0), \tag{23}$$

$$f_k\left(x_0 + \frac{1}{t_{sr}^m} v_{sr}^m\right) = f_k(x_{sr}^m) \leq f_k(x_0), \ k \in L^i. \tag{24}$$

and

$$f_i\left(x_0 + \frac{1}{t_{sr}^m} v_{sr}^m\right) = f_i(x_{sr}^m) \geq f_i(x_0), \tag{25}$$

since x_0 is an efficient solution to Problem (VOP). Using (22)–(25), by definition of Hadamard derivative, we can have

$$dg_j(x_0, v_{sr}) \leq 0, \ j \in I(x_0), \tag{26}$$

$$df_k(x_0, v_{sr}) \leq 0, \ k \in L^i, \tag{27}$$

$$df_i(x_0, v_{sr}) \geq 0. \tag{28}$$

From this system, (20), (21) and Assumptions (A2), (A3), it follows that

$$dg_j(x_0, v) \leq 0, \ j \in I(x_0),$$
$$df_k(x_0, v) \leq 0, \ k \in L^i,$$
$$df_i(x_0, v) \geq 0.$$

This contradicts the system (17)–(19). □

Theorem 6. *Suppose that the assumptions of Theorem 5 hold. Then, there exist vectors $\lambda \in \mathbb{R}^l$ and $\mu \in \mathbb{R}^q$ such that, for any $v \in \mathbb{R}^n$,*

$$\lambda^T df(x_0, v) + \mu^T dg(x_0, v) \geqq 0 \tag{29}$$

$$\mu^T g(x_0) = 0 \tag{30}$$

$$\lambda = (\lambda_1, \ldots, \lambda_l)^T > 0, \mu = (\mu_1, \ldots, \mu_q)^T \geqq 0 \tag{31}$$

Proof. The proof is similar to that of Theorem 3.2 in [23]. □

Remark 1. It is easy to check that if f is Hadamard differentiable then f is also directional differentiable at x_0, but, we do not need the assumption that f and g are quasiconvex at x_0 of [23].

Theorem 7. *Suppose that the assumptions of Theorem 5 hold, and suppose that $df_i(x_0, v) = f_i^\circ(x_0; v)$ and $dg_j(x_0, v) = g_j^\circ(x_0; v)$ for all $i \in L, j \in I(x_0)$. Then, there exist vectors $\lambda \in \mathbb{R}^l$ and $\mu \in \mathbb{R}^q$ such that*

$$0 \in \lambda^T \partial^\circ f(x_0) + \mu^T \partial^\circ g(x_0)$$
$$\mu^T g(x_0) = 0,$$
$$\lambda = (\lambda_1, \ldots, \lambda_l)^T > 0, \ \mu = (\mu_1, \ldots, \mu_q)^T \geqq 0.$$

Proof. By Theorem 6, we have

$$\lambda^T f^\circ(x_0, v) + \mu^T g^\circ(x_0, v) \geq 0, \tag{32}$$

for all $v \in \mathbb{R}^n$. If f_i, g_j are Hadamard differentiable, then they are directional differentiable at x_0, and

$$f_i^\circ(x_0; v) = df_i(x_0, v) = f_i'(x_0, v),$$

$$g_j^\circ(x_0; v) = dg_j(x_0, v) = g_j'(x_0, v),$$

Thus, f_i and g_j are regular for all $i \in L$ and $j \in I(x_0)$. By assumption, for any $v \in \mathbb{R}^n$, we have

$$0 \leq \lambda^T f^\circ(x_0, v) + \mu^T g^\circ(x_0, v) = \left(\lambda^T f + \mu^T g\right)^\circ (x_0, v),$$

or

$$0^T v \leq \left(\lambda^T f + \mu^T g\right)^\circ (x_0, v).$$

So, according to the definition of Clarke's generalized gradient, we have

$$0 \in \lambda^T \partial^\circ f(x_0) + \mu^T \partial^\circ g(x_0).$$

□

5 General Mixed Mond-Weir Type Dual

Duality theory plays a central role in mathematical programming. In this section, we introduce a general mixed Mond-Weir dual program of Problem (VOD) and establish the corresponding dual theorems under the generalized convexity assumptions. However, in order to derive strong duality result, we use the constraint qualification discussed in the previous section. Weakening the assumptions of constraint qualification would be helpful to establish more general strong duality result.

Let M_0, $M_1, \ldots, M_r$ be a partition of Q, i.e., $\bigcup_{k=0}^{r} M_k = Q$, $M_{k_1} \bigcap M_{k_2} = \emptyset$ for $k_1 \neq k_2$. Let e_l be the vector of $\mathbb{R}^l$ whose components are all ones. Motivated by [3, 16, 9], we define the following general mixed Mond-Weir dual of (VOP).

$$\text{(VOD)} \quad \max f(y) + {\mu_{M_0}}^T g_{M_0}(y) e_l$$

$$\text{s.t.} \ \ 0 \in \sum_{i=1}^{l} \lambda_i \partial f_i(y) + \sum_{k=0}^{r} \partial \left(\mu_{M_k}^T g_{M_k} \right)(y), \tag{33}$$

$$h_k(y) \triangleq \left({\mu_{M_k}}^T g_{M_k} \right)(y) \geq 0, \ k = 1, 2, \ldots, r,$$

$$\sum_{i=1}^{l} \lambda_i = 1, \lambda_i > 0 \ (i = 1, 2, \ldots, l), \lambda = (\lambda_1, \ldots, \lambda_l)^T,$$

$$\mu = (\mu_1, \mu_2, \ldots, \mu_q)^T \in \mathbb{R}^q_+, y \in \mathbb{R}^n, \mu_{M_k} \in \mathbb{R}^{|M_k|}_+.$$

Theorem 8 (`Weak Duality`). *Let x_0 be a feasible solution of (VOP), $(y_0, \bar{\lambda}, \bar{\mu})$ be a feasible solution of (VOD) and $h_0(y) \triangleq \bar{\mu}_{M_0}^T g_{M_0}(y)$. Let us use the following notations: $h(y) = (h_1(y), \ldots, h_r(y))$. Suppose that any of the following holds:*
(a)$\left(f + \bar{\mu}_{M_0}^T g_{M_0} e_l, h\right)$ is (C, α, ρ, d)-type I at y_0, f_i $(i = 1, \ldots, l)$ and h_0 are regular at y_0 and

$$\bar{\lambda}^T \frac{\rho^1 d^1(x_0, y_0)}{\alpha^1(x_0, y_0)} + e_r^T \frac{\rho^2 d^2(x_0, y_0)}{\alpha^2(x_0, y_0)} \geq 0, \tag{34}$$

(b) $\left(f + \bar{\mu}_{M_0}^T g_{M_0} e_l, h\right)$ is strong pseudoquasi (C, α, ρ, d)-type I at y_0, f_i $(i = 1, \cdots, l)$ and h_0 are regular at y_0 and (34) is true
(c) $(\bar{\lambda}^T f + \bar{\mu}_{M_0}^T g_{M_0}, \sum_{k=1}^{r} \mu_{M_k}^T g_{M_k})$ is pseudoquasi (C, α, ρ, d)-type I at y_0, f_i $(i = 1, \cdots, p)$ and h_k $(k = 0, 1, \ldots, r)$ are regular at y_0 and

$$\frac{\rho^1 d^1(x_0, y_0)}{\alpha^1(x_0, y_0)} + \frac{\rho^2 d^2(x_0, y_0)}{\alpha^2(x_0, y_0)} \geq 0.$$

Then the following cannot hold.

$$f(x_0) \leqslant f(y_0) + \bar{\mu}_{M_0}^T g_{M_0}(y_0) e_l. \tag{35}$$

Proof. Here we give the proofs of (a) and (b) since (c) can be proven similarly. Suppose to the contrary that (35) holds. Since x_0 is feasible for (VOP) and $\bar{\mu} \geqq 0$, (35) implies that

$$f(x_0) + \bar{\mu}_{M_0}^T g_{M_0}(x_0)e_l \leqslant f(y_0) + \bar{\mu}_{M_0}^T g_{M_0}(y_0)e_l \tag{36}$$

holds.

(a) By (35), (36) and the hypothesis (a), we can write the following statement for any $\bar{\xi}_i \in \partial f_i(y_0)$ and $\bar{\eta}_k \in \partial h_k(y_0)$.

$$\sum_{i=1}^{l} \frac{\bar{\lambda}_i}{\bar{\tau}} \frac{(f_i(x_0)+h_0(x_0)) - (f_i(y_0)+h_0(y_0))}{\alpha_i^1(x_0,y_0)} + \sum_{k=1}^{r} \frac{1}{\bar{\tau}} \frac{-h_k(y_0)}{\alpha_k^2(x_0,y_0)}$$
$$\geqslant C_{(x_0,y_0)}\left(\frac{1}{\bar{\tau}}(\bar{\lambda}^T\bar{\xi} + e_{r+1}^T\bar{\eta})\right) + \frac{1}{\bar{\tau}}\left(\bar{\lambda}^T \frac{\rho^1 d^1(x_0,y_0)}{\alpha^1(x_0,y_0)} + e_r^T \frac{\rho^2 d^2(x_0,y_0)}{\alpha^2(x_0,y_0)}\right),$$

where $\bar{\tau} = r + 2$. From (33), (34) and the above inequality, it follows that

$$\sum_{i=1}^{l} \frac{\bar{\lambda}_i}{\bar{\tau}} \frac{(f_i(x_0)+h_0(x_0)) - (f_i(y_0)+h_0(y_0))}{\alpha_i^1(x_0,y_0)} + \sum_{k=1}^{r} \frac{1}{\bar{\tau}} \frac{-h_k(y_0)}{\alpha_k^2(x_0,y_0)} \geq 0 \tag{37}$$

Since $(y_0, \bar{\lambda}, \bar{\mu})$ is a feasible solution of (VOD), it follows that $-h(y_0) \leqq 0$. Therefore, by (36), we have

$$\sum_{i=1}^{l} \frac{\bar{\lambda}_i}{\bar{\tau}} \frac{(f_i(x_0)+h_0(x_0)) - (f_i(y_0)+h_0(y_0))}{\alpha_i^1(x_0,y_0)} + \sum_{k=1}^{r} \frac{1}{\bar{\tau}} \frac{-h_k(y_0)}{\alpha_k^2(x_0,y_0)} < 0,$$

which is a contradiction to (37).

(b) By (36), $-h(y_0) \leqq 0$, the hypothesis (b) and the convexity of C, we obtain

$$C_{(x_0,y_0)}\left(\frac{1}{\bar{\tau}}(\bar{\lambda}^T\bar{\xi} + e_{r+1}^T\bar{\eta})\right) + \frac{1}{\bar{\tau}}\left(\bar{\lambda}^T \frac{\rho^1 d^1(x_0,y_0)}{\alpha^1(x_0,y_0)} + e_r^T \frac{\rho^2 d^2(x_0,y_0)}{\alpha^2(x_0,y_0)}\right) < 0.$$

Therefore, $C_{(x_0,y_0)}\left(\frac{1}{\bar{\tau}}(\bar{\lambda}^T\bar{\xi} + e_{r+1}^T\bar{\eta})\right) < 0$, which is a contradiction to (33).

□

Theorem 9 (Strong Duality). *Let the assumptions of Theorem 7 be satisfied. If $x_0 \in S$ is an efficient solution of (VOP), then there exist $\bar{\lambda} \in \mathbb{R}^l$, $\bar{\mu} \in \mathbb{R}^q$ such that $(x_0, \bar{\lambda}, \bar{\mu})$ is a feasible solution of (VOD) and the objective function values of (VOP) and (VOD) at the corresponding points are equal. Furthermore if the assumptions about the generalized convexity and the inequality (34) in Theorem 8 are also satisfied, then $(x_0, \bar{\lambda}, \bar{\mu})$ is an efficient solution of (VOD).*

Proof. By Theorem 7, it is obvious that $(x_0, \bar{\lambda}, \bar{\mu})$ is a feasible solution of (VOD). Moreover the objective function values of (VOP) and (VOD) at the corresponding points are equal since the objective functions are the same. Therefore $(x_0, \bar{\lambda}, \bar{\mu})$ is an efficient point of (VOD) due to the weak duality result in Theorem 8. □

6 Conclusions and Future Work

In this chapter we have defined some generalized convex functions. For mathematical programming problems with such functions, we have established sufficient optimality conditions for nonconvex nondifferentiable multiobjective programming problems with the generalized convex functions. We have also introduced a general mixed Mond-Weir type dual program of a multiobjective program and proved a weak duality theorem under the generalized convexity assumptions. Therefore, a strong duality theorem has been proved using a constraint qualification, which was derived after relaxing some assumptions of the constraint qualfication in [23] in terms of the Hadamard derivative, for nondifferentiable multiobjective programming. Weakening the assumptions of constraint qualification would be helpful to establish more general strong duality result. The chapter mainly focuses on theoretical aspects of the generalized convexity. We have not discussed any applications. Future work will include the solutions of real world engineering problems associated with the generalized convexities.

References

1. Aghezzaf B, Hachimi, M (2001) Sufficiency and duality in multiobjective programming involving generalized (F,ρ)-convexity. Journal of Mathematical Analysis and Applications 258:617–628
2. Aubin J-P, Frankowska, H (1990) Set-valued analysis. Birkhäuser, Boston
3. Bector CR (1973) Duality in nonlinear fractional programming. Zeitschrift für Operations Research, 17:183–193.
4. Bector CR, Bhatia BL (1985) Sufficient optimality conditions and duality for a minimax problem. Utilitas Mathematica 27:229–247
5. Bhatia D, Jain P (1994) Generalized (F, ρ)-convexity and duality for nonsmooth multiobjective programs. Optimization 31:153–164
6. Bhatia D, Mehra A (1999) Optimality conditions and duality involving arcwise connected and generalized arcwise connected functions. Journal of Optimization Theory and Applications 100:181–194
7. Chinchuluun A, Yuan DH, Pardalos PM (2006) Optimality conditions and duality for nondifferentiable multiobjective fractional programming with genreralized convexity. Annals of Operations Research, to appear
8. Clarke F (1983) Optimization and nonsmooth analysis. John Wiley Sons Inc.
9. Hachimi M, Aghezzaf B (2004) Sufficiency and duality in differentiable multiobjective programming involving generalized type I functions. Journal of Mathematical Analysis and Applications 296: 382–392
10. Hanson MA (1981) On sufficiency of the Kuhn-Tucker conditions. Journal of Mathematical Analysis and Applications 80:545–550
11. Hanson MA, Mond B (1987) Necessary and sufficient conditions in constrained optimization. Mathematical Programming 37:51–58
12. Jeyakumar V, Mond B (1992) On generalized convex mathematical programming. Journal of the Australian Mathematical Society Series B 34:43–53

13. Kaul RN, Kaur S (1985) Optimality criteria in nonlinear programming involving nonconvex functions. Journal of Mathematical Analysis and Applications 105:104–112
14. Liang ZA, Shi ZW (2003) Optimality conditions and duality for a minimax fractional programming with generalized convexity. Journal of Mathematical Analysis and Applications 277:474–488
15. Liang ZA, Huang HX, Pardalos PM (2001) Optimality conditions and duality for a class of nonlinear fractional programming problems. Journal of Optimization Theory and Application 110:611–619
16. Liang ZA, Huang HX, Pardalos PM (2003) Efficiency conditions and duality for a class of multiobjective fractional programming problems. Journal Global Optimization 27:447–471
17. Maeda T (1994) Constraint qualification in multiobjective problems: differentiable case. Journal of Optimization Theory and Applications 80:483–500
18. Majumdar AAK (1997) Optimality conditions in differentiable multiobjective programming. Journal of Optimizationd Theory and Applications 92:419–427
19. Merkovsky RR, Ward DE (1990) General constraint qualifications in nondifferentiable programming. Mathematical Programming 47:389–405
20. Mukherjee RN (1991) Generalized convex duality for multiobjective fractional programs. Journal of Mathematical Ananlysis and Applications 162:309–316
21. Mond B, Weir T (1981) Generalized concavity and duality. In: S. Schaible, and W.T. Ziemba (eds.), Generalized Convexity in Optimization and Economics. New York: Academic Press.
22. Preda V (1992) On sufficiency and duality for multiobjective programs. Journal of Mathematical Analysis and Applications 166:365–377
23. Preda V, Chitescu I (1999) On constraint qualification in multiobjective optimization problems: Semidifferentiable Case. Journal of Optimization Theory and Applications 100:417–433
24. Rueda NG, Hanson MA (1988) Optimality criteria in mathematical programming involving generalized invexity. Journal of Mathematical Analysis and Applications 130:375–385
25. Suneja SK (1997) Optimality and duality in nondefferentiable multiobjective optimization involving d-type I and related functions. Journal of Mathematical Analysis and Applications 206:465–479
26. Schmitendorf WE (1977) Necessary conditions for static minimax problems. Journal of Mathematical Analysis and Applications 57:683–693
27. Vial JP (1982) Strong convexity of sets and functions. Journal of Mathematical Economics 9:187–205
28. Vial JP (1983) Strong and weak convexity of sets and functions. Mathematical of Operations Research 8:187–205
29. Ward DE, Lee GM (2001) Generalized properly efficient solutions of vector optimization problems. Mathematical Methods of Operations Research 53:215–232
30. Ward D (1995) A comparison of second-order epiderivatives: calculus and optimality conditions. Journal of Mathematical Analysis and Applications 193:465–482
31. Weir T (1992) Pseudoconvex minimax programming. Utilitas Mathematica 27:35–57

32. Yuan DH, Chinchuluun A, Liu XL, Pardalos PM (2005) Generalized convexities and generalized gradients based on algebraic operations. Journal of Mathematical Analysis and Applications, to appear
33. Yuan DH, Liu XL, Chinchuluun A, Pardalos PM (2006) Nondifferentiable minimax fractional programming problems. Journal of Optimization Theory and Applications. to appear

Part II

Contributed Papers

Partitionable Variational Inequalities with Multi-valued Mappings

Elisabetta Allevi[1], Adriana Gnudi[2] and Igor V. Konnov[3]

[1] Department of Quantitative Methods, Brescia University, Contrada Santa Chiara, 50, Brescia 25122, Italy
allevi@eco.unibs.it

[2] Department of Mathematics, Statistics, Computer Science and Applications, Bergamo University, Via dei Caniana, 2, Bergamo 24127, Italy
adriana.gnudi@unibg.it

[3] Department of Applied Mathematics, Kazan University, ul. Kremlevskaya,18, Kazan 420008, Russia
ikonnov@ksu.ru
In this research, the third author was partially supported by RFBR Grant 01-01-00484.

Summary. We consider multi-valued variational inequalities defined on a Cartesian product of finite-dimensional subspaces. We introduce extensions of order monotonicity concepts for set-valued mappings, which are adjusted to the case where the subspaces need not be real lines. These concepts enable us to establish new existence and uniqueness results for the corresponding partitionable multi-valued variational inequalities. Following a parametric coercivity approach, we obtain convergence of the Tikhonov regularization method without monotonicity conditions.

Key words: Partitionable variational inequalities, multi-valued mappings, order monotonicity, existence and uniqueness results, regularization method.

1 Introduction

Many equilibrium problems arising in Mathematical Physics, Economics, Operations Research and other fields possess a partitionable structure which enables one to essentially weaken the conditions for existence and uniqueness results of solutions and for convergence of solution methods. Usually, such results are based on order monotonicity type assumptions, however, they are restricted with the case where subspaces are one-dimensional; see e.g. [1, 2] and references therein. Moreover, most papers in this field are devoted to the classical variational inequalities with single-valued mappings such as the standard complementarity problem, whereas many problems arising in applications involve either multi-valued mappings or non-smooth nonlinear functions; see e.g. [3, 4] and references therein. In two recent papers [5] and [6], several existence and uniqueness results for solutions of such problems involving P type mappings and convex and separable but not necessarily differentiable functions were established. These properties allowed for developing effective solution methods based on the D-gap function approach. In [7], some of these results were extended to mixed variational inequalities defined on a Cartesian product of arbitrary finite-dimensional subsets.

In this work, we suggest extensions of order monotonicity concepts for multi-valued mappings to the case where subspaces also need not be real lines and establish new existence and uniqueness results for partitionable multi-valued variational inequalities. We also obtain convergence of the Tikhonov regularization method under weakened order monotonicity conditions.

Let M be the index set $\{1, \dots, m\}$. We consider a fixed partition of the real Euclidean space R^n associated to M, i.e.

$$R^n = \prod_{s \in M} R^{n_s} = R^{n_1} \times \cdots \times R^{n_m}, \tag{1}$$

hence for each element $x \in R^n$ we can define its partition

$$x = x_1 \times \cdots \times x_m,$$

where $x_s \in R^{n_s}$. For brevity, we will also use the notation $x = (x_s \mid s \in M)$ of this partition associated to M. Let X_s be a nonempty closed convex set in R^{n_s} for every $s \in M$, and let

$$X = \prod_{s \in M} X_s,$$

i.e., X admits the partition associated to M. Let $Q : X \to \Pi(R^n)$ be an arbitrary multi-valued mapping (Here and below $\Pi(R^n)$ denotes the family of all nonempty subsets of a set R^n). Although each its value $Q(x)$ is not the Cartesian product set in general, we can still define the partition associated to M for each element q of $Q(x)$, i.e., $q = (q_s \mid s \in M)$ with $q_s \in R^{n_s}$ for $s \in M$. We consider the partitionable *variational inequality problem* (VI for short) of the form: Find $x^* = (x^*_s \mid s \in M) \in X$ such that

$$\exists q^* = (q^*_s \mid s \in M) \in Q(x^*) : \sum_{s \in M} \langle q^*_s, x_s - x^*_s \rangle \geq 0 \tag{2}$$

$$\forall x_s \in X_s, \ \forall s \in M.$$

We intend to present existence and uniqueness results of solutions and to obtain convergence of the Tikhonov regularization method for this problem under new order monotonicity type assumptions on the cost mapping.

2 Theoretical Background

So, we fix the partition associated to M of the space R^n and consider VI (2) under the following standing assumptions.

(A1) *$Q : X \to \Pi(R^n)$ is a given mapping.*

(A2) *X admits the partition associated to M, i.e.,*

$$X = \prod_{s \in M} X_s,$$

where X_s is a nonempty, convex and closed subset of R^{n_s} for every $s \in M$.

Note that X is obviously convex and closed. As usual, the partitionable VI (2) can be replaced with a system of partial variational inequalities, which are not however independent in general.

Proposition 1. *The following assertions are equivalent:*

(i) $x^* = (x^*_s \mid s \in M)$ is a solution to (2);

(ii) it holds that $x^* = (x^*_s \mid s \in M) \in X$ and there exists $q^* = (q^*_s \mid s \in M) \in Q(x^*)$ such that

$$\langle q^*_s, x_s - x^*_s \rangle \geq 0 \quad \forall x_s \in X_s, \ \forall s \in M. \tag{3}$$

Proof. It is clear that (3) implies (2). Conversely, let x^* solve (2) and there exist an index l and a point $y_l \in X_l$ such that

$$\langle q^*_l, y_l - x^*_l \rangle < 0.$$

Set $\tilde{x} = (x^*_1, \dots, x^*_{l-1}, y_l, x^*_{l+1}, \dots, x^*_n) \in X$, then we have

$$\sum_{s \in M} \langle q^*_s, \tilde{x}_s - x^*_s \rangle = \langle q^*_l, y_l - x^*_l \rangle < 0,$$

which is a contradiction. Hence, (2) implies (3). □

Definition 1. Let M be an index set such that (1) holds, and let $G : X \to \Pi(R^n)$ be a mapping with the partition associated to M. Then the mapping G is said to be

(a) a $P_0(M)$-*mapping*, if for each pair of points $x', x'' \in X$, $x' \neq x''$, and for all $g' = (g'_s \mid s \in M) \in G(x')$, $g'' = (g''_s \mid s \in M) \in G(x'')$, there exists an index i such that $x'_i \neq x''_i$ and

$$\langle x'_i - x''_i, g'_i - g''_i \rangle \geq 0;$$

(b) a $P(M)$-*mapping*, if for each pair of points $x', x'' \in X$, $x' \neq x''$, and for all $g' = (g'_s \mid s \in M) \in G(x')$, $g'' = (g''_s \mid s \in M) \in G(x'')$, there exists an index i such that

$$\langle x'_i - x''_i, g'_i - g''_i \rangle > 0;$$

(c) a *strict* $P(M)$*-mapping*, if there exists $\gamma > 0$ such that $G - \gamma I_n$ is a $P(M)$-mapping, where I_n is the identity map in R^n;

(d) a *uniform* $P(M)$*-mapping*, if for each pair of points $x', x'' \in X$, $x' \neq x''$ and for all $g' = (g'_s \mid s \in M) \in G(x')$, $g'' = (g''_s \mid s \in M) \in G(x'')$, there exists an index i such that

$$\langle x'_i - x''_i, g'_i - g''_i \rangle \geq \mu \|x - y\|^2$$

for some constant $\mu > 0$.

It is clear that each $P(M)$-mapping is a $P_0(M)$-mapping, each strict $P(M)$-mapping is a $P(M)$-mapping, and that each uniform $P(M)$-mapping is a strict $P(M)$-mapping. At the same time, the sum of $P_0(M)$-mappings is not a $P_0(M)$-mapping in general, and this is the case for other classes of mappings given in Definition 1. However, we can present particular classes whose addition does not change the class of the resulting mapping.

Definition 2. Let M be an index set such that (1) holds, and let $G : X \to \Pi(R^n)$ be a mapping with the partition associated to M. Then the mapping G is said to be an *(M)-diagonal* mapping, if

$$G(x) = (G_s(x_s) \mid s \in M).$$

In case $n_s = 1$ for $s \in M$ we obtain the usual diagonal mapping. Clearly, each (M)-diagonal mapping G is (strictly, strongly) monotone if and only if so is G_s for each $s \in M$. From the definitions we obtain the following properties immediately.

Proposition 2. *Suppose that (A2) is fulfilled,* $G : X \to \Pi(R^n)$ *is a* $P_0(M)$ *(respectively,* $P(M)$*, strict* $P(M)$*, uniform* $P(M)$*) -mapping, and* $F : X \to \Pi(R^n)$ *is a monotone (M)-diagonal mapping. Then* $G + F$ *is a* $P_0(M)$ *(respectively,* $P(M)$*, strict* $P(M)$*, uniform* $P(M)$*) -mapping.*

Moreover, the resulting order monotonicity property can be strengthened. We first give a simplified version, which is very useful for the regularization method.

Proposition 3. *Suppose that (A2) is fulfilled. If* $G : X \to R^n$ *is a* $P_0(M)$*-mapping, then, for any* $\varepsilon > 0$, $G + \varepsilon I_n$ *is a strict* $P(M)$*-mapping.*

Proof. First we show that $G^{(\varepsilon)} = G + \varepsilon I_n$ is a $P(M)$-mapping for each $\varepsilon > 0$. Choose $x', x'' \in X$, $x' \neq x''$, set $S = \{s \mid x'_s \neq x''_s\}$ and fix $\varepsilon > 0$. Since G is $P_0(M)$, there exists an index $k \in S$ such that $x'_k \neq x''_k$ and

$$\langle g'_k - g''_k, x'_k - x''_k \rangle \geq 0$$

for all $g' = (g'_s \mid s \in M) \in G(x')$, $g'' = (g''_s \mid s \in M) \in G(x'')$, moreover,

$$\varepsilon \langle x'_k - x''_k, x'_k - x''_k \rangle > 0.$$

Adding these inequalities yields

$$\langle t'_k - t''_k, x'_k - x''_k \rangle > 0$$

for all $t' = (t'_s \mid s \in M) \in G^{(\varepsilon)}(x')$, $t'' = (t''_s \mid s \in M) \in G^{(\varepsilon)}(x'')$, hence, $G^{(\varepsilon)}$ is a $P(M)$-mapping. Since $G^{(\varepsilon'')} = G^{(\varepsilon')} - (\varepsilon' - \varepsilon'')I_n = G + \varepsilon'' I_n$ is a $P(M)$-mapping, if $0 < \varepsilon'' < \varepsilon'$, we conclude that $G^{(\varepsilon)}$ is also a strict $P(M)$-mapping. □

Thus, the regularized problem becomes well-defined. Of course, we can replace the identity map with an arbitrary strongly monotone (M)-diagonal mapping.

Corollary 1. *Suppose that (A2) is fulfilled, $G : X \to \Pi(R^n)$ is a $P_0(M)$-mapping, and $F : X \to \Pi(R^n)$ is a strongly monotone (M)-diagonal mapping. Then, for any $\varepsilon > 0$, $G + \varepsilon F$ is a strict $P(M)$-mapping.*

Observe that the above regularization is not sufficient for obtaining the uniform $P(M)$ property. Thus, the concept of the strict $P(M)$-mapping is essentially weaker than the strong monotonicity.

3 General Existence and Uniqueness Results

In this section, we establish existence and uniqueness results for VI (2). First we remind that a mapping $G : X \to \Pi(R^n)$, is said to be a *K-mapping* if it is upper semicontinuous on X and has nonempty convex and compact values; see e.g. [4, Definition 2.1.1].

Proposition 4. *(see [8]) Suppose (A1) and (A2) are fulfilled, X is a bounded set, and Q is a K-mapping. Then VI (2) has a solution.*

Proposition 5. *Suppose (A1) and (A2) are fulfilled and Q is a $P(M)$-mapping. Then VI (2) has at most one solution.*

Proof. Suppose for contradiction that there exist at least two different solutions x' and x'' of VI (2). Then, by definition

$$\exists q' = (q'_i \mid i \in M) \in Q(x'), \quad \langle q'_i, x''_i - x'_i \rangle \geq 0$$

and

$$\exists q'' = (q''_i \mid i \in M) \in Q(x''), \quad \langle q''_i, x'_i - x''_i \rangle \geq 0$$

for all $i \in M$. Adding these inequalities gives

$$\langle q'_i - q''_i, x''_i - x'_i \rangle \geq 0 \text{ for all } i \in M.$$

Hence Q is not a $P(M)$-mapping, a contradiction. □

Combining both the propositions yields the following result.

Corollary 2. *Suppose (A1) and (A2) are fulfilled, X is a bounded set, and Q is a K- and $P(M)$-mapping. Then VI (2) has a unique solution.*

Now we present an existence and uniqueness result for the unbounded case.

Theorem 1. *Suppose (A1) and (A2) are fulfilled, Q is a K-mapping and a strict $P(M)$-mapping. Then VI (2) has a unique solution.*

Proof. Due to Proposition 5, it suffices to show that VI (2) is solvable. If X is bounded, then this is the case due to Proposition 4. Hence, we have to only consider the unbounded case. Fix a point $z = (z_s \mid s \in M) \in X$. For a number $\rho > 0$ we set

$$B_s(z_s, \rho) = \{x_s \in R^n_s \mid \|x_s - z_s\| \leq \rho\}$$

for each $s \in M$. Let x^ρ denote a unique solution of the problem (2) over the set

$$X_\rho = \{x \in R^n \mid x_s \in X_s \bigcap B_s(z_s, \rho), s \in M\}.$$

By Corollary 2, this is the case if X_ρ is nonempty. Then, by Proposition 1, there exists $q^\rho = (q^\rho_s \mid s \in M) \in Q(x^\rho)$ such that

$$\langle q^\rho_s, x_s - x^\rho_s \rangle \geq 0$$

for all $x_s \in X_s \bigcap B_s(z_s, \rho)$, $s \in M$.

We now proceed to show that $\|x^\rho_s - z_s\| < \rho$, $s \in M$ for $\rho > 0$ large enough. Assume for contradiction that $\|x^\rho - z\| \to +\infty$ as $\rho \to +\infty$. Choose an arbitrary sequence $\{\rho_k\} \to +\infty$ and set $y^k = x^{\rho_k}$. Choose the index set $J = \{s \mid \|y^k_s\| \to \infty \text{ as } k \to \infty\}$. Letting

$$\tilde{z}^k_s = \begin{cases} y^k_s, & \text{if } s \notin J, \\ z_s, & \text{if } s \in J, \end{cases}$$

we have

$$\langle y^k_{s_k} - \tilde{z}^k_{s_k}, q^k_{s_k} - \tilde{q}^k_{s_k} \rangle > \gamma \|y^k_{s_k} - \tilde{z}^k_{s_k}\|^2 \quad \forall q^k \in Q(y^k), \ \forall \tilde{q}^k \in Q(\tilde{z}^k), \tag{4}$$

for some s_k. Since s_k is taken from the finite set $M = \{1, ..., m\}$, without loss of generality we can suppose that s_k is fixed, i.e. $s_k = l$. Note that (4) yields $l \in J$, hence

$$\langle y^k_l - z_l, q^k_l - \tilde{q}^k_l \rangle > \gamma \|y^k_l - z_l\|^2 \quad \forall q^k \in Q(y^k), \ \forall \tilde{q}^k \in Q(\tilde{z}^k),$$

or, equivalently,

$$\langle q^k_l, y^k_l - z_l \rangle > \gamma \|y^k_l - z_l\|^2 - \langle \tilde{q}^k_l, z_l - y^k_l \rangle.$$

Since $\{\tilde{z}^k\}$ is bounded, we must have $\|\tilde{q}^k_l\| < C$ for all $\tilde{q}^k \in Q(\tilde{z}^k)$. Hence, it holds that $\|y^k_l - z_l\| \to +\infty$ and

$$\gamma\|y_l^k - z_l\|^2 - \langle \tilde{q}_l^k, z_l - y_l^k \rangle \to +\infty \text{ as } k \to \infty,$$

i.e.

$$\langle q_l^k, y_l^k - z_l \rangle > 0 \quad \forall q^k \in Q(y^k)$$

for k large enough, which contradicts the definition of y^k.

Thus there exists a number k' such that $\|y_i^k - z_i\| < \rho_k$ for all $i \in M$ if $k \geq k'$. Take an arbitrary point $x = (x_s \mid s \in M) \in X$, then there exists a number $\delta > 0$ such that

$$x_i^\rho + \delta(x_i - x_i^\rho) \in X_i \bigcap B_i(z_i, \rho) \text{ for } i \in M,$$

if $\rho \geq \rho_{k'}$. It follows that

$$\exists \tilde{q} = (\tilde{q}_s \mid s \in M) \in Q(x^\rho), \quad \langle \tilde{q}_i, x_i^\rho + \delta(x_i - x_i^\rho) - x_i^\rho \rangle \geq 0 \text{ for } i \in M,$$

or, equivalently,

$$\langle \tilde{q}_i, x_i - x_i^\rho \rangle \geq 0 \text{ for } i \in M.$$

Therefore, x^ρ solves VI (2) and the result follows. □

Observe that the existence and uniqueness results above extend those in [5]–[7] from the cases of mixed box- constrained variational inequalities and partitionable mixed variational inequalities.

4 Regularization Method

Many equilibrium type problems, which can be formulated as partitionable mixed variational inequalities and arise in applications, do not possess strengthened $P(M)$ type properties; see e.g. [6]. The regularization approach allows us to overcome this drawback by replacing the initial problem with a sequence of perturbed problems with strengthened order monotonicity properties. For the usual box-constrained variational inequalities, this method was investigated by many authors; see e.g. [2] and references therein. Here we follow the approach suggested in [9, 10], which ensures convergence of the regularization method without monotonicity conditions in the case when the feasible set is unbounded.

We also consider VI (2) under assumptions (A1) and (A2). Given a number $\varepsilon > 0$, we define the perturbed VI: Find $x^\varepsilon = (x_s^\varepsilon \mid s \in M) \in X$ such that

$$\exists q^\varepsilon = (q_s^\varepsilon \mid s \in M) \in Q(x^\varepsilon) : \sum_{s \in M} \langle q_s^\varepsilon + \varepsilon x_s^\varepsilon, x_s - x_s^\varepsilon \rangle \geq 0 \qquad (5)$$

$$\forall x_s \in X_s, \ \forall s \in M.$$

From Proposition 1 it follows that VI (5) can be rewritten equivalently as follows: there exists $q^\varepsilon = (q_s^\varepsilon \mid s \in M) \in Q(x^\varepsilon)$ such that

$$\langle q_s^\varepsilon + \varepsilon x_s^\varepsilon, x_s - x_s^\varepsilon \rangle \geq 0 \quad \forall x_s \in X_s,$$

for all $s \in M$.

We first consider convergence of the sequence $\{x^\varepsilon\}$ in the bounded case.

Theorem 2. *Suppose (A1) and (A2) are fulfilled, X is a bounded set, and $Q : X \to R^n$ is a K- and $P_0(M)$-mapping. Then VI (5) has the unique solution x^ε for each $\varepsilon > 0$, the sequence $\{x^{\varepsilon_k}\}$, where $\{\varepsilon_k\} \searrow 0$, has some limit points, and all these points are contained in the solution set of VI (2).*

Proof. From Proposition 3 it follows that $Q + \varepsilon I_n$ is a strict $P(M)$-mapping. By Corollary 2, for each $\varepsilon > 0$, VI (5) has the unique solution x^ε. Since the sequence $\{x^\varepsilon\}$ is contained in the bounded set X, it has some limit points. If x^* is an arbitrary limit point of $\{x^\varepsilon\}$, then taking the corresponding limit in (5) gives

$$\exists q^* = (q^*_s \mid s \in M) \in Q(x^*) : \sum_{s \in M} \langle q^*_s, x_s - x^*_s \rangle \geq 0 \quad \forall x_s \in X_s, \ \forall s \in M.$$

i.e. x^* solves VI (2). □

In the unbounded case, we follow the approach suggested in [9], which is based on introducing an auxiliary bounded VI and on the parametric coercivity conditions. Let us define the set

$$\tilde{X} = \prod_{s \in M} \tilde{X}_s,$$

where $\tilde{X}_s$ is a nonempty compact convex set in R^{n_s}, $\tilde{X}_s \subseteq X_s$ for every $s \in M$, i.e. it corresponds to the same fixed partition of the space R^n associated to M. Let us consider the reduced partitionable VI: Find $z^* = (z^*_s \mid s \in M) \in X$ such that

$$\exists \tilde{q} = (\tilde{q}_s \mid s \in M) \in Q(z^*) : \sum_{s \in M} \langle \tilde{q}_s, x_s - z^*_s \rangle \geq 0 \quad \forall x_s \in \tilde{X}_s, \ \forall s \in M. \tag{6}$$

We denote by X^* and $\tilde{X}^*$ the solution sets of VIs (2) and (6), respectively. Let us also consider the corresponding regularized VI: Find $z^\varepsilon = (z^\varepsilon_s \mid s \in M) \in \tilde{X}$ such that

$$\exists \tilde{q}^\varepsilon = (\tilde{q}^\varepsilon_s \mid s \in M) \in Q(z^\varepsilon) : \sum_{s \in M} \langle \tilde{q}^\varepsilon_s + \varepsilon z^\varepsilon_s, x_s - z^\varepsilon_s \rangle \geq 0 \tag{7}$$

$$\forall x_s \in \tilde{X}_s, \ \forall s \in M.$$

It is clear that $\tilde{X}^* \neq \emptyset$ and that (7) has a unique solution under the corresponding assumptions on Q. However, the strict inclusion $X^* \bigcap \tilde{X} \subset \tilde{X}^*$ may prevent to convergence of the regularization method to a solution of VI (2). We now give sufficient conditions, which ensure the precise reduction of the solution set X^*.

(A3) *There exist sets $\tilde{D} \subseteq D \subseteq R^n$ such that, for each point $y \in X \backslash D$ there exists a point $x \in \tilde{D} \bigcap X$ such that*

$$\max_{s \in M} \langle q_s, y_s - x_s \rangle > 0 \quad \forall q = (q_i \mid i \in M) \in Q(y).$$

The sets D and $\tilde{D}$ may be called the absorbing and blocking sets for the solutions of VI (2), respectively.

Proposition 6. *Suppose (A1)–(A3) are fulfilled, moreover,* $\tilde{D} \subseteq \tilde{X}$ *and* $D \bigcap \tilde{X} \subseteq X^* \bigcap \tilde{X}$. *Then* $\tilde{X}^* = \tilde{X} \bigcap X^*$.

Proof. First we note that $X^* \subseteq X \bigcap D$ due to (A3). Next, clearly, $\tilde{X} \bigcap X^* \subseteq \tilde{X}^*$. Suppose that there is a point $y \in \tilde{X}^* \backslash X^*$, then $y \in \tilde{X}$ and $y \notin D$. Applying (A3), we see that there exists a point $x \in \tilde{X}$ such that

$$\max_{s \in M} \langle q_s, y_s - x_s \rangle > 0 \quad \forall q = (q_i \mid i \in M) \in Q(y),$$

i.e. $y \notin \tilde{X}^*$, so we get a contradiction, and the result follows. □

Observe that the solution set X^* need not be bounded in the above proposition. Thus, replacing the unbounded VI (2) with a suitable bounded VI (6), which has the same solution set, we obtain convergence for the regularization method to a solution of the initial problem.

Theorem 3. *Suppose (A1)–(A3) are fulfilled,* $Q : X \rightarrow R^n$ *is a* K*- and* $P_0(M)$*-mapping, moreover,* $\tilde{D} \subseteq \tilde{X}$ *and* $D \bigcap \tilde{X} = X^* \bigcap \tilde{X}$. *Then VI (7) has the unique solution* z^{ε} *for each* $\varepsilon > 0$, *the sequence* $\{z^{\varepsilon_k}\}$, *where* $\{\varepsilon_k\} \searrow 0$, *has some limit points, and all these points are contained in the solution set of VI (2).*

Proof. Following the proof of Theorem 2, applied to the reduced VI (6), we see that VI (7) has the unique solution z^{ε} for each $\varepsilon > 0$, and that the sequence $\{z^{\varepsilon_k}\}$ has some limit points and all these points belong to $\tilde{X}^*$. Since all the assumptions of Proposition 6 hold, we obtain $\tilde{X}^* = X^* \bigcap \tilde{X}$ and the result follows. □

References

1. Ortega JM, Rheinboldt WC (1970): Iterative solution of nonlinear equations in several variables. Academic Press, New York
2. Facchinei F, Pang JS (2003): Finite-dimensional variational inequalities and complementarity problems. Springer-Verlag, Berlin (two volumes)
3. Baiocchi C, Capelo A (1984): Variational and quasivariational inequalities. Applications to free boundary problems. John Wiley and Sons, New York
4. Konnov IV (2001): Combined relaxation methods for variational inequalities. Springer-Verlag, Berlin
5. Konnov IV (2000): Properties of gap functions for mixed variational inequalities. Siberian Journal of Numerical Mathematics 3: 259–270
6. Konnov IV, Volotskaya EO (2002): Mixed variational inequalities and ecomonic equilibrium problems. Journal of Applied Mathematics 6: 289–314
7. Allevi E, Gnudi A, Konnov IV, Mazurkevich EO (2005): Partitionable mixed variational inequalities. In Giannessi F, Maugeri A (eds) Variational analysis and applications. Kluwer Academic Publishers, Dordrecht, 133–146
8. Saigal R (1976): Extensions of the generalized complementarity problem. Mathematics of Operations Research 1: 260–266

9. Konnov IV (2004): Regularization method for variational inequalities under order monotonicity conditions, In: Badriev IB et al (eds) Mesh methods for boundary value problems and applications. Kazan University Publishing Office, Kazan, 116–118
10. Konnov IV, Mazurkevich E, Ali MSS (2005): On the regularization method for variational inequalities with P_0 mappings: International Journal of Applied Mathematics and Computer Science 15: 35–44

Almost Convex Functions: Conjugacy and Duality

Radu Ioan Boţ[1], Sorin-Mihai Grad[2], and Gert Wanka[3]

[1] Faculty of Mathematics, Chemnitz University of Technology, D-09107 Chemnitz, Germany radu.bot@mathematik.tu-chemnitz.de
[2] Faculty of Mathematics, Chemnitz University of Technology, D-09107 Chemnitz, Germany sorin-mihai.grad@mathematik.tu-chemnitz.de
[3] Faculty of Mathematics, Chemnitz University of Technology, D-09107 Chemnitz, Germany gert.wanka@mathematik.tu-chemnitz.de

Summary. We prove that the formulae of the conjugates of the precomposition with a linear operator, of the sum of finitely many functions and of the sum between a function and the precomposition of another one with a linear operator hold even when the convexity assumptions are replaced by almost convexity or nearly convexity. We also show that the duality statements due to Fenchel hold when the functions involved are taken only almost convex, respectively nearly convex.

Key words: Fenchel duality, conjugate functions, almost convex functions, nearly convex functions

1 Introduction

Convexity is an important tool in many fields of Mathematics having applications in different areas, including optimization. Various generalizations of the convexity were given in the literature, so a natural consequence was to verify their applicability in optimization. We mention here the papers [3], [4], [6], [8], [9], [10], [12], [13] and [15], where properties of the convex functions and statements in convex analysis and optimization were extended by using functions and sets that are not convex but nearly convex, closely convex, convexlike, evenly convex, quasiconvex or weakly convex. Comparisons between some classes of generalized convexities were also performed, let us remind here just [6] and [9] among many others.

Within this article we work with three types of generalized convexity. Our main results concern almost convex functions, which are defined as they were introduced by Frenk and Kassay in [9]. We need to mention this because there are in the literature some other types of functions called almost convex, too. We wrote our paper motivated by the lack of known results concerning almost convex functions (cf. [9]), but also in order to introduce new and to rediscover some of our older ([3]) statements for nearly convex functions. Introduced by Aleman ([1]) as p-convex functions, the latter ones were quite intensively studied recently under the name of nearly convex functions in papers like [3], [4], [6], [10], [12], [15] and [17], while for studies on nearly convex sets we refer to [7] and [14]. Closely convexity (cf. [2], [17]) is used to illustrate some properties of the already mentioned types of functions. We have also shown that there are differences between the classes of almost convex functions and nearly convex functions, both of them being moreover larger than the one of the convex functions.

Our paper is dedicated to the extension of some results from Convex Analysis in the sense that we prove that they hold not only when the functions involved are convex, but also when they are only almost convex, respectively nearly convex. The statements we generalize concern conjugacy and duality, as follows. We prove that the formulae of some conjugates, namely of the precomposition with a linear operator, of the sum of finitely many functions and of the sum between a function and the precomposition of another one with a linear operator hold even when the convexity assumptions are replaced by almost or nearly convexity. After these, we show that the well-known duality statements due to Fenchel hold when the functions involved are taken only almost convex, respectively nearly convex. The paper is divided into five sections. After the introduction and the necessary preliminaries we give some properties of the almost convex functions, then we deal with conjugacy and Fenchel duality for this kind of functions. Some short but comprehensive conclusions and the list of references close the paper.

2 Preliminaries

This section is dedicated to the exposition of some notions and results used within our paper. Not all the results we present here are so widely-known, thus we consider necessary to recall them.

As usual, $\mathbb{R}^n$ denotes the *n-dimensional real space*, for $n \in \mathbb{N}$, and $\mathbb{Q}$ is the set of all *rational* real numbers. Throughout this paper all the vectors are considered as column vectors belonging to $\mathbb{R}^n$, unless otherwise specified. An upper index ^{T} transposes a column vector to a row one and vice versa. The *inner product* of two vectors $x = \big(x_1, ..., x_n\big)^T$ and $y = \big(y_1, ..., y_n\big)^T$ in the n-dimensional real space is denoted by $x^T y = \sum_{i=1}^n x_i y_i$. The *closure* of a certain set is distinguished from the set itself by the preceding particle cl, while the leading ri denotes the *relative interior* of the set. If $A : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, then by $A^* : \mathbb{R}^m \to \mathbb{R}^n$ we denote its *adjoint* defined by $(Ax)^T y = x^T (A^* y)\ \forall x \in \mathbb{R}^n\ \forall y \in \mathbb{R}^m$. For some set $X \subseteq \mathbb{R}^n$ we have the *indicator* function $\delta_X : \mathbb{R}^n \to \overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ defined by

$$\delta_X(x) = \begin{cases} 0, & \text{if } x \in X, \\ +\infty, & \text{if } x \notin X. \end{cases}$$

Definition 1. *For a function* $f : \mathbb{R}^n \to \overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ *we consider the following notions*

(*i*) epigraph*:* $\mathrm{epi}(f) = \big\{(x,r) \in \mathbb{R}^n \times \mathbb{R} : f(x) \le r\big\}$,
(*ii*) (effective) domain*:* $\mathrm{dom}(f) = \big\{x \in \mathbb{R}^n : f(x) < +\infty\big\}$,
(*iii*) f *is called* proper *if* $\mathrm{dom}(f) \neq \emptyset$ *and* $f(x) > -\infty \ \forall x \in \mathbb{R}^n$,
(*iv*) $\bar{f}$ *is called the* lower-semicontinuous hull of f *if* $\mathrm{epi}(\bar{f}) = \mathrm{cl}(\mathrm{epi}(f))$.
(*v*) subdifferential of f at x *(where* $f(x) \in \mathbb{R}$*):*

$$\partial f(x) = \{p \in \mathbb{R}^n : f(y) - f(x) \ge p^T(y-x) \ \forall y \in \mathbb{R}^n\}.$$

Remark 1. For any function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ we have $\mathrm{dom}(f) \subseteq \mathrm{dom}(\bar{f}) \subseteq \mathrm{cl}(\mathrm{dom}(f))$, which implies $\mathrm{cl}(\mathrm{dom}(f)) = \mathrm{cl}(\mathrm{dom}(\bar{f}))$.

Definition 2. *A set* $X \subseteq \mathbb{R}^n$ *is called* nearly convex *if there is a constant* $\alpha \in]0,1[$ *such that for any* x *and* y *belonging to* X *one has* $\alpha x + (1-\alpha)y \in X$.

An example of a nearly convex set which is not convex is $\mathbb{Q}$. Important properties of the nearly convex sets follow.

Lemma 1. *([1]) For every nearly convex set* $X \subseteq \mathbb{R}^n$ *the following properties are valid*

(*i*) $\mathrm{ri}(X)$ *is convex (may be empty),*
(*ii*) $\mathrm{cl}(X)$ *is convex,*
(*iii*) *for every* $x \in \mathrm{cl}(X)$ *and* $y \in \mathrm{ri}(X)$ *we have* $tx + (1-t)y \in \mathrm{ri}(X)$ *for each* $0 \le t < 1$.

Definition 3. *([6], [9]) A function* $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ *is called*

(*i*) almost convex *if* $\bar{f}$ *is convex and* $\mathrm{ri}(\mathrm{epi}(\bar{f})) \subseteq \mathrm{epi}(f)$,
(*ii*) nearly convex *if* $\mathrm{epi}(f)$ *is nearly convex,*
(*iii*) closely convex *if* $\mathrm{epi}(\bar{f})$ *is convex (i.e.* $\bar{f}$ *is convex).*

Connections between these kinds of functions arise from the following observations, while to show that there are differences between them we give Example 1 within the next section.

Remark 2. Any almost convex function is also closely convex.

Remark 3. Any nearly convex function has a nearly convex effective domain. Moreover, as its epigraph is nearly convex, the function is also closely convex, according to Lemma 1(*ii*).

Although cited from the literature, the following auxiliary results are not so widely known, thus we have included them here.

Lemma 2. *([4], [9]) For a convex set* $C \subseteq \mathbb{R}^n$ *and any non-empty set* $X \subseteq \mathbb{R}^n$ *satisfying* $X \subseteq C$ *we have* $\mathrm{ri}(C) \subseteq X$ *if and only if* $\mathrm{ri}(C) = \mathrm{ri}(X)$.

Lemma 3. *([4]) Let $X \subseteq \mathbb{R}^n$ be a non-empty nearly convex set. Then* $\mathrm{ri}(X) \neq \emptyset$ *if and only if* $\mathrm{ri}(\mathrm{cl}(X)) \subseteq X$.

Lemma 4. *([4]) For a non-empty nearly convex set $X \subseteq \mathbb{R}^n$,* $\mathrm{ri}(X) \neq \emptyset$ *if and only if* $\mathrm{ri}(X) = \mathrm{ri}(\mathrm{cl}(X))$.

Using the last remark and Lemma 3 we deduce the following statement.

Proposition 1. *If $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is a nearly convex function satisfying* $\mathrm{ri}(\mathrm{epi}(f)) \neq \emptyset$*, then it is almost convex.*

Remark 4. Each convex function is both nearly convex and almost convex.

The first observation is obvious, while the second can be easily proven. Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a convex function. If $f(x) = +\infty$ everywhere then $\mathrm{epi}(f) = \emptyset$, which is closed, so $\bar{f} = f$ and it follows f almost convex. Otherwise, $\mathrm{epi}(f)$ is non-empty and, being convex because of f's convexity, it has a non-empty relative interior (cf. Theorem 6.2 in [16]) so, by Proposition 1, is almost convex.

3 Properties of the Almost Convex Functions

Within this part of our paper we present some properties of the almost convex functions and some examples that underline the differences between this class of functions and the nearly convex functions.

Theorem 1. *([9]) Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ having non-empty domain. The function f is almost convex if and only if $\bar{f}$ is convex and $\bar{f}(x) = f(x)\ \forall x \in \mathrm{ri}(\mathrm{dom}(\bar{f}))$.*

Proof. "$\Rightarrow$" When f is almost convex, $\bar{f}$ is convex. As $\mathrm{dom}(f) \neq \emptyset$, we have $\mathrm{dom}(\bar{f}) \neq \emptyset$. It is known (cf. [16]) that

$$\mathrm{ri}(\mathrm{epi}(\bar{f})) = \big\{(x, r) : \bar{f}(x) < r, x \in \mathrm{ri}(\mathrm{dom}(\bar{f}))\big\} \tag{1}$$

so, as the definition of the almost convexity includes $\mathrm{ri}(\mathrm{epi}(\bar{f})) \subseteq \mathrm{epi}(f)$, it follows that for any $x \in \mathrm{ri}(\mathrm{dom}\bar{f}))$ and $\varepsilon > 0$ one has $(x, \bar{f}(x) + \varepsilon) \in \mathrm{epi}(f)$. Thus $\bar{f}(x) \geq f(x)\ \forall x \in \mathrm{ri}(\mathrm{dom}(\bar{f}))$ and the definition of $\bar{f}$ yields the coincidence of f and $\bar{f}$ over $\mathrm{ri}(\mathrm{dom}(\bar{f}))$.

"$\Leftarrow$" We have $\bar{f}$ convex and $\bar{f}(x) = f(x)\ \forall x \in \mathrm{ri}(\mathrm{dom}(\bar{f}))$. Thus $\mathrm{ri}(\mathrm{dom}(\bar{f})) \subseteq \mathrm{dom}(f)$. By Lemma 2 and Remark 1 one gets $\mathrm{ri}(\mathrm{dom}(\bar{f})) \subseteq \mathrm{dom}(f)$ if and only if $\mathrm{ri}(\mathrm{dom}(\bar{f})) = \mathrm{ri}(\mathrm{dom}(f))$, therefore this last equality holds. Using this and (1) it follows $\mathrm{ri}(\mathrm{epi}(\bar{f})) = \big\{(x, r) : f(x) < r, x \in \mathrm{ri}(\mathrm{dom}(f))\big\}$, so $\mathrm{ri}(\mathrm{epi}(\bar{f})) \subseteq \mathrm{epi}(f)$. This and the hypothesis $\bar{f}$ convex yield that f is almost convex. □

Remark 5. From the previous proof we obtain also that if f is almost convex and has a non-empty domain then $\text{ri}(\text{dom}(f)) = \text{ri}(\text{dom}(\bar{f})) \neq \emptyset$. We have also $\text{ri}(\text{epi}(\bar{f})) \subseteq \text{epi}(f)$, from which, by the definition of $\bar{f}$, follows

$$\text{ri}(\text{cl}(\text{epi}(f))) \subseteq \text{epi}(f) \subseteq \text{cl}(\text{epi}(f)).$$

Applying Lemma 2 we get $\text{ri}(\text{epi}(f)) = \text{ri}(\text{cl}(\text{epi}(f))) = \text{ri}(\text{epi}(\bar{f}))$.

In order to avoid confusions between the nearly convex functions and the almost convex functions we give below some examples showing that there is no inclusion between these two classes of functions. Their intersection is not empty, as Remark 4 states that the convex functions are concomitantly almost convex and nearly convex.

Example 1. (i) Let $f : \mathbb{R} \to \mathbb{R}$ be any discontinuous solution of Cauchy's functional equation $f(x + y) = f(x) + f(y)\ \forall x, y \in \mathbb{R}$. For each of these functions, whose existence is guaranteed in [11], one has

$$f\Big(\frac{x+y}{2}\Big) = \frac{f(x) + f(y)}{2}\ \forall x, y \in \mathbb{R},$$

i.e. these functions are nearly convex. None of these functions is convex because of the absence of continuity. We have that $\text{dom}(f) = \mathbb{R} = \text{ri}(\text{dom}(f))$. Suppose f is almost convex. Then Theorem 1 yields $\bar{f}$ convex and $f(x) = \bar{f}(x)$ $\forall x \in \mathbb{R}$. Thus f is convex, but this is false. Therefore f is nearly convex, but not almost convex.

(ii) Consider the set $X = \big([0, 2] \times [0, 2]\big) \setminus \big(\{0\} \times]0, 1[\big)$ and let $g : \mathbb{R}^2 \to \overline{\mathbb{R}}$, $g = \delta_X$. We have $\text{epi}(g) = X \times [0, +\infty)$, so $\text{epi}(\bar{g}) = \text{cl}(\text{epi}(g)) = [0, 2] \times [0, 2] \times [0, +\infty)$. As this is a convex set, $\bar{g}$ is a convex function. We also have $\text{ri}(\text{epi}(\bar{g})) =]0, 2[\times]0, 2[\times]0, +\infty)$, which is clearly contained inside $\text{epi}(g)$. Thus g is almost convex. On the other hand, $\text{dom}(g) = X$ and X is not a nearly convex set, because for any $\alpha \in]0, 1[$ we have $\alpha(0, 1) + (1 - \alpha)(0, 0) = (0, \alpha) \notin X$. By Remark 3 it follows that the almost convex function g is not nearly convex.

Using Remark 4 and the facts above we see that there are almost convex and nearly functions which are not convex, i.e. both these classes are larger than the one of convex functions.

The following assertion states an interesting and important property of the almost convex functions that is not applicable for nearly convex functions.

Theorem 2. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ and $g : \mathbb{R}^m \to \overline{\mathbb{R}}$ be proper almost convex functions. Then the function $F : \mathbb{R}^n \times \mathbb{R}^m \to \overline{\mathbb{R}}$ defined by $F(x, y) = f(x) + g(y)$ is almost convex, too.*

Proof. Consider the linear operator $L : (\mathbb{R}^n \times \mathbb{R}) \times (\mathbb{R}^m \times \mathbb{R}) \to \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}$ defined as $L(x, r, y, s) = (x, y, r+s)$. Let us first show that $L(\text{epi}(f) \times \text{epi}(g)) = \text{epi}(F)$.

Taking the pairs $(x,r) \in \text{epi}(f)$ and $(y,s) \in \text{epi}(g)$ we have $f(x) \le r$ and $g(y) \le s$, so $F(x,y) = f(x) + g(y) \le r + s$, i.e. $(x,y,r+s) \in \text{epi}(F)$. Thus $L(\text{epi}(f) \times \text{epi}(g)) \subseteq \text{epi}(F)$.

On the other hand, for $(x,y,t) \in \text{epi}(F)$ one has $F(x,y) = f(x)+g(y) \le t$, so $f(x)$ and $g(y)$ are finite. It follows $(x, f(x), y, t - f(x)) \in \text{epi}(f) \times \text{epi}(g)$, i.e. $(x,y,t) \in L(\text{epi}(f) \times \text{epi}(g))$ meaning $\text{epi}(F) \subseteq L(\text{epi}(f) \times \text{epi}(g))$.

Therefore $L(\text{epi}(f) \times \text{epi}(g)) = \text{epi}(F)$. We prove that $\text{cl}(\text{epi}(F))$ is convex, which means $\bar{F}$ convex.

Let (x,y,r) and (u,v,s) in $\text{cl}(\text{epi}(F))$. There are two sequences,

$$(x_k, y_k, r_k)_{k\ge1} \text{ and } (u_k, v_k, s_k)_{k\ge1}$$

in $\text{epi}(F)$, the first converging towards (x,y,r) and the second to (u,v,s). Then we also have the sequences of reals $(r_k^1)_{k\ge1}$, $(r_k^2)_{k\ge1}$, $(s_k^1)_{k\ge1}$ and $(s_k^2)_{k\ge1}$ fulfilling for each $k \ge 1$ the following $r_k^1 + r_k^2 = r_k$, $s_k^1 + s_k^2 = s_k$, $(x_k, r_k^1) \in \text{epi}(f)$, $(y_k, r_k^2) \in \text{epi}(g)$, $(u_k, s_k^1) \in \text{epi}(f)$ and $(v_k, s_k^2) \in \text{epi}(g)$. Let $\lambda \in [0,1]$. We have, due to the convexity of the lower-semicontinuous hulls of f and g, $(\lambda x_k + (1-\lambda)u_k, \lambda r_k^1 + (1-\lambda)s_k^1) \in \text{cl}(\text{epi}(f)) = \text{epi}(\bar{f})$ and $(\lambda y_k + (1-\lambda)v_k, \lambda r_k^2 + (1-\lambda)s_k^2) \in \text{cl}(\text{epi}(g)) = \text{epi}(\bar{g})$. Further, $(\lambda x_k + (1-\lambda)u_k, \lambda y_k + (1-\lambda)v_k, \lambda r_k + (1-\lambda)s_k) \in L(\text{cl}(\text{epi}(f)) \times \text{cl}(\text{epi}(g))) = L(\text{cl}(\text{epi}(f) \times \text{epi}(g))) \subseteq \text{cl}(L(\text{epi}(f) \times \text{epi}(g)))$ for all $k \ge 1$. Letting k converge towards $+\infty$ we get $(\lambda x + (1-\lambda)u, \lambda y + (1-\lambda)v, \lambda r + (1-\lambda)s) \in \text{cl}(L(\text{epi}(f) \times \text{epi}(g))) = \text{cl}(\text{epi}(F))$. As this happens for any $\lambda \in [0,1]$ it follows $\text{cl}(\text{epi}(F))$ convex, so $\text{epi}(\bar{F})$ is convex, i.e. $\bar{F}$ is a convex function.

Therefore, in order to obtain that F is almost convex we have to prove only that $\text{ri}(\text{cl}(\text{epi}(F))) \subseteq \text{epi}(F)$. Using some basic properties of the closures and relative interiors and also that f and g are almost convex we have $\text{ri}(\text{cl}(\text{epi}(f) \times \text{epi}(g))) = \text{ri}(\text{cl}(\text{epi}(f)) \times \text{cl}(\text{epi}(g))) = \text{ri}(\text{cl}(\text{epi}(f))) \times \text{ri}(\text{cl}(\text{epi}(g))) \subseteq \text{epi}(f) \times \text{epi}(g)$. Applying the linear operator L to both sides we get $L(\text{ri}(\text{cl}(\text{epi}(f) \times \text{epi}(g)))) \subseteq L(\text{epi}(f) \times \text{epi}(g)) = \text{epi}(F)$. One has $\text{cl}(\text{epi}(f) \times \text{epi}(g)) = \text{cl}(\text{epi}(f)) \times \text{cl}(\text{epi}(g)) = \text{epi}(\bar{f}) \times \text{epi}(\bar{g})$, which is a convex set, so also $L(\text{cl}(\text{epi}(f) \times \text{epi}(g)))$ is convex. As for any linear operator $A : \mathbb{R}^n \to \mathbb{R}^m$ and any convex set $X \subseteq \mathbb{R}^n$ one has $A(\text{ri}(X)) = \text{ri}(A(X))$ (see for instance Theorem 6.6 in [16]), it follows

$$\text{ri}(L(\text{cl}(\text{epi}(f) \times \text{epi}(g)))) = L(\text{ri}(\text{cl}(\text{epi}(f) \times \text{epi}(g)))) \subseteq \text{epi}(F). \qquad (2)$$

On the other hand, $\text{epi}(F) = L(\text{epi}(f) \times \text{epi}(g)) \subseteq L(\text{cl}(\text{epi}(f) \times \text{epi}(g))) \subseteq \text{cl}(L(\text{epi}(f) \times \text{epi}(g)))$, so $\text{cl}(L(\text{epi}(f) \times \text{epi}(g))) = \text{cl}(L(\text{cl}(\text{epi}(f) \times \text{epi}(g))))$ and further

$$\text{ri}(\text{cl}(L(\text{epi}(f) \times \text{epi}(g)))) = \text{ri}(\text{cl}(L(\text{cl}(\text{epi}(f) \times \text{epi}(g))))).$$

As for any convex set $X \subseteq \mathbb{R}^n$ $\text{ri}(\text{cl}(X)) = \text{ri}(X)$ (see Theorem 6.3 in [16]), we have $\text{ri}(\text{cl}(L(\text{cl}(\text{epi}(f) \times \text{epi}(g))))) = \text{ri}(L(\text{cl}(\text{epi }(f) \times \text{epi}(g))))$, which implies $\text{ri}(\text{cl}(L(\text{epi}(f) \times \text{epi}(g)))) = \text{ri}(L(\text{cl}(\text{epi}(f) \times \text{epi}(g))))$. Using (2) it follows that $\text{ri}(\text{epi}(\bar{F})) = \text{ri}(\text{cl}(\text{epi}(F))) = \text{ri}(L(\text{cl}(\text{epi}(f) \times \text{epi}(g)))) \subseteq \text{epi}(F)$. Because $\bar{F}$ is a convex function it follows by definition that F is almost convex. $\square$

Corollary 1. *Using the previous statement it can be shown that if* $f_i : \mathbb{R}^{n_i} \to \overline{\mathbb{R}}$, $i = 1, ..., k$, *are proper almost convex functions, then* $F : R^{n_1} \times ... \times \mathbb{R}^{n_k} \to \overline{\mathbb{R}}$, $F(x^1, ..., x^k) = \sum_{i=1}^{k} f_i(x^i)$ *is almost convex, too.*

Next we give an example that shows that the property just proven to hold for almost convex functions does not apply for nearly convex functions.

Example 2. Consider the sets

$$X_1 = \bigcup_{n \geq 1} \left\{ \tfrac{k}{2^n} : 0 \leq k \leq 2^n \right\} \quad \text{and} \quad X_2 = \bigcup_{n \geq 1} \left\{ \tfrac{k}{3^n} : 0 \leq k \leq 3^n \right\}.$$

They are both nearly convex, X_1 for $\alpha = 1/2$ and X_2 for $\alpha = 1/3$, for instance. It is easy to notice that δ_{X_1} and δ_{X_2} are nearly convex functions. Taking $F : \mathbb{R}^2 \to \overline{\mathbb{R}}$, $F(x_1, x_2) = \delta_{X_1}(x_1) + \delta_{X_2}(x_2)$, we have $\mathrm{dom}(F) = X_1 \times X_2$, which is not nearly convex, thus F is not a nearly convex function. To show this, we have $(0,0), (1,1) \in \mathrm{dom}(F)$ and assuming $\mathrm{dom}(F)$ nearly convex with the constant $\bar{\alpha} \in]0,1[$, one gets $(\bar{\alpha}, \bar{\alpha}) \in \mathrm{dom}(F)$. This yields $\bar{\alpha} \in X_1 \cap X_2$ and, so, $\bar{\alpha} \in \{0, 1\}$, which is false. Therefore F is not nearly convex.

4 Conjugacy and Fenchel Duality for Almost Convex Functions

This section is dedicated to the generalization of some well-known results concerning the conjugate of convex functions. We prove that they keep their validity when the functions involved are taken almost convex, too. Moreover, these results are proven to stand also when the functions are nearly convex and their epigraphs have non-empty relative interiors.

First we deal with the conjugate of the precomposition with a linear operator (see, for instance, Theorem 16.3 in [16]).

Theorem 3. *Let* $f : \mathbb{R}^m \to \overline{\mathbb{R}}$ *be an almost convex function and* $A : \mathbb{R}^n \to \mathbb{R}^m$ *a linear operator such that there is some* $x' \in \mathbb{R}^n$ *satisfying* $Ax' \in \mathrm{ri}(\mathrm{dom}(f))$. *Then for any* $p \in \mathbb{R}^m$ *one has*

$$(f \circ A)^*(p) = \inf \left\{ f^*(q) : A^* q = p \right\},$$

and the infimum is attained.

Proof. We first prove that $(f \circ A)^*(p) = (\bar{f} \circ A)^*(p)\ \forall p \in \mathbb{R}^n$. By Remark 5 we get $Ax' \in \mathrm{ri}(\mathrm{dom}(\bar{f}))$. Assume first that $\bar{f}$ is not proper. Corollary 7.2.1 in [16] yields $\bar{f}(y) = -\infty\ \forall y \in \mathrm{dom}(\bar{f})$. As $\mathrm{ri}(\mathrm{dom}(\bar{f})) = \mathrm{ri}(\mathrm{dom}(f))$ and $\bar{f}(y) = f(y)\ \forall y \in \mathrm{ri}(\mathrm{dom}(\bar{f}))$, one has $\bar{f}(Ax') = f(Ax') = -\infty$. It follows easily $(\bar{f} \circ A)^*(p) = (f \circ A)^*(p) = +\infty\ \forall p \in \mathbb{R}^n$.

Now take $\bar{f}$ proper. By definition one has $(\bar{f} \circ A)(x) \leq (f \circ A)(x)\ \forall x \in \mathbb{R}^n$ and, by simple calculations, one gets $(\bar{f} \circ A)^*(p) \geq (f \circ A)^*(p)$ for any $p \in \mathbb{R}^n$.

Take some $p \in \mathbb{R}^n$ and denote $\beta := (\bar{f} \circ A)^*(p) \in]-\infty, +\infty]$. Assume $\beta \in \mathbb{R}$. We have $\beta = \sup_{x \in \mathbb{R}^n} \{p^T x - \bar{f} \circ A(x)\}$. Let $\varepsilon > 0$. Then there is an $\bar{x} \in \mathbb{R}^n$ such that $p^T \bar{x} - \bar{f} \circ A(\bar{x}) \geq \beta - \varepsilon$, so $A\bar{x} \in \mathrm{dom}(\bar{f})$. As $Ax' \in \mathrm{ri}(\mathrm{dom}(\bar{f}))$, we get, because of the linearity of A and of the convexity of $\mathrm{dom}(\bar{f})$, by Theorem 6.1 in [16] that for any $\lambda \in]0,1]$ it holds $A((1-\lambda)\bar{x} + \lambda x') = (1-\lambda)A\bar{x} + \lambda A x' \in \mathrm{ri}(\mathrm{dom}(\bar{f}))$. Applying Theorem 1 and using the convexity of $\bar{f}$ we have

$$p^T((1-\lambda)\bar{x} + \lambda x') - f(A((1-\lambda)\bar{x} + \lambda x')) = p^T((1-\lambda)\bar{x} + \lambda x')$$

$$-\bar{f}(A((1-\lambda)\bar{x} + \lambda x')) \geq p^T((1-\lambda)\bar{x} + \lambda x') - (1-\lambda)\bar{f} \circ A(\bar{x})$$

$$-\lambda \bar{f} \circ A(x') = p^T \bar{x} - \bar{f} \circ A(\bar{x}) + \lambda\big[p^T(x' - \bar{x}) - (\bar{f} \circ A(x') - \bar{f} \circ A(\bar{x}))\big].$$

As Ax' and $A\bar{x}$ belong to the domain of the proper function $\bar{f}$, there is a $\bar{\lambda} \in]0,1]$ such that $\bar{\lambda}\big[p^T(x' - \bar{x}) - (\bar{f} \circ A(x') - \bar{f} \circ A(\bar{x}))\big] > -\varepsilon$.

The calculations above lead to

$$(f \circ A)^*(p) \geq p^T((1-\bar{\lambda})\bar{x} + \bar{\lambda} x') - (\bar{f} \circ A)((1-\bar{\lambda})\bar{x} + \bar{\lambda} x') \geq \beta - 2\varepsilon.$$

As ε is an arbitrarily chosen positive number, let it converge towards 0. We get $(f \circ A)^*(p) \geq \beta = (\bar{f} \circ A)^*(p)$. Because the opposite inequality is always true, we get $(f \circ A)^*(p) = (\bar{f} \circ A)^*(p)$.

Consider now the last possible situation, $\beta = +\infty$. Then for any $k \geq 1$ there is an $x_k \in \mathbb{R}^n$ such that $p^T x_k - \bar{f}(Ax_k) \geq k+1$. Thus $Ax_k \in \mathrm{dom}(\bar{f})$ and by Theorem 6.1 in [16] we have, for any $\lambda \in]0,1]$,

$$p^T((1-\lambda)x_k + \lambda x') - f \circ A((1-\lambda)x_k + \lambda x') = p^T((1-\lambda)x_k + \lambda x')$$

$$-\bar{f} \circ A((1-\lambda)x_k + \lambda x') \geq p^T((1-\lambda)x_k + \lambda x') - (1-\lambda)\bar{f} \circ A(x_k)$$

$$-\lambda \bar{f} \circ A(x') = p^T x_k - \bar{f} \circ A(x_k) + \lambda\big[p^T(x' - x_k) - (\bar{f} \circ A(x') - \bar{f} \circ A(x_k))\big].$$

Like before, there is some $\bar{\lambda} \in]0,1[$ such that

$$\bar{\lambda}\big[p^T(x' - x_k) - (\bar{f} \circ A(x') - \bar{f} \circ A(x_k))\big] \geq -1.$$

Denoting $z_k := (1-\bar{\lambda})x_k + \bar{\lambda} x'$ we have $z_k \in \mathbb{R}^n$ and $p^T z_k - f \circ A(z_k) \geq k + 1 - 1 = k$. As $k \geq 1$ is arbitrarily chosen, one gets

$$(f \circ A)^*(p) = \sup_{x \in \mathbb{R}^n} \{p^T x - f \circ A(x)\} = +\infty,$$

so $(f \circ A)^*(p) = +\infty = (\bar{f} \circ A)^*(p)$. Therefore, as $p \in \mathbb{R}^n$ has been arbitrary chosen, we get

$$(f \circ A)^*(p) = (\bar{f} \circ A)^*(p) \ \forall p \in \mathbb{R}^n. \tag{3}$$

By Theorem 16.3 in [16] we have, as $\bar{f}$ is convex and $Ax' \in \mathrm{ri}(\mathrm{dom}(\bar{f})) = \mathrm{ri}(\mathrm{dom}(f))$,

$$(\bar{f} \circ A)^*(p) = \inf \{(\bar{f})^*(q) : A^* q = p\},$$

with the infimum attained at some $\bar{q}$. But $f^* = (\bar{f})^*$ (cf. [16]), so the relation above gives

$$(\bar{f} \circ A)^*(p) = \inf \{ f^*(q) : A^*q = p \}.$$

Finally, by (3), this turns into

$$(f \circ A)^*(p) = \inf \{ f^*(q) : A^*q = p \},$$

and the infimum is attained at $\bar{q}$. □

The following statement follows from Theorem 3 immediately by Proposition 1.

Corollary 2. *If $f : \mathbb{R}^m \to \overline{\mathbb{R}}$ is a nearly convex function satisfying* $\text{ri}(\text{epi}(f)) \neq \emptyset$ *and $A : \mathbb{R}^n \to \mathbb{R}^m$ is a linear operator such that there is some $x' \in \mathbb{R}^n$ fulfilling $Ax' \in \text{ri}(\text{dom}(f))$, then for any $p \in \mathbb{R}^m$ one has*

$$(f \circ A)^*(p) = \inf \{ f^*(q) : A^*q = p \},$$

and the infimum is attained.

Now we give a statement concerning the conjugate of the sum of finitely many proper functions, which is actually the infimal convolution of their conjugates also when the functions are almost convex functions, provided that the relative interiors of their domains have a point in common.

Theorem 4. *Let $f_i : \mathbb{R}^n \to \overline{\mathbb{R}}$, $i = 1, ..., k$, be proper and almost convex functions whose domains satisfy $\cap_{i=1}^k \text{ri}(\text{dom}(f_i)) \neq \emptyset$. Then for any $p \in \mathbb{R}^n$ we have*

$$(f_1 + ... + f_k)^*(p) = \inf \left\{ \sum_{i=1}^k f_i^*(p^i) : \sum_{i=1}^k p^i = p \right\}, \tag{4}$$

with the infimum attained.

Proof. Let $F : \mathbb{R}^n \times ... \times \mathbb{R}^n \to \overline{\mathbb{R}}$, $F(x^1, ..., x^k) = \sum_{i=1}^k f_i(x^i)$. By Corollary 1 we know that F is almost convex. We have $\text{dom}(F) = \text{dom}(f_1) \times ... \times \text{dom}(f_k)$, so $\text{ri}(\text{dom}(F)) = \text{ri}(\text{dom}(f_1)) \times ... \times \text{ri}(\text{dom}(f_k))$. Consider also the linear operator $A : \mathbb{R}^n \to \mathbb{R}^n \times ... \times \mathbb{R}^n$, $Ax = (\underbrace{x, ..., x}_{k})$. The existence of the element $x' \in \cap_{i=1}^k \text{ri}(\text{dom}(f_i))$ gives $(x', ..., x') \in \text{ri}(\text{dom}(F))$, so $Ax' \in \text{ri}(\text{dom}(F))$. By Theorem 3 we have for any $p \in \mathbb{R}^n$

$$(F \circ A)^*(p) = \inf \{ F^*(q) : A^*q = p \}, \tag{5}$$

with the infimum attained at some $\bar{q} \in \mathbb{R}^n \times ...\mathbb{R}^n$. For the conjugates above we have for any $p \in \mathbb{R}^n$

$$(F \circ A)^*(p) = \sup_{x \in \mathbb{R}^n} \left\{ p^T x - \sum_{i=1}^k f_i(x) \right\} = \left(\sum_{i=1}^k f_i \right)^* (p)$$

and for every $q = (p^1, ..., p^k) \in \mathbb{R}^n \times ... \times \mathbb{R}^n$,

$$F^*(q) = \sup_{\substack{x^i \in \mathbb{R}^n, \\ i=1,...,k}} \left\{ \sum_{i=1}^k (p^i)^T x^i - \sum_{i=1}^k f_i(x^i) \right\} = \sum_{i=1}^k f_i^*(p^i),$$

so, as $A^* q = \sum_{i=1}^k p^i$, (5) delivers (4). □

In [16] the formula (4) is given assuming the functions f_i, $i = 1, ..., k$, proper and convex and the intersection of the relative interiors of their domains non-empty. We have proven above that it holds even under the much weaker than convexity assumption of almost convexity imposed on these functions, when the other two conditions, i.e. their properness and the non-emptiness of the intersection of the relative interiors of their domains, stand. As the following assertion states, the formula is valid under the assumption regarding the domains also when the functions are proper and nearly convex, provided that the relative interiors of their epigraphs are non-empty.

Corollary 3. *If $f_i : \mathbb{R}^n \to \overline{\mathbb{R}}$, $i = 1, ..., k$, are proper nearly convex functions whose epigraphs have non-empty relative interiors and with their domains satisfying $\cap_{i=1}^k \mathrm{ri}(\mathrm{dom}(f_i)) \neq \emptyset$, then for any $p \in \mathbb{R}^n$ one has*

$$(f_1 + ... + f_k)^*(p) = \inf \left\{ \sum_{i=1}^k f_i^*(p_i) : \sum_{i=1}^k p_i = p \right\},$$

with the infimum attained.

Next we show that another important conjugacy formula remains true when imposing almost convexity (or near convexity) instead of convexity for the functions in discussion.

Theorem 5. *Given two proper almost convex functions $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ and $g : \mathbb{R}^m \to \overline{\mathbb{R}}$ and the linear operator $A : \mathbb{R}^n \to \mathbb{R}^m$ for which is guaranteed the existence of some $x' \in \mathrm{ri}(\mathrm{dom}(f))$ satisfying $Ax' \in \mathrm{ri}(\mathrm{dom}(g))$, one has for all $p \in \mathbb{R}^n$*

$$(f + g \circ A)^*(p) = \inf \left\{ f^*(p - A^* q) + g^*(q) : q \in \mathbb{R}^m \right\}, \tag{6}$$

with the infimum attained.

Proof. Consider the linear operator $B : \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^m$ defined by $Bz = (z, Az)$ and the function $F : \mathbb{R}^n \times \mathbb{R}^m \to \overline{\mathbb{R}}$, $F(x, y) = f(x) + g(y)$. By Theorem 2, F is an almost convex function and we have $\mathrm{dom}(F) = \mathrm{dom}(f) \times \mathrm{dom}(g)$. From the hypothesis one gets

$$\begin{aligned} Bx' = (x', Ax') &\in \mathrm{ri}(\mathrm{dom}(f)) \times \mathrm{ri}(\mathrm{dom}(g)) \\ &= \mathrm{ri}(\mathrm{dom}(f) \times \mathrm{dom}(g)) = \mathrm{ri}(\mathrm{dom}(F)), \end{aligned}$$

thus $Bx' \in \text{ri}(\text{dom}(F))$. Theorem 3 is applicable, leading to

$$(F \circ B)^*(p) = \inf \big\{ F^*(q_1, q_2) : B^*(q_1, q_2) = p, (q_1, q_2) \in \mathbb{R}^n \times \mathbb{R}^m \big\}$$

where the infimum is attained for any $p \in \mathbb{R}^n$. Since for each $p \in \mathbb{R}^n$

$$\begin{aligned}(F \circ B)^*(p) &= \sup_{x \in \mathbb{R}^n} \big\{ p^T x - F(B(x)) \big\} = \sup_{x \in \mathbb{R}^n} \big\{ p^T x - F(x, Ax) \big\} \\ &= \sup_{x \in \mathbb{R}^n} \big\{ p^T x - f(x) - g(Ax) \big\} = (f + g \circ A)^*(p),\end{aligned}$$

$F^*(q_1, q_2) = f^*(q_1) + g^*(q_2)$ $\forall (q_1, q_2) \in \mathbb{R}^n \times \mathbb{R}^m$ and

$$B^*(q_1, q_2) = q_1 + A^* q_2 \ \forall (q_1, q_2) \in \mathbb{R}^n \times \mathbb{R}^m,$$

the relation above becomes

$$\begin{aligned}(f + g \circ A)^*(p) &= \inf \big\{ f^*(q_1) + g^*(q_2) : q_1 + A^* q_2 = p \big\} \\ &= \inf \big\{ f^*(p - A^* q_2) + g^*(q_2) : q_2 \in \mathbb{R}^m \big\},\end{aligned}$$

where the infimum is attained for any $p \in \mathbb{R}^n$, i.e. (6) stands. □

Corollary 4. *Let the proper nearly convex functions* $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ *and* $g : \mathbb{R}^m \to \overline{\mathbb{R}}$ *satisfying* $\text{ri}(\text{epi}(f)) \neq \emptyset$ *and* $\text{ri}(\text{epi}(g)) \neq \emptyset$ *and the linear operator* $A : \mathbb{R}^n \to \mathbb{R}^m$ *such that there is some* $x' \in \text{ri}(\text{dom}(f))$ *fulfilling* $Ax' \in \text{ri}(\text{dom}(g))$. *Then* (6) *holds for any* $p \in \mathbb{R}^n$ *and the infimum is attained.*

Remark 6. Assuming the hypotheses of Theorem 5, respectively, Corollary 4 fulfilled, one has from (6) that the following so-called subdifferential sum formula holds (for the proof see, for example, [5])

$$\partial(f + g \circ A)(x) = \partial f(x) + A^* \partial g(Ax) \ \forall x \in \text{dom}(f) \cap A^{-1}(\text{dom}(g)).$$

After weakening the conditions under which some widely-used formulae concerning the conjugation of functions take place, we switch to duality where we prove important results which hold even when replacing the convexity with almost convexity or near convexity.

The following duality statements are immediate consequences of Theorem 5, respectively Corollary 4, by taking $p = 0$ in (6).

Theorem 6. *Given two proper almost convex functions* $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ *and* $g : \mathbb{R}^m \to \overline{\mathbb{R}}$ *and the linear operator* $A : \mathbb{R}^n \to \mathbb{R}^m$ *for which is guaranteed the existence of some* $x' \in \text{ri}(\text{dom}(f))$ *satisfying* $Ax' \in \text{ri}(\text{dom}(g))$, *one has*

$$\inf_{x \in \mathbb{R}^n} \big[f(x) + g(Ax) \big] = -(f + g \circ A)^*(0) = \sup_{q \in \mathbb{R}^m} \big\{ -f^*(A^* q) - g^*(-q) \big\}, \quad (7)$$

with the supremum in the right-hand side attained.

Remark 7. This statement generalizes Corollary 31.2.1 in [16] as we take the functions f and g almost convex instead of convex and, moreover, we remove the closedness assumption required in the mentioned book. It is easy to notice that when f and g are convex there is no need to consider them moreover closed in order to obtain the formula (7).

Remark 8. Theorem 6 states actually the so-called strong duality between the primal problem (P_A) $\inf_{x\in\mathbb{R}^n}\big[f(x)+g(Ax)\big]$ and its Fenchel dual (D_A) $\sup_{q\in\mathbb{R}^m}\big\{-f^*(A^*q)-g^*(-q)\big\}$.

Using Proposition 1 and Theorem 6 we rediscover the assertion in Theorem 4.1 in [3], which follows.

Corollary 5. *Let $f:\mathbb{R}^n\to\overline{\mathbb{R}}$ and $g:\mathbb{R}^m\to\overline{\mathbb{R}}$ two proper nearly convex functions whose epigraphs have non-empty relative interiors and consider the linear operator $A:\mathbb{R}^n\to\mathbb{R}^m$. If there is an $x'\in\operatorname{ri}(\operatorname{dom}(f))$ such that $Ax'\in\operatorname{ri}(\operatorname{dom}(g))$, then (7) holds and the dual problem (D_A) has a solution.*

In the end we give a generalization of the well-known Fenchel's duality theorem (Theorem 31.1 in [16]). It follows immediately from Theorem 6, for A the identity mapping, thus we skip the proof.

Theorem 7. *Let f and g be proper almost convex functions on $\mathbb{R}^n$ with values in $\overline{\mathbb{R}}$. If $\operatorname{ri}(\operatorname{dom}(f))\cap\operatorname{ri}(\operatorname{dom}(g))\neq\emptyset$, one has*

$$\inf_{x\in\mathbb{R}^n}\big[f(x)+g(x)\big]=\sup_{q\in\mathbb{R}^n}\big\{-f^*(q)-g^*(-q)\big\},$$

with the supremum attained.

When f and g are nearly convex functions we have, as in Theorem 3.1 in [3], the following statement.

Corollary 6. *Let f and g be proper nearly convex functions on $\mathbb{R}^n$ with values in $\overline{\mathbb{R}}$. If $\operatorname{ri}(\operatorname{epi}(f))\neq\emptyset$, $\operatorname{ri}(\operatorname{epi}(g))\neq\emptyset$ and $\operatorname{ri}(\operatorname{dom}(f))\cap\operatorname{ri}(\operatorname{dom}(g))\neq\emptyset$, one has*

$$\inf_{x\in\mathbb{R}^n}\big[f(x)+g(x)\big]=\sup_{q\in\mathbb{R}^n}\big\{-f^*(q)-g^*(-q)\big\},$$

with the supremum attained.

Remark 9. The last two assertions give actually the strong duality between the primal problem (P) $\inf_{x\in\mathbb{R}^n}\big[f(x)+g(x)\big]$ and its Fenchel dual (D) $\sup_{q\in\mathbb{R}^m}\big\{-f^*(q)-g^*(-q)\big\}$. In both cases we have weakened the initial assumptions required in [16] to guarantee strong duality between (P) and (D) by asking the functions f and g to be almost convex, respectively nearly convex, instead of convex.

Remark 10. Let us notice that the relative interior of the epigraph of a proper nearly convex function f with $\operatorname{ri}(\operatorname{dom}(f))\neq\emptyset$ may be empty (see for instance the function in Example 1(i)).

As proven in Example 1 there are almost convex functions which are not convex, so our Theorems 3–7 extend some results in [16]. An example given in [3] shows that also the Corollaries 2–6 generalize indeed the corresponding results from Rockafellar's book [16], as a nearly convex function whose epigraph has a non-empty interior is not necessarily convex.

5 Conclusions

After recalling the definitions of three generalizations of the convexity, we have shown that there are differences between the classes of almost convex functions and nearly convex functions, both of them being indeed larger than the one of the convex functions. Then we proved that the formulae of some conjugates, namely of the precomposition with a linear operator, of the sum of finitely many functions and of the sum between a function and the precomposition of another one with a linear operator hold even when the convexity assumptions are replaced by almost (or near) convexity. The last results we give show that the well-known duality statements due to Fenchel hold when the functions involved are taken only almost convex, respectively nearly convex.

References

1. Aleman A (1985) On some generalizations of convex sets and convex functions. Mathematica - Revue d'Analyse Numérique et de la Théorie de l'Approximation 14:1–6
2. Blaga L, Kolumbán J (1994) Optimization on closely convex sets. In: Generalized Convexity (Pécs, 1992), Lecture Notes in Economics and Mathematical Systems 405. Springer, Berlin, pp. 19–34
3. Boţ RI, Grad SM, Wanka G (2005) Fenchel's duality theorem for nearly convex functions. Preprint 13/2005, Faculty of Mathematics, Chemnitz University of Technology
4. Boţ RI, Kassay G, Wanka G (2005) Strong duality for generalized convex optimization problems. Journal of Optimization Theory and Applications 127(1):45–70
5. Boţ RI, Wanka G (2006) A weaker regularity condition for subdifferential calculus and Fenchel duality in infinite dimensional spaces. Nonlinear Analysis: Theory, Methods & Applications 64(12):2787–2804
6. Breckner WW, Kassay G (1997) A systematization of convexity concepts for sets and functions. Journal of Convex Analysis 4(1):109–127
7. Cobzaş Ş, Muntean I (1987) Duality relations and characterizations of best approximation for p-convex sets. L'Analyse Numérique et la Théorie de l'Approximation 16(2):95–108
8. Frenk JBG, Kassay G (2005) Introduction to convex and quasiconvex analysis. In: Handbook of Generalized Convexity and Generalized Monotonicity, Nonconvex Optimization and its Applications 76. Springer, New York, pp. 3–87

9. Frenk JBG, Kassay G, Lagrangian duality and convexlike functions. Journal of Optimization Theory and Applications (to appear)
10. Frenk JBG, Kassay G (2002) Minimax results and finite-dimensional separation. Journal of Optimization Theory and Applications 113(2):409–421
11. Hamel G (1905) Eine Basis aller Zahlen und die unstetigen Lösungen der Funktionalgleichung: $f(x+y) = f(x) + f(y)$. Mathematische Annalen 60:459–462
12. Jeyakumar V, Gwinner J (1991) Inequality systems and optimization. Journal of Mathematical Analysis and Applications 159(1):51–71
13. Kolumbán J, Blaga L (1990) On the weakly convex sets. Studia Universitatis Babeş-Bolyai Mathematica 35(4):13–20
14. Muntean I (1985) Support points of p-convex sets. In: Proceedings of the Colloquium on Approximation and Optimization (Cluj-Napoca, 1985). University of Cluj-Napoca, Cluj-Napoca, pp. 293–302
15. Paeck S (1992) Convexlike and concavelike conditions in alternative, minimax, and minimization theorems. Journal of Optimization Theory and Applications 74(2):317–332
16. Rockafellar RT (1970) Convex analysis. Princeton University Press, Princeton
17. Sach PH (2005) New generalized convexity notion for set-valued maps and application to vector optimization. Journal of Optimization Theory and Applications 125(1):157-179

Pseudomonotonicity of a Linear Map on the Interior of the Positive Orthant

Alberto Cambini[1] and Laura Martein[2]

[1] Department of Statistics and Applied Mathematics, University of Pisa,Italy
acambini@ec.unipi.it

[2] Department of Statistics and Applied Mathematics, University of Pisa,Italy
lmartein@ec.unipi.it

Summary. In this paper we will establish some necessary and/or sufficient conditions for both a nonsingular and a singular matrix A (interpreted as a linear map) to be pseudomonotone. The given results are in terms of the sign of the determinants of the principal submatrices and of the cofactors of A in the nonsingular case and in terms of the structure of A in the singular case. A complete characterization of pseudomonotonicity in terms of the coefficients of a 3×3 matrix is given and a method for constructing a merely pseudomonotone matrix is suggested.

Key words: Pseudomonotonicity, pseudoconvexity, principal submatrices, cofactors, linear map.

1 Introduction

Pseudomonotonicity of a linear map on the interior of the positive orthant is of particular interest for its relationship with the pseudoconvexity of a quadratic function over a cone and because of its relevance in complementarity problems. Starting from the pioneer work of Karamardian [2], this subject has been studied by several authors (see for instance [3], [5] [6], [7] and [8]) which have established various characterizations involving the bordered Hessian or the Moore-Penrose inverse or copositive and subdefinite matrices. The given approaches are interesting but not very useful in constructing or in testing pseudomonotonicity in an easily way. The aim of this paper is to move in this direction. By means of some reformulations of a result given by Crouzeix in [6], we will establish some necessary and/or sufficient conditions for both a nonsingular and a singular matrix A (interpreted as a linear map) to be pseudomonotone. The given results are in terms of the sign of the determinants of the principal submatrices and of the cofactors of A in the nonsingular case and in terms of the structure of A in the singular case. In particular, a complete characterization of pseudomonotonicity in terms of the coefficients of a 3×3 matrix is given and a method for constructing a merely pseudomonotone matrix is suggested.

2 Preliminary Results

In this section we will establish two main results which are fundamentals in characterizing the pseudomonotonicity of a linear map in the interior of the positive orthant.

Let B be a square matrix of order n. Following [1] we introduce a concise notation for determinants formed from elements of B:

$$B\begin{pmatrix} i_1 \ i_2 \ .... \ i_p \\ k_1 \ k_2 \ .... \ k_p \end{pmatrix} = \begin{vmatrix} b_{i_1k_1} \ b_{i_1k_2} \ ..... \ b_{i_1k_p} \\ .... \\ .... \\ b_{i_pk_1} \ b_{i_pk_2} \ ..... \ b_{i_pk_p} \end{vmatrix} \tag{1}$$

The determinants (1) in which $i_1 = k_1, i_2 = k_2, ... \ i_p = k_p$ are called principal minors of order p. In particular we will use the following notations: $d_p = B\begin{pmatrix} 1 \ 2 \ .... \ p \\ 1 \ 2 \ .... \ p \end{pmatrix}$; $|B_{ij}|$ is the determinant of the submatrix obtained by deleting the i-th row and the j-th column of B; $(-1)^{i+j}|B_{ij}|, i \neq j$, $i, j = 1, ..., n$ are the cofactors of B.

Applying the algorithm of Gauss, B is reduced, after p steps, to the following matrix:

$$C^{(p)} = \begin{bmatrix} b_{11} & b_{12} & ..... & b_{1p} & b_{1p+1} & .... & b_{1n} \\ 0 & c_{22} & ..... & c_{2p} & c_{2p+1} & .... & c_{2n} \\ ..... & ..... & ...... & ..... & ..... & ..... & ..... \\ 0 & 0 & ..... & c_{pp} & c_{pp+1} & .... & c_{pn} \\ 0 & 0 & ..... & 0 & c_{p+1p+1} & .... & c_{p+1n} \\ ..... & ..... & ...... & ..... & ..... & ..... & ..... \\ 0 & 0 & ..... & 0 & c_{np+1} & ...... & c_{nn} \end{bmatrix}. \tag{2}$$

This reduction can be carry out if and only if in the process the pivot elements $b_{11}, c_{22}, ..., c_{pp}$ turn out to be different from zero. Set

$$C_{n-p} = \begin{bmatrix} c_{p+1p+1} & .... & c_{p+1n} \\ & .... & \\ c_{np+1} & .... & c_{nn} \end{bmatrix}. \tag{3}$$

It is known [1] that

$$c_{kk} = \frac{d_k}{d_{k-1}}, \ 1 < k \leq p+1; \ c_{kk} = \frac{B\begin{pmatrix} 1 \ 2 \ .... \ p \ k \\ 1 \ 2 \ .... \ p \ k \end{pmatrix}}{d_p}, \ k > p+1 \tag{4}$$

and

$$c_{ij} = \frac{B\begin{pmatrix} 1\ 2\ ....\ p\ i \\ 1\ 2\ ....\ p\ j \end{pmatrix}}{d_p},\ i,j \geq p,\ i \neq j$$

$$c_{ij} = \frac{B\begin{pmatrix} 1\ 2\ ....i \\ 1\ 2\ ....\ j \end{pmatrix}}{d_{i-1}},\ 1 < i \leq p,\ j > i$$

In the particular case $p = n-2$, (3) reduces to

$$C_2 = \frac{1}{d_{n-2}} \begin{bmatrix} |B_{nn}| & |B_{nn-1}| \\ |B_{n-1n}| & |B_{n-1n-1}| \end{bmatrix}.$$

$$C_2 = \frac{1}{d_{n-2}} \begin{bmatrix} |B_{n-1n-1}| & -|B_{n-1n}| \\ -|B_{nn-1}| & |B_{nn}| \end{bmatrix}^{-1}.$$

More generally, if the algorithm of Gauss involves $p = n-2$ pivot elements c_{ss} with $s \neq i, j,\ i < j$ and B is a symmetric matrix, we obtain

$$C_2 = \frac{1}{d_{n-2}} \begin{bmatrix} |B_{jj}| & (-1)^{i+j}|B_{ij}| \\ (-1)^{i+j}|B_{ij}| & |B_{ii}| \end{bmatrix}. \quad (5)$$

where B_{jj} (B_{ii}) is the submatrix of order $n-1$ obtained by deleting the j-th (i-th) row and the j-th (i-th) column of B.
Our first result is stated in the following theorem.

Theorem 1. *Let B a symmetric matrix with $|B| < 0$ and assume that the principal submatrices of order $n-1$ are positive semidefinite. Then the principal submatrices of order $n-2$ are positive definite.*

Proof. Let B_{n-2} be a principal submatrix of order $n-2$. Without loss of generality we can assume that B_{n-2} is obtained by deleting the last two rows and the last two columns of B. If B_{n-2} is not positive definite, a pivot element c_{ii} in $C^{(n-2)}$ turns out to be zero so that we have $c_{ii} = c_{ii+1} = ... = c_{in-2} = 0$. Consider now the principal submatrix B_{n-1} obtained by deleting the last (second-last) row and the last (second-last) column of B. Since B_{n-1} is positive semidefinite we obtain $c_{in-1} = 0$ ($c_{in} = 0$). Consequently $C^{(n-2)}$ turns out to have a null row so that $|B| = 0$ and this is a contradiction.

Consider now the quadratic form $\psi(y) = y^T By$; $\psi(y)$ can be interpreted as a trinomial in the variable y_1 and its discriminant $\frac{\Delta_1}{4}(y_2, ..., y_n)$ turns out to be a quadratic form in the variables $y_2, ..., y_n$. Analogously, the discriminant $\frac{\Delta_2}{4}(y_3, ..., y_n)$ of $\frac{\Delta_1}{4}(y_2, ..., y_n)$ with respect to y_2 turns out to be a quadratic form in the variables $y_3, ..., y_n$. We obtain, after p steps, a quadratic form $\frac{\Delta_p}{4}(y_{p+1}, ..., y_n)$ depending to $n-p$ variables. This process can be carry out if and only if the coefficient of the variable $y_h, y = 2, ..., p+1$ in $\frac{\Delta_{h-1}}{4}$ turns out to be different from zero.

Denotes with $D_{n-1}, ..., D_{n-p}$ the symmetric matrices associate to the quadratic forms $\frac{\Delta_1}{4}, ..., \frac{\Delta_p}{4}$, respectively. When B is a symmetric matrix we are interested to find a relation between C_{n-p} and D_{n-p}. With this aim we will establish a preliminary result.

Let F be a symmetric matrix of order m and assume that $f_{11} \neq 0$. Setting $F = \begin{bmatrix} f_{11} & (f^*)^T \\ f^* & F_{m-1} \end{bmatrix}$ and performing a pivot operation on the first element $\frac{f_{11}}{\alpha}$ of the matrix $\frac{F}{\alpha}$, $\alpha \neq 0$, we obtain the matrix $\Gamma = \begin{bmatrix} \frac{f_{11}}{\alpha} & \frac{(f^*)^T}{\alpha} \\ 0 & H_{m-1} \end{bmatrix}$.

Let D_{m-1} be the symmetric matrix associate to the discriminant of the quadratic form $\psi(v) = v^T F v$ with respect to the variable v_1.

Set $\gamma_{11} = \frac{f_{11}}{\alpha}$, $v^T = (v_1, (v^*)^T)$.

Lemma 1. *We have* $D_{m-1} = -\alpha^2\ \gamma_{11}\ H_{m-1}$.

Proof. $\psi(v) = f_{11}v_1^2 + 2((f^*)^T v^*)v_1 + (v^*)^T F_{m-1} v^*$, so that $\frac{\Delta_1}{4}(y_2, ..., y_n) = ((f^*)^T v^*)^2 - f_{11}(v^*)^T F_{m-1} v^* = (v^*)^T (f^*(f^*)^T - f_{11}F_{m-1})v^*$ and thus $D_{m-1} = f^*(f^*)^T - f_{11}F_{m-1}$. On the other hand performing a pivot operation on the element γ_{11} of the matrix $\frac{F}{\alpha}$ we obtain $H_{m-1} = -\frac{f^*}{f_{11}}(f^*)^T + \frac{F_{m-1}}{\alpha} = -\frac{D_{m-1}}{\alpha f_{11}}$. Since $f_{11} = \alpha\gamma_{11}$ the thesis is achieved.

Now we are able to state the following fundamental result (the convention $z^j = 0$ if $j < 0$ is used).

Theorem 2. *Let B be a symmetric matrix of order n and assume the validity of (2). Then*

$$D_{n-p} = -b_{11}^{p-2} \cdot d_2^{p-3} \cdot ... \cdot d_p \cdot C_{n-p} \tag{6}$$

Proof. Firstly we prove the following relation between D_{n-p} and C_{n-p} in terms of the pivot elements:

$$D_{n-p} = -(b_{11})^{2^{p-1}} \cdot (c_{22})^{2^{p-2}} \cdot ... \cdot c_{pp} \cdot C_{n-p}. \tag{7}$$

The proof is given by induction. When $p = 1$, applying Lemma 1 to the case $F = B$, $\alpha = 1$, we have $D_{n-1} = -b_{11}C_{n-1}$ so that (7) is verified. Assume the validity of (7) for $p = s$. Applying Lemma 1 to the case $F = \frac{D_{n-s}}{\alpha}$, $\alpha = (b_{11})^{2^{s-1}} \cdot (c_{22})^{2^{s-2}} \cdot ... \cdot c_{ss}$, we have $D_{n-s-1} = -\alpha^2 c_{s+1s+1} C_{n-s-1}$ that is (7) for $p = s+1$. Substituting $c_{kk} = \frac{d_k}{d_{k-1}}$ in (7) (see (4)) we obtain (6).

Remark 1. In Theorem 2 we have assumed, for the sake of simplicity, that the pivot elements are associate to indices $22, ..., pp$. If the algorithm of Gauss is applied with respect to the indices $i_1 i_1, i_2 i_2, ..., i_p i_p$, denoting with $d_{i_1 i_2 \dots i_k} = B\begin{pmatrix} i_1 & i_2 & .... & i_k \\ i_1 & i_2 & .... & i_k \end{pmatrix}$, (6) is substituted with

$$D_{n-p} = -b_{i_1 i_1}^{p-2} \cdot d_{i_1 i_2}^{p-3} \cdot ... \cdot d_{i_1 i_2 \dots i_p} \cdot C_{n-p}. \tag{8}$$

In particular, if the principal submatrices of order $n-1$ of B are positive semidefinite, we have $d_{i_1 i_2 \ldots i_k} > 0, k = 2, ..., p$ (see Theorem 1) and furthermore the diagonal elements of C_{n-p} are positive (see (4)).

Remark 2. Let B a symmetric matrix with $|B| < 0$ and such that the principal submatrices of order $n-1$ are positive semidefinite. Taking into account Theorem 1, we can perform $n-2$ steps of the algorithm of Gauss on the matrix B with respect to $n-2$ diagonal elements (pivots) choosen arbitrarily; so that we can calculate, according to (8), $n-2$ discriminants of the quadratic form $\psi(v) = v^T B v$ with respect to the $n-2$ variables associate to the choosen pivots.

3 Equivalent Formulations of Pseudomonotonicity

Through the paper we will use the following notations:

- $\Re^n_+ = \{x = (x_1, ..., x_n) \in \Re^n : x_i \geq 0, i = 1, ..., n\}$;
- $int\Re^n_+ = \{x = (x_1, ..., x_n) \in \Re^n : x_i > 0, i = 1, ..., n\}$.

Let A be a square matrix of order n. We recall that the linear map $A : \Re^n \to \Re^n$ is pseudomonotone on $int\Re^n_+$ if the following logical implication holds:

$$x, y \in \ int\Re^n_+, \quad (y-x)^T Ax > 0 \Rightarrow (y-x)^T Ay > 0 \tag{9}$$

The definition (9) is equivalent to the following implication ([6]):

$$x \in \ int\Re^n_+, \ v \in \Re^n, \quad v^T Ax = 0 \Rightarrow v^T Av \geq 0. \tag{10}$$

We will say that A is pseudomonotone on $int\Re^n_+$ if it is pseudomonotone on $int\Re^n_+$, as a linear map.
In order to describe the structure of a pseudomonotone matrix, in what follows we will utilize the following reformulation of (10):
a matrix A is pseudomonotone on $int\Re^n_+$ if and only if the following implication holds:

$$v \in \Re^n, \ A^T v = y, \quad x \in \ int\Re^n_+, \quad y^T x = 0 \ \Rightarrow y^T v \geq 0. \tag{11}$$

This reformulation will be used in section 5 to find an explicit form of the scalar product $y^T v$ with the aim to study the pseudomonotonicity of a singular matrix.
When A is a nonsingular matrix, (11) assume one of the simple forms stated in the following theorem.

Theorem 3. *The nonsingular matrix A is pseudomonotone on $int\Re^n_+$ if and only if (12) or (13) holds:*

$$x \in \ int\Re^n_+, \quad y^T x = 0 \ \Rightarrow y^T A^{-1} y \geq 0 \tag{12}$$

$$y^T A^{-1} y < 0 \Rightarrow \ y \in \ \Re^n_+ \cup \Re^n_- \tag{13}$$

Proof. The equivalence between (11) and (12) follows by noting that $A^T v = y$ if and only if $v = (A^T)^{-1} y = (A^{-1})^T y$ and that $y^T (A^{-1})^T y = y^T A^{-1} y$.
The equivalence between (12) and (13) follows taking into account that for any fixed $y \notin \Re^n_+ \cup \Re^n_-$ there exists some $x \in int\Re^n_+$ such that $y^T x = 0$.

Remark 3. Condition (12) implies that the study of the pseudomonotonicity of a nonsingular linear map is equivalent to the study of the sign of a quadratic function subject to a complementarity condition; condition (13) implies that a matrix A is pseudomonotone if and only if the quadratic form $y^T A^{-1} y$ is semidefinite positive or its assume negative values for vectors having non null components of the same sign.

Let us note that any matrix A such that $M = \frac{A+A^T}{2}$ is positive semidefinite verifies (10) so that the main problem in finding pseudodomonotone matrices is related to matrices for which M is not positive semidefinite. We will refer to these last matrices as **merely pseudomonotone matrices**.
By means of (10) we will establish an important property of a pseudomonotone matrix. Denote with A_k a principal submatrix of A of order k, that is the submatrix obtained by removing $n-k$ rows and $n-k$ columns of the same indices. In what follows we will consider proper principal matrices, that is $k \neq n$. The following theorem holds.

Theorem 4. *If A is pseudomonotone on $int\Re^n_+$ then A_k, $k = 2, .., n-1$ is pseudomonotone on $int\Re^k_+$.*

Proof. Obviously, it is sufficient to give the proof for $k = n-1$ (by iterations, the proof is valid also for $k = n-2, n-3, .., 2$) . Without loss of generality set $A = \begin{bmatrix} A_{n-1} & a \\ c^T & \alpha \end{bmatrix}$ with $a, c \in \Re^{n-1}$, $\alpha \in \Re$ and assume that A_{n-1} is not pseudomonotone on $int\Re^{n-1}_+$. Then there exist $x^* \in int\Re^{n-1}_+$, $v^* \in \Re^{n-1}$ such that

$$(v^*)^T A_{n-1} x^* = 0 \ \ and \ \ (v^*)^T A_{n-1} v^* < 0. \tag{14}$$

If $(v^*)^T a = 0$, setting $x = ((x^*)^T, 0)^T \in int\Re^n_+$, $v = ((v^*)^T, 0)^T \in \Re^n$ we have $v^T A x = 0$, $v^T A v = (v^*)^T A_{n-1} v^* < 0$ and this contradicts the pseudomonotonicity of A.
If $(v^*)^T a \neq 0$ we can suppose, without loss of generality, $(v^*)^T a > 0$ substituting in (14) v^* with $-v^*$ if necessary. Setting $x = (\beta x^* + v^*, t)^T$, $v = ((v^*)^T, 0)^T$, we have $v^T A x = (v^*)^T A_{n-1} v^* + (v^*)^T a t$, so that $v^T A x = 0$ for $t^* = -\frac{(v^*)^T A_{n-1} v^*}{(v^*)^T a} > 0$. It follows that $\overline{x} = ((\beta x^* + v^*)^T, t^*)^T$ belongs to $int\Re^n_+$ for β large enough, $v^T A \overline{x} = 0$, $v^T A v < 0$ and this contradicts, once again, the pseudomonotonicity of A.

Let us note that the converse of Theorem 4 does not hold as it is shown in the following example.

Example 1. Consider the matrix $A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{bmatrix}$. It is easy to verify that the principal submatrices $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 4 \\ 0 & 1 \end{bmatrix}$ are pseudomonotone on $int\Re^2_+$ but A is not pseudomonotone on $int\Re^3_+$ since for $x^T = (10, 1, 1)$, $v^T = (1, -5, 13)$ we have $v^T Ax = 0$ and $v^T Av < 0$ contradicting (10).

4 Pseudomonotonicity of a Nonsingular Matrix

In this section we will characterize the pseudomonotonicity of a nonsingular matrix by means of our reformulations (12) and (13).

4.1 Pseudomonotonicity of a Nonsingular 2 × 2 Matrix

This case has been studied recently in [9] and the following characterization in terms of the elements of the matrix are obtained (further details can be also found in [4]).

Theorem 5. *Consider the nonsingular matrix* $A = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$. *Then i) and ii) hold.*
i) A is merely pseudomonotone on $int\Re^2_+$ *if and only if (15) or (16) holds.*

$$| A |> 0,\ a \geq 0,\ d \geq 0,\ b + c \geq 0,\ \Delta = (b + c)^2 - 4ad > 0 \tag{15}$$

$$| A |< 0,\ a \leq 0,\ d \leq 0,\ b + c \leq 0. \tag{16}$$

ii) A is pseudomonotone (not merely) on $int\Re^2_+$ *if and only if (17) holds.*

$$a \geq 0,\ d \geq 0,\ \Delta = (b + c)^2 - 4ad \leq 0. \tag{17}$$

4.2 Pseudomonotonicity of a Nonsingular Matrix of Order (≥ 3)

The following theorem establishes a necessary condition for a nonsingular matrix to be pseudomonotone in terms of the symmetric matrix associate to its inverse.

Theorem 6. *Let A be a nonsingular matrix of order* $n \geq 3$.
If A is pseudomonotone on $int\Re^n_+$ *then all proper principal submatrices of* $B = \frac{A^{-1}+(A^{-1})^T}{2}$ *are positive semidefinite.*

Proof. It is sufficient to prove the theorem for a principal submatrix of order $n-1$. Set $A^{-1} = \begin{bmatrix} B_{n-1} & b \\ c^T & \beta \end{bmatrix}$ with $b, c \in \Re^{n-1},\ \beta \in \Re$. We must prove that the quadratic form associate to B_{n-1} is non negative that is $(y^{n-1})^T B_{n-1} y^{n-1} \geq$

$0, \forall\ y^{n-1}$.

Consider $x \in int\Re^n_+$ and $y \in \Re^n$ such that $y_n = \frac{-\sum_{i=1}^{n-1} x_i y_i}{x_n},\ y_i \in \Re$. From (12) we have

$$y^T A^{-1} y = (y^{n-1})^T B_{n-1} y^{n-1} + (y^{n-1})^T (b+c) y_n + \beta y_n^2 \geq 0. \qquad (18)$$

The thesis trivially holds if $b+c=0$ and $\beta=0$. Assume the existence of $(\overline{y})^{n-1}$ such that $(\overline{y}^{n-1})^T B_{n-1} \overline{y}^{n-1} < 0$. Setting $x_i = 1,\ y_i = \overline{y}_i,\ i = 1,...,n-1$ and taking into account that $y_n \to 0$ when $x_n \to +\infty$, we can choose $\overline{y}_n$ such that $(\overline{y}^{n-1})^T (b+c)\overline{y}_n + \beta \overline{y}_n^2 < -(\overline{y}^{n-1})^T B_{n-1} \overline{y}^{n-1}$ and this contradicts (18).

The following example shows that the necessary pseudomonotonicity condition stated in Theorem 6 is not sufficient.

Example 2. Consider the matrix $A = \begin{bmatrix} -3 & -\frac{3}{2} & 1 \\ -\frac{3}{2} & -\frac{1}{2} & \frac{1}{2} \\ 1 & \frac{1}{2} & 0 \end{bmatrix}$. We have $B = A^{-1} = \begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & 0 \\ 1 & 0 & 3 \end{bmatrix}$. It is easy to verify that the principal submatrices of B are positive semidefinite but A is not pseudomonotone on $int\Re^3_+$ since for $x^T = (1,2,4)$, $y^T = (-2,-1,1)$ we have $y^T x = 0$ and $y^T A^{-1} y < 0$ contradicting (12).

Now we are able to state a necessary and sufficient pseudomonotonicity condition for a nonsingular matrix.

Theorem 7. *Let A be a nonsingular matrix of order $n \geq 3$ and let A^{-1} be its inverse. Then A is pseudomonotone on $int\Re^n_+$ if and only if i) or ii) holds.*
i) $B = \frac{A^{-1}+(A^{-1})^T}{2}$ is positive semidefinite;
ii) $detB < 0$, the principal submatrices of B of order $n-1$ are positive semidefinite and all the cofactors $(-1)^{i+j} \mid B_{ij} \mid,\ i \neq j$ are positive.
Furthermore A is merely pseudomonotone if and only if ii) holds.

Proof. If $detB \geq 0$, from Theorem 3 and Theorem 6, the pseudomonotonicity of A is equivalent to condition i). Assume now that the principal submatrices of B of order $n-1$ are positive semidefinite and furthermore that $detB < 0$. Referring to Section 2 and in particular to Remark 1 and Remark 2, setting $\psi(y) = y^T B y$, we can calculate the $n-2$ discriminants $\frac{\Delta_1}{4}(y_2,...,y_n)$, $\frac{\Delta_2}{4}(y_3,...,y_n)$, ..., $\frac{\Delta_{n-2}}{4}(y_i,y_j)$.
Since B is indefinite, there exists a solution $y^* = (y_1^*, y_2^*, ..., y_n^*)$ of the inequality $\psi(y) < 0$. It follows that y_1^* is a solution of $\psi(y_1, y_2^*, ..., y_n^*) < 0$ and since the coefficient of y_1 is negative (see Remark 1) necessarily we have $\frac{\Delta_1}{4}(y_2^*,...,y_n^*) > 0$. It follows that y_2^* is a solution of $\psi(y_2, y_3^*, ..., y_n^*) > 0$ and since the coefficient of y_2 is negative (see Remark 1) necessarily we have $\frac{\Delta_2}{4}(y_3^*,...,y_n^*) > 0$. After $n-2$ times we arrive to have $\frac{\Delta_{n-2}}{4}(y_i^*, y_j^*) > 0$, that is (see (5))

$$- |B_{jj}| (y_i^*)^2 + 2(-1)^{i+j} |B_{ij}| y_i^* y_j^* - |B_{ii}| (y_j^*)^2 > 0, \ i \neq j, \tag{19}$$

where $|B_{jj}| \geq 0, \ |B_{ii}| \geq 0$. If A is pseudomonotone then from (13) we have $y^* \in \Re_+^n \cup \Re_-^n$ so that necessarily we have $(-1)^{i+j} |B_{ij}| > 0$ and this implies ii) taking into account Theorem 6. Viceversa, the validity of ii) implies $y_i^* y_j^* > 0, \ i,j = 1,...,n, \ i \neq j$; consequently $y^* \in \Re_+^n \cup \Re_-^n$ and thus A is pseudomonotone.
The last assertion of the theorem follows by noting that $\frac{A^{-1}+(A^{-1})^T}{2}$ is semidefinite positive if and only if $\frac{A+A^T}{2}$ is semidefinite positive. The proof is complete.

Remark 4. In order to deduce $y^* \in \Re_+^n \cup \Re_-^n$ in the proof given in Theorem 7 it is sufficient to consider the couples $(i, i+1), i = 1, 2, ..., n-1$. Consequently, the condition "the cofactors $(-1)^{i+j} \mid B_{ij} \mid, \ i \neq j$ are positive" in ii) can be substituted with the less restrictive condition "the cofactors $- \mid B_{ii+1} \mid, \ i = 1, ..., n-1$ are positive".

4.3 On Constructing a Nonsingular Pseudomonotone Matrix of Order 3

In this subsection we suggest, as an application of Theorem 7, the way of constructing a merely pseudomonotone matrix of order 3.
First Step
Construct a symmetric matrix B such that:
$B_{11}, \ B_{22}, \ B_{33}$ are positive semidefinite submatrices;
$\mid B_{12} \mid < 0, \ \mid B_{13} \mid > 0, \ \mid B_{23} \mid < 0; \mid B \mid < 0$.
Second Step
Any nonsingular matrix A such that $\frac{A^{-1}+(A^{-1})^T}{2} = B$ is merely pseudomonotone.

Example 3. Consider the symmetric matrix $B = \begin{bmatrix} 1 & -2 & -1 \\ -2 & 4 & 0 \\ -1 & 0 & 3 \end{bmatrix}$.
The proper principal submatrices $B_{11}, \ B_{22}, \ B_{33}$ are positive semidefinite and furthermore $\mid B_{12} \mid = -6 < 0, \ \mid B_{13} \mid = 4 > 0, \ \mid B_{23} \mid = -2 < 0$.
The matrices $A^{-1} = B, \ A^{-1} = \begin{bmatrix} 1 & -6 & -3 \\ 2 & 4 & -2 \\ 1 & 2 & 3 \end{bmatrix}$ verify the condition $\frac{A^{-1}+(A^{-1})^T}{2}$ $= B$ so that $A = B^{-1} = \frac{1}{2} \begin{bmatrix} -6 & -3 & -2 \\ -3 & -1 & -1 \\ -2 & -1 & 0 \end{bmatrix}$, $A = \frac{1}{32} \begin{bmatrix} 8 & 6 & 12 \\ -4 & 3 & -2 \\ 0 & -4 & 8 \end{bmatrix}$ or
$A = k \begin{bmatrix} 8 & 6 & 12 \\ -4 & 3 & -2 \\ 0 & -4 & 8 \end{bmatrix}$, $k > 0$ are merely pseudomonotone.

5 Pseudomonotonicity of a Singular Linear Map

In this section we will study the pseudomonotonicity of a singular matrix by means of our reformulation (11).
First of all we will prove that if A is a singular pseudomonotone matrix with $rank(A) \geq 2$, then there exists a proper principal submatrix having the same rank of A. More exactly we have the following theorem.

Theorem 8. *Let A be a singular matrix of order $n \geq 3$ with $rank(A) = s \geq 2$ and let k be the maximum of the ranks of the proper principal submatrices of order s. If A is pseudomonotone on $int\Re^n_+$ then $s = k$.*

Proof. Assume that $rank(A) = s > k$ and consider the system $A^T v = y$; without loss of generality we can suppose $rank(A_k^T) = k$.
From Kronecker's theorem there exists a submatrix of order s containing A_k^T. Set $I = \{1, .., k, i_1, ..., i_{s-k}\}$, with $k < i_1 < .. < i_{s-k}$, the indices of the s rows in A^T which are linearly independent and consider the matrix $A_{s,n}^T$ whose rows are a_i, $i \in I$. Since the proper principal submatrix whose rows and columns are associated to the indices of I has rank k with $k < s$, the variables $v_{i_1}, ..., v_{i_{s-k}}$ cannot be explicited in the system $A^T v = y$.
For the sake of simplicity assume that $I = \{1, .., k, k+1, .., s\}$.
Set $I^* = \{1, 2, ..., n\}$, $I_k = \{1, 2, ..., k\}$, $I_{s-k} = \{k+1, k+2, ..., s\}$, $J = \{j_1, j_2, ..., j_{s-k}\}$ with $s < j_1$ and $H = I^* \setminus (I_k \cup I_{s-k} \cup J)$ (H may be empty). Partitioning the vector v according to the described indices, that is $v = (v^k, v^{s-k}, v^{|J|}, v^{|H|})^T$ with $v^k = (v_1, ..., v_k)^T$, $v^{s-k} = (v_{k+1}, ..., v_s)^T$, $v^{|J|} = (v_{j_1}, ..., v_{j_{s-k}})^T$, the solutions of the system $A^T v = y$ are of the kind

$$v^k = B_1 y^k + C_1 v^{s-k} + D_1 v^{|H|} + E_1 y^{s-k} \tag{20}$$

$$v^{|J|} = B_2 y^k + D_2 v^{|H|} + E_2 y^{s-k} \tag{21}$$

$$y^{n-s} = F y^s \tag{22}$$

where $y^k = (y_1, ..., y_k)^T$, $y^{s-k} = (y_{k+1}, ..., v_s)^T$, $y^{n-s} = (y_{s+1}, ..., y_n)^T$ and the dimensions of the matrices B_i, C_i, D_i, E_i, F, $i = 1, 2$, are according the product rule between matrices (notes that $v^{|J|}$ does not depend from v^{s-k} since the variables $v_{k+1}, ..., v_s$ cannot be explicited in the system $A^T v = y$).
Taking into account (20) and (21), $y^T v$ is of the following kind:

$$y^T v = ((y^{s-k})^T + (y^k)^T C_1) v^{s-k} + \psi(y, v^{|H|}). \tag{23}$$

Choose $\overline{y} = (1, 0, .., 0, y_{k+1}, 0, .., 0, y^{n-s})$ with y^{n-s} verifying (22) and $y_{k+1} < 0$, such that $y_{k+1} + c_{11} < 0$ where c_{11} is the first element of C_1. Then it is possible to find $x \in int\Re^n_+$ such that $\overline{y}^T x = x_1 + y_{k+1} x_{k+1} + (y^{n-s})^T x^{n-s} = 0$. On the other hand we have $\overline{y}^T v = (y_{k+1} + c_{11}) v_{k+1} + \psi(\overline{y}, v^{|H|}) \to -\infty$ when $v_{k+1} \to +\infty$ and this contradicts (11). The proof is complete.

Remark 5. The assumption $rank(A) \geq 2$ in Theorem 8 cannot be relaxed. Consider for instance the matrix $A = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. It is easy to prove that A is pseudomonotone on $int\Re^3_+$ but $rank(A) = 1$ while $k = 0$.

Now we are able to establish a necessary and sufficient condition for a singular matrix A to be pseudomonotone on $int\Re^n_+$ in terms of its structure. With this aim let $k \geq 2$ be the rank of A. From Theorem 8 there exists a proper principal submatrix A_k with $rank(A_k) = k$. Without loss of generality assume that A_k is obtained by deleting the last $n - k$ rows and columns of A. Set $A = \begin{bmatrix} A_k & A_{k,n-k} \\ A_{n-k,k} & A_{n-k,n-k} \end{bmatrix}$, $x = (x^k, x^{n-k})^T$ and $y = (y^k, y^{n-k})^T$.
The following theorem holds.

Theorem 9. *Let A be a singular matrix with $rank(A) = rank(A_k) = k \geq 2$. A is pseudomonotone on $int\Re^n_+$ if and only if the following conditions hold:*

$$A_{n-k,n-k} = A_{n-k,k} A_k^{-1} A_{k,n-k} \tag{24}$$

$$A_{n-k,k} = A^T_{k,n-k} (A_k^{-1})^T A_k \tag{25}$$

$$x \in int\Re^n_+, \ (y^k)^T (x^k + A_k^{-1} A_{k,n-k} \ x^{n-k}) = 0 \Rightarrow (y^k)^T A_k^{-1} y^k \geq 0 \tag{26}$$

Proof. First of all observe that (24) follows from the assumption $rank(A) = rank(A_k) = k$.
From the system $A^T v = y$, we have

$$v^k = (A_k^{-1})^T y^k - (A_k^{-1})^T A^T_{n-k,k} v^{n-k} \tag{27}$$

$$(y^{n-k})^T = (y^k)^T A_k^{-1} \ A_{k,n-k} \tag{28}$$

so that the pseudomonotonicity of A is equivalent to the following conditions:

$$y^T x = (y^k)^T [x^k + A_k^{-1} \ A_{k,n-k} x^{n-k}] = 0, \tag{29}$$

$$y^T v = (y^k)^T A_k^{-1} y^k + (y^k)^T \Gamma v^{n-k} \geq 0, \tag{30}$$

where $\Gamma = A_k^{-1} \ A_{k,n-k} - (A_k^{-1})^T A^T_{n-k,k}$. We are going to prove that Γ is the null matrix. With this aim let $\gamma^j, \ j = 1,..,n-k$ be the columns of Γ and assume that $\gamma_{sj} \neq 0$ for some $s \in \{1,...,k\}$. Setting in (29), (30) $y_h = 0, \ h \neq 1, s$, with obvious notations we obtain

$$y_1(x_1 + \sum_{i=k+1}^{n} \beta_i x_i) + y_s(x_s + \sum_{i=k+1}^{n} \delta_i x_i) = 0 \tag{31}$$

$$\alpha_1 y_1^2 + \alpha_2 y_1 y_s + \alpha_3 y_s^2 + \sum_{i=1}^{n-k} (y_1 \gamma_1^i + y_s \gamma_s^i) v_{k+i} \geq 0 \tag{32}$$

From (31) we can choose $x_1, ..., x_n > 0$ such that $y_1\gamma_1^j + y_s\gamma_s^j \neq 0$ and this is absurd since v_{k+j} is a free variable and thus the inequality (32) cannot be verified for all $v_{k+j} \in \Re$. Then necessarily we have $\gamma^j = 0, j = 1, .., n-k$ and this implies $\Gamma = 0$; consequently (25) and (26) hold. Vice versa (25) and (26) imply (29) and (30). The proof is complete.

The assumption $rank(A) = k$ implies the existence of a $k \times (n-k)$ matrix $C_{k,n-k}$ such that $A_{k,n-k} = A_k C_{k,n-k}$. From (25), (24) we have: $A_{n-k,k} = C^T_{k,n-k} A^T_k (A_k^{-1})^T A_k = C^T_{k,n-k} A_k$, $A_{n-k,n-k} = C^T_{k,n-k} A_k C_{k,n-k}$, so that A has the following structure:

$$A = \begin{bmatrix} A_k & A_k C_{k,n-k} \\ C^T_{k,n-k} A_k & C^T_{k,n-k} A_k C_{k,n-k} \end{bmatrix} \tag{33}$$

A simple interpretation of (33) is the following: if the $j-th$ column of the matrix $A_{k,n-k}$, $j = k+1, .., n$ is a linear combination of the columns of A_k with multipliers $\alpha_1, ..., \alpha_k$, then the $j-th$ row of the matrix $A_{n-k,k}$, $j = k+1, .., n$ is a linear combination of the rows of A_k with the same multipliers $\alpha_1, ..., \alpha_k$. Theorem 9 can be specified with respect to pseudomonotone matrices which are not merely (that is $\frac{A_k + A_k^T}{2}$ is semidefinite positive) and with respect to merely pseudomonotone matrices.

Theorem 10. *Let A be a singular matrix with $rank(A) = rank(A_k) = k \geq 2$.*
i) A is pseudomonotone (not merely) if and only if $\frac{A_k + A_k^T}{2}$ is positive semidefinite and (33) holds.
ii) A is merely pseudomonotone if and only if (33) holds, A_k is merely pseudomonotone and furthermore

$$x \in int\Re^n_+, \ (y^k)^T(x^k + C_{k,n-k}\ x^{n-k}) = 0 \Rightarrow (y^k)^T A_k^{-1} y^k \geq 0 \tag{34}$$

6 Special Cases

In this section, as an application of Theorem 10, we specialize condition (34) to the case $rank(A) = 2$ and, in particular, we characterize a 3×3 merely pseudomonotone matrix. At last, for the sake of completeness, we will consider also the case $rank(A) = 1$. Set:

- $A_2 = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$, $\begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix}$, the columns of $C_{2,n-2}$, $i = 3, .., n$;
- $J = \{i : \alpha_i \leq 0, \ \beta_i < 0\} \cup \{i : \alpha_i < 0, \ \beta_i \leq 0\}$;
- $I_1 = \{i : \alpha_i \geq 0, \ \beta_i \geq 0\}$;
- $I_2 = \{i : \alpha_i > 0, \ \beta_i < 0\}$;
- $I_3 = \{i : \alpha_i < 0, \ \beta_i > 0\}$;
- $\frac{-\beta_s}{\alpha_s} = \max_{i \in I_2} \frac{-\beta_i}{\alpha_i}$, $\frac{-\beta_k}{\alpha_k} = \min_{i \in I_3} \frac{-\beta_i}{\alpha_i}$;
- $\psi(m) = \frac{d}{|A_2|} m^2 - \frac{b+c}{|A_2|} m + \frac{a}{|A_2|}$, $\Delta = (b+c)^2 - 4ad > 0$;

- $0 \leq m_1 < m_2$ where $m_1,\ m_2$ are the roots of the trinomial $\psi(m)$ when $d \neq 0$.

The following theorem holds.

Theorem 11. *Let A be a singular matrix with rank $A = 2$. Then A is merely pseudomonotone if and only if the following conditions hold.*
i) $A = \begin{bmatrix} A_2 & A_2 C_{2,n-2} \\ C^T_{2,n-2} A_2 & C^T_{2,n-2} A_2 C_{2,n-2} \end{bmatrix}$;
ii) $\frac{a}{|A_2|} \geq 0, \quad \frac{d}{|A_2|} \geq 0, \quad \frac{b+c}{|A_2|} \geq 0, \quad \Delta = (b+c)^2 - 4ad > 0$;
iii) $J = \emptyset$;
iv) one of the following conditions holds:
- $a \cdot d \neq 0, \quad \frac{-\beta_s}{\alpha_s} \leq m_1 < m_2 \leq \frac{-\beta_k}{\alpha_k}$;
- $d \neq 0,\ a = 0, \quad I_2 = \emptyset, \quad \frac{-\beta_k}{\alpha_k} \geq \frac{b+c}{d}$;
- $d = 0,\ a \neq 0, \quad I_3 = \emptyset, \quad \frac{-\beta_s}{\alpha_s} \leq \frac{a}{b+c}$;
- $d = 0,\ a = 0, \quad I_2 = I_3 = \emptyset$.

Proof. Taking into account of ii) of Theorem 10 and i) of Theorem 5, it remains to prove that $iii), iv)$ are equivalent to (34) which becomes

$$x \in int\Re^n_+,\ y_1(x_1 + \sum_{t \in I_1} \alpha_t x_t + \sum_{i \in I_2} \alpha_i x_i + \sum_{j \in I_3} \alpha_j x_j + \sum_{l \in J} \alpha_l x_l) +$$
$$+ y_2(x_2 + \sum_{t \in I_1} \beta_t x_t + \sum_{i \in I_2} \beta_i x_i + \sum_{j \in I_3} \beta_j x_j + \sum_{l \in J} \beta_l x_l) = 0 \Rightarrow$$
$$\Rightarrow \psi(y_1, y_2) = \frac{d}{|A_2|} y_1^2 - \frac{b+c}{|A_2|} y_1 y_2 + \frac{a}{|A_2|} y_2^2 \geq 0 \quad (35)$$

Setting $m = \frac{y_1}{y_2}$, the inequality $\psi(y_1, y_2) \geq 0$ is equivalent to the inequality $\psi(m) = \frac{1}{y_2^2} \psi(y_1, y_2) \geq 0$ which is verified when:
- $m \notin (m_1, m_2)$ if $a \cdot d \neq 0$ or $d \neq 0, a = 0$;
- $m \leq \frac{a}{b+c}$ if $d = 0,\ a \neq 0$;
- $m \leq 0$ if $a = d = 0$.

iii), iv) $\Rightarrow$ (35)
From the left hand side of (35) we have

$$m = \frac{-x_2 - \sum_{t \in I_1} \beta_t x_t - \sum_{i \in I_2} \beta_i x_i - \sum_{j \in I_3} \beta_j x_j}{x_1 + \sum_{t \in I_1} \alpha_t x_t + \sum_{i \in I_2} \alpha_i x_i + \sum_{j \in I_3} \alpha_j x_j} \quad (36)$$

that is

$$m x_1 + x_2 + \sum_{t \in I_1} (m\alpha_t + \beta_t) x_t + \sum_{i \in I_2} (m\alpha_i + \beta_i) x_i + \sum_{j \in I_3} (m\alpha_j + \beta_j) x_j = 0. \quad (37)$$

Consider the case $d \neq 0$; we must prove that any m verifying (37) is such that $m \notin (m_1, m_2)$. A positive value of m is a solution of (37) if and only if for

some $i \in I_2$ and/or for some $j \in I_3$ we have $m\alpha_i + \beta_i < 0$, $m\alpha_j + \beta_j < 0$ that is $m \leq \frac{-\beta_s}{\alpha_s}$, $m \geq \frac{-\beta_k}{\alpha_k}$, so that $m \notin (\frac{-\beta_s}{\alpha_s}, \frac{-\beta_k}{\alpha_k})$; from *iv*) $m \notin (m_1, m_2)$.
In the case $d = 0, a \neq 0, I_3 = \emptyset$, any m verifying (37) is such that $m \leq \frac{-\beta_s}{\alpha_s} \leq \frac{a}{b+c}$, so that $\psi(m) \geq 0$.
In the case $a = d = 0,\ I_2 = I_3 = \emptyset$ (37) implies $m \leq 0$ so that $\psi(m) \geq 0$.
(35) ⇒ iii), iv)
Consider the principal submatrix of A obtained by deleting all rows and columns with index $k \neq 1, 2, i$; (36) becomes

$$m = \frac{-x_2 - \beta_i x_i}{x_1 + \alpha_i x_i} \tag{38}$$

Case $i \in J$.
If $\beta_i = 0, \alpha_i < 0$, setting in (38) $x_1 = x_2 = 1$ we have $m = \frac{-1}{1+\alpha_i x_i}$ so that $m \to 0$ when $x_i \to +\infty$ and $m \to +\infty$ when $x_i \to (-\frac{1}{\alpha_i})^+$. It follows that $m \in (0, +\infty)$ contradicting (35).
If $\beta_i < 0, \alpha_i \leq 0$, setting in (38) $x_i = 1,\ x_2 = \frac{-\beta_i}{2}$, we have $m \to 0^+$ when $x_1 \to +\infty$ and $m \to +\infty$ when $x_1 \to -\alpha_i^+$. It follows that $m \in (0, +\infty)$ contradicting (35) once again and thus $J = \emptyset$.
Case $i \in I_2$.
Since $m = \frac{-x_2 - \beta_i x_i}{x_1 + \alpha_i x_i} \leq \frac{-\beta_i x_i}{\alpha_i x_i} = \frac{-\beta_i}{\alpha_i}$ we have $\psi(m) \geq 0\ \ \forall m \leq \frac{-\beta_i}{\alpha_i}$ and this implies $\frac{-\beta_i}{\alpha_i} \leq m_1$ and , in particular, $\frac{-\beta_s}{\alpha_s} \leq m_1$. When $a = 0,\ d \neq 0$ we have $m_1 = 0$ so that $I_2 = \emptyset$ since $\frac{-\beta_s}{\alpha_s} > 0$.
Case $i \in I_3$.
Since $\frac{1}{m} = \frac{-x_1 - \alpha_i x_i}{x_2 + \beta_i x_i} \leq \frac{-\alpha_i}{\beta_i}$ we have $m \geq \frac{-\beta_i}{\alpha_i}$ so that the validity of $\psi(m) \geq 0\ \ \forall\ m \notin (m_1, m_2)$ implies $\frac{-\beta_i}{\alpha_i} \geq m_2$ and , in particular, $\frac{-\beta_k}{\alpha_k} \geq m_2$.
If $a \neq 0,\ d = 0,\ m \leq \frac{a}{b+c}$ and $m \geq \frac{-\beta_i}{\alpha_i}$ imply $I_3 = \emptyset$.
At last if $a = d = 0,\ \psi(m) \geq 0$ for $m \leq 0$ and since $\frac{-\beta_s}{\alpha_s}, \frac{-\beta_k}{\alpha_k}$ are positive, necessarily we have $I_2 = I_3 = \emptyset$. The proof is complete

As a particular case of the above theorem, we have the following characterization of a merely pseudomonotone matrix of order 3 with rank 2.

Corollary 1. *Consider the following matrix* $A = \begin{bmatrix} a & c & e \\ b & d & f \\ g & h & i \end{bmatrix}$ *where* $A_2 = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$ *is non singular and set* $(\alpha, \beta)^T = A_2^{-1}(e, f)^T$.
Then A *is merely pseudomonotone if and only if i), ii), iii) hold.*
i) $(g,\ h,\ i) = \alpha\ (a,\ c,\ e) + \beta(b,\ d,\ f)$.
ii) $\frac{a}{|A_2|} \geq 0,\quad \frac{d}{|A_2|} \geq 0,\quad \frac{b+c}{|A_2|} \geq 0,\quad \Delta = (b+c)^2 - 4ad > 0$.
iii) one of the following conditions holds:
$\alpha \geq 0,\ \beta \geq 0$;
$a \cdot d \neq 0,\ \alpha > 0,\ \beta < 0,\ \ \frac{-\beta}{\alpha} \leq m_1$;
$a \cdot d \neq 0,\ \alpha < 0,\ \beta \geq 0,\ \ \frac{-\beta}{\alpha} \geq m_2$;

$d \neq 0,\ a = 0,\ \alpha > 0,\ \beta < 0,\ \frac{-\beta}{\alpha} \geq \frac{b+c}{d};$
$d = 0,\ a \neq 0,\ \alpha < 0,\ \beta \geq 0,\ \frac{-\beta}{\alpha} \leq \frac{a}{b+c};$
$d = 0,\ a = 0,\ \alpha \geq 0,\ \beta \geq 0.$

Example 4. Consider the matrix $A = \begin{bmatrix} 0 & -1 & -2 \\ 3 & 2 & 1 \\ 6 & 5 & 4 \end{bmatrix}$. We have that $A_2 = \begin{bmatrix} 0 & -1 \\ 3 & 2 \end{bmatrix}$ is a 2×2 merely pseudomonotone principal submatrix, $\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = A_2^{-1} \begin{pmatrix} -2 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$, $-1(0,-1,-2) + 2(3,2,1) = (6,5,4)$, the trinomial $\psi(m) \geq 0$ is verified for $m \notin (0,1)$, $\frac{-\beta}{\alpha} = 2 > m_2 = 1$ and thus A is merely pseudomonotone.

Example 5. Consider the matrix $A = \begin{bmatrix} 1 & -1 & 6 & -3 \\ 6 & 4 & 26 & 2 \\ -1 & -9 & 4 & -17 \\ 11 & 9 & 46 & 7 \end{bmatrix}$. We have that $A_2 = \begin{bmatrix} 1 & -1 \\ 6 & 4 \end{bmatrix}$ is a 2×2 merely pseudomonotone principal submatrix, $\begin{pmatrix} \alpha_3 \\ \beta_3 \end{pmatrix} = A_2^{-1} \begin{pmatrix} 6 \\ 26 \end{pmatrix} = \begin{pmatrix} 5 \\ -1 \end{pmatrix}$, $\begin{pmatrix} \alpha_4 \\ \beta_4 \end{pmatrix} = A_2^{-1} \begin{pmatrix} -3 \\ 2 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$, $I_2 = \{3\}$, $I_3 = \{4\}$, $5(1,-1,6,-3) - 1 \cdot (6,4,26,2) = (-1,-9,4,-17)$, $-1 \cdot (1,-1,6,-3) + 2 \cdot (6,4,26,2) = (11,9,46,7)$, the trinomial $\psi(m) \geq 0$ is verified for $m \notin (\frac{1}{4}, 1)$, $\frac{-\beta_3}{\alpha_3} = \frac{1}{5} < m_1 = \frac{1}{4} < m_2 = 1 < \frac{-\beta_4}{\alpha_4} = 2$ and thus A is merely pseudomonotone.

At last, for sake of completeness, we consider the case rank(A)=1. A very simple sufficient condition for pseudomonotonicity can be found in [7], Proposition 3.3.
In what follows, the columns of A (which are all proportional), are denoted by $\alpha_1 a$, $\alpha_2 a$,..., $\alpha_n a$, $a \neq 0$, α_1, α_2,..., α_n not all zero.

Theorem 12. *Let A be a matrix with $rank(A) = 1$. Then A is pseudomonotone if and only if i) or ii) hold.*
i) $\alpha_i \geq 0,\ i = 1, ..., n$;
ii) $\frac{A+A^T}{2}$ is positive semidefinite.

Proof. . Setting $x = (x_1, ..., x_n)^T$, $\alpha^T = (\alpha_1, ..., \alpha_n)$, we have $Ax = (\alpha^T x)a$. If $\alpha_i \geq 0,\ i = 1, ..., n$ (this implies that at least one of them is positive), then $\alpha^T x > 0$ so that $v^T Ax = 0$ if and only if $v^T a = 0$ and this implies $v^T Av = 0\ \forall v \in \Re^n$ and thus A is pseudomonotone (see (10)). If there exist α_i, α_j with $\alpha_i \alpha_j < 0$ then it is possible to choose $x^* = (x_1^*, ..., x_n^*) \in int\Re_+^n$ such that $\alpha_1 x_1^* + \alpha_2 x_2^* + ..., \alpha_n x_n^* = 0$, so that $Ax^* = 0$, $v^T Ax^* = 0\ \ \forall v \in \Re^n$ and consequently A is pseudomonotone if and only if the quadratic form $v^T Av$ is semidefinite positive (see (10)).

Remark 6. Let us note that if $rank(A) = 1$ then $\frac{A+A^T}{2}$ is semidefinite (positive or negative) if and only if A is symmetric; in such a case A assume a particular form: let i be the index corresponding to the first non null column of A. Then $a^T = (0, .., 0, a_{ii}, ..., a_{in})$ with $a_{ii} \neq 0$ and we have $A = a_{ii} a^*(a^*)^T$ where $a^* = \frac{1}{a_{ii}} a = (0, .., 0, 1, \alpha_{i+1}, ..., \alpha_n)$.

Remark 6 and Theorem 12 imply the following characterization of the pseudomonotonicity of a matrix with $rank(A) = 1$.

Theorem 13. *Let A be a matrix with $rank(A) = 1$. Then A is merely pseudomonotone if and only if i) or ii) hold.*
i) A is not symmetric and $\alpha_i \geq 0, \ i = 1, ..., n$;
ii) A is symmetric, $a_{ii} < 0, \ \alpha_{i+s} \geq 0, \ s = 1, ..., n - i$ where a_{ii} is the first diagonal element different from zero.

In the following example we present some merely and not merely pseudomonotone matrices having rank equal to 1.

Example 6. Merely pseudomonotone matrices:

$$A = \begin{bmatrix} 1 & 2 & 4 \\ -3 & -6 & -12 \\ 5 & 10 & 20 \end{bmatrix}, \quad A = \mu \begin{bmatrix} 1 & \alpha & \beta \\ \alpha & \alpha^2 & \alpha\beta \\ \beta & \alpha\beta & \beta^2 \end{bmatrix}, \ \mu < 0, \ \alpha \geq 0, \beta \geq 0.$$

Some canonical form of not merely pseudomonotone matrices:

$$A = \mu \begin{bmatrix} 1 & \alpha & \beta \\ \alpha & \alpha^2 & \alpha\beta \\ \beta & \alpha\beta & \beta^2 \end{bmatrix}, \ \mu > 0, \quad A = \mu \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & \alpha \\ 0 & \alpha & \alpha^2 \end{bmatrix}, \ \mu > 0,$$

$$A = \mu \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \ \mu > 0.$$

References

1. Gantmacher FR (2000) The theory of Matrices, volume one, AMS Chelsea Publishing.
2. Karamardian S (1976) Complementarity problems over cones with monotone and pseudomonotone maps, Journal of Optimization Theory and Applications 18: 445-454.
3. Karamardian S, Schaible S, Crouzeix JP (1993) Characterizations of generalized monotone maps, Journal of Optimization Theory and Applications 76: 399-413.
4. Pini R, Schaible S (1994) Invariance properties of generalized monotonicity, Optimization 28: 211-222.
5. Crouzeix JP, Schaible S (1996) Generalized monotone affine maps, SIAM Journal Matrix Anal. Appl.17: 992-997.
6. Crouzeix JP (1998) Characterizations of generalized convexity and monotonicity, a survey. In Crouzeix P, Martinez-Legaz JE, Volle M (eds) Generalized convexity, generalized monotonicity, Kluwer Academic Publisher, Dordrecht.

7. Crouzeix JP, Hassouni A, Lahlou A, Schaible S (2000) Positive subdefinite matrices, generalized monotonicity, and linear complementarity problems, SIAM Journal Matrix Anal. Appl. 22: 66-85.
8. Hadjsavvas N, Schaible S (2005) Generalized monotone maps. In Hadjsavvas N, Komlosi S, Schaible S (eds) Handbook of Generalized Convexity and Generalized Monotonicity, Springer.
9. Marchi A, Martein L (2005) Pseudomonotonicity of an affine map: the two dimensional case, Technical report, Department of Statistics and Applied Mathematics, University of Pisa, Italy.

An Approach to Discrete Convexity and Its Use in an Optimal Fleet Mix Problem

Riccardo Cambini, Rossana Riccardi[1], and Ümit Yüceer[2]

[1] Dept. of Statistics and Applied Mathematics, University of Pisa
Faculty of Economics, Via Cosimo Ridolfi 10, 56124 Pisa, ITALY
cambric@ec.unipi.it - riccardi@ec.unipi.it

[2] Department of Industrial Engineering, Çankaya University
Öğretmenler Caddesi 14, Ankara 06530, TURKEY
uyuceer@cankaya.edu.tr

Summary. A notion of convexity for discrete functions is first introduced, with the aim to guarantee both the increasing monotonicity of marginal increments and the convexity of the sum of convex functions. Global optimality of local minima is then studied both for single variable functions and for multi variables ones. Finally, a concrete optimal fleet mix problem is studied, pointing out its discrete convexity properties.

Key words: Discrete programming, discrete convexity, optimal fleet mix.

1 Introduction

Concrete problems are often discrete, in the sense that the variables are defined over the set of integers. This happens, for instance, whenever the variables represent the number of units, such as workforce units, number of ambulances, number of vehicles, and so on.

Due to their importance in applications, discrete problems have been widely studied in the mathematical programming literature, especially from the algorithmic point of view. Some approaches to convexity properties of discrete functions have been proposed too (see for example [2, 3, 5]), pointing out the difficulty of this research field.

The aim of this paper is twofold. First, we propose an approach to the notion of convexity for discrete functions, with the aim to guarantee both the increasing monotonicity of marginal increments and the convexity of the sum of convex functions. Some properties of the defined class of functions are then studied, especially with respect to the global optimality of local minima. Then, a concrete problem of optimal fleet mix is analyzed. In particular, we consider a model involving both internal workforce units and external technicians; quality of service requirements and penalties for unfulfilled services are also considered. The model is then studied from a theoretical point of view, pointing out that some of the variables can be parametrically fixed to their optimal value, thus obtaining a parametrical discrete convex objective function.

2 Discrete Convex Functions

Convexity property has been widely used in Mathematics and in Economics due to its usefulness in optimization problems (both critical points and local minima are global optimum points). As it is well known, such a concept regards to functions defined over convex sets. Unfortunately, many applicative problems arising in Operations Research and in Management Science deal with integer programming. As a consequence, some efforts have been done in the literature in order to determine a convexity concept suitable for discrete problems.

In this section, we aim to propose a new definition of discrete convexity by using an approach different from the ones already appeared in the literature. In particular, our aim is to guarantee two properties which results to be useful in Economics and in applicative problems, that are the increasing monotonicity of marginal increments and the discrete convexity of the sum of two discrete convex functions.

2.1 A Brief Overview

For the sake of completeness, let us now briefly recall some of the results already appeared in the literature.

Favati and Tardella in [2] introduced the concept of integer convexity extending a function f, defined over a discrete rectangle $X \subset Z^n$, to a piecewise-linear function $\overline{f}$ defined over the convex hull of X, denoted with $co(X) \subseteq \Re^n$.

Definition 1. *A set $X \subset Z^n$ is said to be a discrete rectangle if there exist $a, b \in Z^n$ such that:*

$$X = \{x \in Z^n : a_i \leq x_i \leq b_i,\ i = 1, \ldots, n\}$$

Given a number $x \in \Re$ it is denoted with $N(x)$ the so called discrete neighborhood of x, defined as the set

$$N(x) = \{z \in Z^n : |x_i - z_i| < 1, i = 1, \ldots, n\}$$

Definition 2. *Let $f : X \to \Re$, where $X \subset Z^n$ is a discrete rectangle. The so called extension of f is the function $\overline{f} : co(X) \to \Re$ defined as follows:*

$$\overline{f}(y) = \min \left\{ \sum_{i=1}^{k} \alpha_i f(z^i) : z^i \in N(y), \sum_{i=1}^{k} \alpha_i z^i = y, \sum_{i=1}^{k} \alpha_i = 1, \alpha_i \geq 0 \right\}$$

where $y \in co(X)$ and $k = card(N(y))$. Then, function f is said to be integrally convex if its extension $\overline{f} : co(X) \to \Re$ is convex.

This discrete convexity property is not easy to be verified. In any case, the authors have been able to state some useful properties and a global optimality results, which deserve to be recalled for the sake of completeness.

Proposition 1. *Let $f, g : X \to \Re$, where X is a discrete rectangle, then*

$$\overline{f}(x) + \overline{g}(x) \leq (\overline{f+g})(x), \ \forall x \in co(X) \tag{1}$$

furthermore, if over any unit hypercube contained in $co(X)$ at least one of the functions $\overline{f}(x)$ and $\overline{g}(x)$ is linear, then

$$\overline{f}(x) + \overline{g}(x) = (\overline{f+g})(x), \ \forall x \in co(X) \tag{2}$$

Proposition 2. *A point $x \in X$ is a local minimum point for $\overline{f}$ over $co(X)$ if and only if it is a local minimum point for f over X.*

Proposition 3. *Let f be an integrally convex function on a discrete rectangle X. If x is a local minimum point for f over X, then x is a global minimum point.*

Unfortunately, the class of integrally convex functions is not closed under addition (see Proposition 1). However, if f and g are integrally convex on X and condition (2) holds, then $f + g$ is also integrally convex. This happens, for example, when f and g are submodular integrally convex functions.

A branch of the literature, then has concentrated its attention to this particular class of functions. Murota in [3] defines a concept of convexity for integer valued functions and investigates its relationship to submodularity. Yüceer in [5] establishes the equivalence of discrete convexity (in the sense of Yüceer) and increasing first forward differences of functions of a single variable.

Definition 3. *Let S be a subspace of a discrete n-dimensional space. A function $f : S \to \Re$ is discretely convex (in the sense of Yüceer) if for all $x, y \in S$ and for all $\alpha \in (0, 1)$*

$$\alpha f(x) + (1 - \alpha) f(y) \geq \min_{u \in N(z)} f(u)$$

where $N(z) = \{u \in S : \|u - z\| < 1\}$, $z = \alpha x + (1 - \alpha) y$ and $\|u\| = \max_{1 \leq i \leq n} \{|u_i|\}$.

Then, Yüceer proposes the concept of strong discrete convexity by imposing additional conditions to discrete convexity such as submodularity.

2.2 A New Approach

Let us now introduce a new notion of convexity for discrete functions by means of an approach not based neither on extended functions nor on submodular ones, hence different from the ones proposed in [2, 3, 5]. With this aim, let us first introduce the definition of discrete reticulum.

Definition 4. *Let $ret(x, y)$ be the set*

$$ret(x, y) = \{z \in Z^n : \min\{x_i, y_i\} \leq z_i \leq \max\{x_i, y_i\}, \ i = 1, .., n\}$$

A set $X \subseteq Z^n$ is said to be a discrete reticulum if $ret(x, y) \subseteq X$ $\forall x, y \in X$.

Obviously, any discrete rectangle is also a discrete reticulum; notice also that Z^n_+ is a discrete reticulum but not a discrete rectangle.

From now on the infinite norm will be used to determine the length of vectors, that is to say that the norm of an n-dimensional vector x will be denoted as follows:

$$\|x\| = \|x\|_\infty = \max_{i=1,\ldots,n} |x_i|$$

As usual, if $\|x\| = 1$ then x is said to be an unitary vector.

The following further notations will be used in the rest of the paper.

Definition 5. *Let $f : X \to \Re$, where $X \subset Z^n$ is a discrete reticulum. The following first and second order differences are introduced:*

$$\Delta f(x; v) = f(x + v) - f(x) \tag{3}$$

$$\Delta^2 f(x; v) = f(x + 2v) - 2f(x + v) + f(x) \tag{4}$$

where $x \in X$, $v \in Z^n$ with $\|v\| = 1$, and $x + 2v \in X$.

Notice that it is $\Delta^2 f(x; v) = \Delta f(x+v; v) - \Delta f(x; v)$. Let us now introduce the definition of discrete convex function.

Definition 6. *Let $f : X \to \Re$, where $X \subset Z^n$ is a discrete reticulum. Function f is said to be a discrete convex function if for all $x \in X$, for all $v \in Z^n$, $\|v\| = 1$, such that $x + 2v \in X$, it is:*

$$\Delta^2 f(x; v) \geq 0 \tag{5}$$

Let us point out that any continuous convex function restricted over a discrete reticulum verifies the proposed definition.

Remark 1. Let $f : X \to \Re$, where $X \subset Z^n$ is a discrete reticulum, and let $x \in X$ and $v \in Z^n$, $\|v\| = 1$, be such that $x + 2v \in X$ and $x - 2v \in X$. By simply renaming the variables ($\bar{x} = x + 2v$), it follows that:

$$\Delta^2 f(x; v) \geq 0 \quad \Leftrightarrow \quad \Delta^2 f(x; -v) \geq 0$$

In other words, if inequality (5) holds for a certain direction v then it is necessarily verified also for the direction $-v$.

First of all, it is worth noticing that from Definition 6 it follows straightforward that the sum of two discrete convex functions is discrete convex too.

Theorem 1. *Let $f, g : X \to \Re$, where X is a discrete reticulum, be two discrete convex functions and let $\alpha \in \Re$, $\alpha > 0$. Then, $(f+g)(x)$ and $\alpha f(x)$ are discrete convex functions.*

Let us now prove the following characterization of discrete convex functions which points out that the proposed definition guarantees the increasing monotonicity of marginal increments.

Theorem 2. *Let $f : X \to \Re$, where $X \subset Z^n$ is a discrete reticulum. Function f is a discrete convex function if and only if for all $x \in X$, for all $k, h \in Z$, with $h \geq 0$ and $k \geq h$, for all $v \in Z^n$, $\|v\| = 1$, such that $x + (k+1)v \in X$, it is:*

$$\Delta f(x + kv; v) \geq \Delta f(x + hv; v) \tag{6}$$

Proof. The sufficiency follows just assuming $h = 0$ and $k = 1$ since (6) holds for any $k \geq h \geq 0$.

The necessity is proved noticing that the result is trivial when $k = h$, while in the case $k > h$ from the discreteness of the function it yields:

$$\begin{aligned}\Delta f(x + kv; v) - \Delta f(x + hv; v) &= \sum_{j=h}^{k-1} (\Delta f(x + (j+1)v; v) - \Delta f(x + jv; v)) \\ &= \sum_{j=h}^{k-1} \Delta^2 f(x + jv; v)\end{aligned}$$

so that the result follows directly from the discrete convexity of f.

The following further result will be useful in the next section in order to prove some global optimality conditions.

Theorem 3. *Let $f : X \to \Re$, where $X \subset Z^n$ is a discrete reticulum, be a discrete convex function. Then, for all $x \in X$, for all $k, h \in Z$, with $h \geq 0$ and $k \geq h$, for all $v \in Z^n$, $\|v\| = 1$, such that $x + kv \in X$, it is:*

$$f(x + kv) - f(x + hv) \geq (k - h)\Delta f(x + hv; v) \tag{7}$$

Proof. If $k = h$ the result is trivial; if $k > h$ notice that the discreteness of the function yields:

$$f(x+kv) - f(x+hv) = \sum_{j=h}^{k-1} (f(x+(j+1)v) - f(x+jv)) = \sum_{j=h}^{k-1} \Delta f(x+jv; v)$$

The result then follows by noticing that Theorem 2 implies for any $j \geq h$ that $\Delta f(x + jv; v) \geq \Delta f(x + hv; v)$.

3 Local and Global Optimality

In this section we aim to study the global optimality properties of discrete convex functions; in particular we are going to deepen on the behavior of local minima.

3.1 Definitions and Preliminary Results

For the sake of convenience, let us first introduce the following notations and definitions.

Definition 7. *Given a point $x \in Z^n$ the following sets are defined:*

$$H(x) = \{y \in Z^n : \; y = x + v, \; v \in Z^n, \; \|v\| = 1\}$$
$$S(x) = \{y \in Z^n : \; y = x + kv, \; k \in Z, \; v \in Z^n, \; \|v\| = 1\}$$

The set $H(x)$ represents the surface of a sort of discrete unitary hypercube around point x, so that it may be intended as a sort of neighborhood of x itself; $S(x)$ is a discrete star shaped set centered in x and generated by the discrete unitary directions. Obviously, it is $H(x) \subset S(x)$.

Definition 8. *Let $f : X \to \Re$, where $X \subset Z^n$ is a discrete reticulum. A point $x \in X$ is said to be a local minimum if:*

$$f(x) \leq f(y) \qquad \forall y \in X \cap H(x)$$

while it is said to be a global minimum if:

$$f(x) \leq f(y) \qquad \forall y \in X$$

The next preliminary result follows straightforward from Theorem 3.

Corollary 1. *Let $f : X \to \Re$, where $X \subset Z^n$ is a discrete reticulum, be a discrete convex function. If $x \in X$ is a local minimum then $f(x) \leq f(y)$ for all $y \in X \cap S(x)$.*

Proof. The result follows from Theorem 3 assuming $h = 0$ and noticing that the local optimality assumption implies that $\Delta f(x; v) = f(x + v) - f(x) \geq 0$.

3.2 Convexity and Optimality in Z

It is worth focusing on the attention to single variable discrete functions, due to their usefulness in applicative problems. First of all, it is worth noticing that for Remark 1 the first and second order differences can be simplified as follows:

$$\Delta f(x) = f(x+1) - f(x) \tag{8}$$

$$\Delta^2 f(x) = f(x+2) - 2f(x+1) + f(x) \tag{9}$$

Let us now show that single variable discrete convex functions can be characterized with properties which result to be easier to be verified with respect of the general definition.

Theorem 4. *Let $f : X \to \Re$, where $X \subset Z$ is a discrete reticulum. Function f is discrete convex if and only if for all $x \in X$ such that $x + 2 \in X$, it is:*

$$\Delta^2 f(x) \geq 0 \tag{10}$$

Proof. The result follows directly from Definition 6 and Remark 1.

Corollary 2. *Let $f : X \to \Re$, where $X \subset Z$ is a discrete reticulum. Function f is discrete convex if and only if for all $x, y \in X$ such that $y \geq x$ and $y + 1 \in X$, it is:*

$$\Delta f(y) \geq \Delta f(x)$$

Proof. The sufficiency follows trivially assuming $y = x + 1$. The necessity follows from Theorem 2 by assuming $v = 1$ and $y = x + k$.

Let us finally point out that for single variable functions the discrete convexity property guarantees the global optimality of local optima.

Corollary 3. *Let $f : X \to \Re$, where $X \subset Z$ is a discrete reticulum, be a discrete convex function. If $x \in X$ is a local minimum then it is also a global one.*

Proof. Follows directly from Corollary 1 since in the single variable case it is $S(x) = Z$.

As a conclusion, it is worth noticing that in the case of single variable functions the proposed definition of discrete convexity verifies all the typical properties of continuous convexity, such as the increasing monotonicity of the marginal increments, the global optimality of local optima, the discrete convexity of the sum of discrete convex functions.

3.3 Convexity and Optimality in Z^n, $n \geq 2$

Unlike the single variable case, when two or more discrete variables are involved then the discrete convexity of the function is not sufficient to guarantee the global optimality of a local optima. With this regard, it is worth noticing that Corollary 1 is not a complete global optimality result, since it states the global optimality of a local optimum only with respect to the set $X \cap S(x)$. This behavior is pointed out in the next example.

Example 1. Let us consider the following function defined over $X = Z^2$:

$$f(x_1, x_2) = (x_2 - 2x_1)^2 + \frac{1}{2}\left|x_2 + \frac{1}{2}x_1\right|$$

This is clearly a convex function over $\Re^2$ and hence it is also discrete convex over Z^2. Point $x = (0,0)$ is the unique global minimum, but by means of simple calculations it can be seen that, for example, the points $(1,2)$, $(2,4)$, $(3,6)$, are local optima (with respect of Definition 8) but not global ones.

As a consequence, some additional regularity assumptions are required to extend the optimality range of a local optimum. A first tentative regularity assumption is proposed in the next definition.

Definition 9. *Let $f : X \to \Re$, where $X \subset Z^n$ is a discrete reticulum. Let also be $W = \{w^{(1)}, \ldots, w^{(n)}\} \subset Z^n$ be a set of n linearly independent unitary vectors. The following regularity condition (R1) is then defined:*

- *for all $x \in X$, for all $i, j = 1, \ldots, n$, $i \neq j$, such that $x + w^{(i)} + w^{(j)} \in X$, it is $\Delta f(x + w^{(j)}; w^{(i)}) \geq \Delta f(x; w^{(i)})$*

In the case of discrete convex functions property (R1) can be characterized as follows.

Theorem 5. *Let $f : X \to \Re$, where $X \subset Z^n$ is a discrete reticulum, be a discrete convex function. Let also be $W = \{w^{(1)}, \ldots, w^{(n)}\} \subset Z^n$ be a set of n linearly independent unitary vectors. The regularity condition (R1) holds if and only if for all $x \in X$, for all $i = 1, \ldots, n$, for all $y \in Z^n \cap cone\{W\}$, such that $x + y + w^{(i)} \in X$, it is:*

$$\Delta f(x + y; w^{(i)}) \geq \Delta f(x; w^{(i)})$$

Proof. The sufficiency follows immediately by setting $y = w^{(j)}$. For the necessity, let us first prove the result for $y = kw^{(j)}$, $k \in \aleph$; in other words let us first prove that:

$$\Delta f(x + kw^{(j)}; w^{(i)}) \geq \Delta f(x; w^{(i)}) \tag{11}$$

In the case $j = i$ inequality (11) follows directly from Theorem 2 by assuming $v = w^{(i)}$ and $h = 0$. Consider now the case $j \neq i$; for $k = 0$ the result is trivial, while for $k = 1$ it follows directly from condition (R1). Let now be $k > 1$ and assume by induction that the inequality holds for $k-1$. By applying the regularity condition (R1) and the induction assumption it yields:

$$\begin{aligned}\Delta f(x + kw^{(j)}; w^{(i)}) &= \Delta f(x + (k-1)w^{(j)} + w^{(j)}; w^{(i)}) \\ &\geq \Delta f(x + (k-1)w^{(j)}; w^{(i)}) \\ &\geq \Delta f(x; w^{(i)})\end{aligned}$$

Let now y be any vector in $cone\{W\}$; then, it can be expressed in the form

$$y = \sum_{j=1}^{n} k^{(j)} w^{(j)}$$

where $k^{(j)} \in \aleph$, $j = 1, \ldots, n$, so that the thesis that we are going to prove can be rewritten as follows:

$$\begin{aligned} \Delta f(x + y; w^{(i)}) &= \Delta f(x + k^{(1)} w^{(1)} + \cdots + k^{(n)} w^{(n)}; w^{(i)}) \\ &\geq \Delta f(x; w^{(i)}) \end{aligned}$$

The result then follows directly by applying n times, one for every component $k^{(j)} w^{(j)}$ of y, the preliminary result (11).

The previous result allows us to improve the range of optimality of a local minimum.

Theorem 6. *Let $f : X \to \Re$, where $X \subset Z^n$ is a discrete reticulum, be a discrete convex function. Assume also that the regularity condition (R1) holds. If $x \in X$ is a local minimum, then x is a global minimum with respect to the sets $x + cone\{W\}$ and $x - cone\{W\}$.*

Proof. Assume by contradiction that x is not a global minimum with respect to $x + cone\{W\}$, that is to say that there exists $z \in X \cap (x + cone\{W\})$ such that $f(z) < f(x)$. It is now possible to construct a finite sequence of k elements $\{z^{(j)}\} \in (x + cone\{W\}) \cap (z - cone\{W\})$ such that $z^{(0)} = x$, $z^{(k)} = z$ and $z^{(j+1)} - z^{(j)} \in W$ for all $j = 0, \ldots, k-1$. Since $f(z) < f(x)$ there exists $\bar{k} \in [0, k-1]$ such that $f(z^{(\bar{k})}) > f(z^{(\bar{k}+1)})$. Let us define $y = z^{(\bar{k})} - x$ and let i be such that $w^{(i)} = z^{(\bar{k}+1)} - z^{(\bar{k})} \in W$; then we have $f(x+y) > f(x+y+w^{(i)})$, that is $\Delta f(x+y; w^{(i)}) < 0$. From Theorem 5 and from the local optimality of x it then yields:

$$0 \leq f(x + w^{(i)}) - f(x) = \Delta f(x; w^{(i)}) \leq \Delta f(x + y; w^{(i)}) < 0$$

which is a contradiction. Analogously, it can be proved that x is a global minimum with respect to $x - cone\{W\}$.

4 Convexity in an Optimal Fleet Mix Problem

Discrete optimization has many applications in everyday life and for this reason it has been widely studied in the literature.

This kind of problems are algorithmically difficult to be solved from a complexity point of view and are usually approached with integer programming techniques, branch and bound algorithms, local search, genetic algorithms.

In this section we aim to study a concrete optimal fleet mix problem, which is a discrete variables model related to the management of internal and external workforce units.

A theoretical study will point out that this problem can be solved with a polynomial complexity by means of a sort of parametrical approach. It will be also proved that this approach will provide a discrete convex parametrical objective function. This property allows to solve the problem very efficiently, that is with a very small CPU time, so that it could be used in a real time environment, such as in connection with real time routing problems.

4.1 Optimal Fleet Mix: An Integer Programming Problem

This concrete problem is referred to routing of maintenance units (see for example [1, 4]). The firm employes internal and external technicians for repairing ATMs. Customers signal technical malfunctions to the call center. After the signalling the company has a contractual time window to repair the machine. If the time elapses the firm has to pay a penalty. Main targets are: to minimize call rates, repair time, travel time, and penalty costs. Call rates depend on product reliability, repair times on service diagnostic and service tools, while the travel time is dependent on transportation methods and environmental conditions. The first three aspects concern internal politics of renovating machines and personal training. The last one is the one we treat in this work.

We introduce a suitable objective function that takes into account both fixed and variable costs. The aim is to minimize this objective function subject to quality of service (QoS) constraints. Let us study the problem with respect to a particular geographic area and within a period of one year and let us denote by I the set of days of the year. The variables represent the number of internal and external technicians to be employed. The input data are:

- the daily cost of the technicians
- the penalty costs
- the minimum service level the firm wants to guarantee.

First of all we examine the available historical series of calls for failures (without distinguishing among different types of failures) and we establish two benchmarks: the minimum and the maximum number of calls per day. From these parameters we can extrapolate the range of workforce necessary to reply to the failure calls. In particular, M_i and m_i are, respectively, the maximum number of calls that the firm's call center receives the day i according to the data of the historical series and the minimum number.

These two values determine the unique constraint of the model; in fact, the total number of calls that an employee is able to fulfill cannot be less than the minimum m_i for each i, that is $\beta_x x \geq m_i \ \forall i = 1, \ldots, I$. Actually, in order to guarantee a sort of quality of service, the firm may want to guarantee a higher minimum level of calls fulfilled; this can be represented by means of a parameter $\rho \in [0, 1]$. In order to define more in detail the model structure, let us introduce the following definition.

Definition 10. *Let us consider the following data and parameters:*

$M \in \aleph^I$: *estimated maximum number of calls*
$m \in \aleph^I$: *estimated minimum number of calls*
$I \in \aleph$: *number of working days under consideration*
$x \in \aleph$: *the number of employees of the firm*
$z_i \in \aleph, i = 1, \dots, I$: *number of external technicians employed at the i-th day*
$\beta_x \in \aleph$: *average number of calls fulfilled per single technician in a working day*
$p \in \Re_+$: *daily cost of the single internal technician*
$c_w \in \Re_+$: *penalty cost, proportional to the lack of technicians to repair the faults*
$c_z \in \Re_+$: *cost per call of the external technician.*
$\rho \in [0, 1]$: *penalty coefficient.*

The optimization problem can be modeled as follows:

$$P : \begin{cases} \min f(x, z) \\ (x, z) \in S \end{cases}$$

where the objective function represents the cost of the internal and/or external crews of technicians employed and is given by:

$$f(x, z) = Ixp + c_z \sum_{i=1}^{I} z_i + c_w w(x, z) \tag{12}$$

and the number of not fulfilled calls is:

$$\begin{aligned} w(x, z) &= \sum_{i=1}^{I} \max\{0; M_i - \beta_x x - z_i\} \\ &= \frac{1}{2} \sum_{i=1}^{I} (M_i - \beta_x x - z_i + |M_i - \beta_x x - z_i|) \end{aligned}$$

while the feasible region is given by the following daily constraints:

$$S = \left\{ x \in \aleph, z \in \aleph^I \,|\, M_i - \rho(M_i - m_i) \le \beta_x x + z_i \quad \forall i = 1, \dots, I \right\} \tag{13}$$

External technicians are employed not every day. In the days during which the call center receives many calls, the firm can decide to employ an unlimited number of external technicians and it pays them for the whole day. On the other hand, if internal employees can cover all the demand peaks, z_i will be equal to zero. This kind of mixed fleet is usually employed in firms with a high volatility of demand and a stochastic trend.

Remark 2. Since $(x, z) \in S$, i.e. $M_i - \beta_x x - z_i \leq \rho(M_i - m_i)$ $\forall i = 1, \ldots, I$, then $w(x, z) \leq \rho \sum_{i=1}^{I}(M_i - m_i)$. In this light, $\rho \sum_{i=1}^{I}(M_i - m_i)$ is the maximum number of calls which might be left unfulfilled. Note also that, since the objective function has to be minimized, we can restrict the study of the problem to the following interval of variable x:

$$0 \leq x \leq \bar{M} = \max_{i=1..I}\{\left\lceil \frac{M_i}{\beta_x} \right\rceil\}. \tag{14}$$

4.2 Fundamental Properties of the Problem

Problem P is a discrete variable minimum problem, and can be solved with any of the known discrete programming algorithms. Clearly, due to the great number of variables ($I + 1$ with I equal to the number of working days in the year), the complexity of such algorithms could make impossible the use of this problem in real time environments.

Actually, deepening the study of the problem, we can state properties which will allow to solve it with just a linear complexity and a very small CPU time requirement. First of all, let us notice that the objective function of problem P can be rewritten as

$$f(x, z) = Ipx + \sum_{i=1}^{I} \psi(x, z_i)$$

where for all $i = 1, \ldots, I$ it is:

$$\psi(x, z_i) = c_z z_i + c_w \max\{0; M_i - \beta_x x - z_i\}$$

In other words, the z_i variables are independent one each other, so that whenever x is considered as a parameter then problem P can be solved separately with respect to each variable z_i. This suggest us to state the following result.

Theorem 7. *Let us consider problem P and assume x to be a fixed parameter. For any $i \in \{1, ..., I\}$ the optimal solution of the following problem:*

$$\begin{cases} \min\limits_{z_i} \; g(z_i) = c_z z_i + c_w \max\{0; M_i - \beta_x x - z_i\} \\ \qquad\qquad z_i \geq M_i - \rho(M_i - m_i) - \beta_x x \end{cases}$$

is given by:

$$\hat{z}_i(x) = \begin{cases} \max\{0, M_i - \beta_x x\} & \text{if } c_z < c_w \\ \max\{0, M_i - \beta_x x - \lfloor \rho(M_i - m_i) \rfloor\} & \text{if } c_z \geq c_w \end{cases} \tag{15}$$

Proof. (Case $c_z \geq c_w$) We just need to prove that $g(z_i)$ is monotone increasing for $z_i \geq 0$, that is to say that $g(z_i + 1) - g(z_i) \geq 0$ for all $z_i \geq 0$. Noticing that

$$g(z_i+1)-g(z_i)=c_z-c_w\left(\max\{0;M_i-\beta_x x-z_i\}\right.$$
$$\left.-\max\{0;M_i-\beta_x x-z_i-1\}\right)$$

and taking into account that $M_i-\beta_x x-z_i$ is an integer value, it results:

$$g(z_i+1)-g(z_i)=\begin{cases} c_z & \text{if } M_i-\beta_x x-z_i\leq 0\\ c_z-c_w & \text{if } M_i-\beta_x x-z_i\geq 1\end{cases}$$

and the result is proved since $c_z\geq 0$ and $c_z\geq c_w$, taking into account that $\lceil M_i-\rho(M_i-m_i)-\beta_x x\rceil = M_i-\beta_x x-\lfloor\rho(M_i-m_i)\rfloor$ since β_x and x are nonnegative integers.

(Case $c_z<c_w$) Being z_i a nonnegative integer it yields:

$$\max\{0;M_i-\beta_x x-z_i\}=\begin{cases} 0 & \text{if } z_i\geq\max\{0,M_i-\beta_x x\}\\ M_i-\beta_x x-z_i & \text{if } 0\leq z_i<\max\{0,M_i-\beta_x x\}\end{cases}$$

and hence:

$$g(z_i)=\begin{cases} c_z z_i & \text{if } z_i\geq\max\{0,M_i-\beta_x x\}\\ z_i(c_z-c_w)+c_w(M_i-\beta_x x) & \text{if } 0\leq z_i<\max\{0,M_i-\beta_x x\}\end{cases}$$

Taking into account that $c_z>0$ and $c_z-c_w<0$, the minimum is reached in the feasible value $\hat{z}_i(x)=\max\{0,M_i-\beta_x x\}$.

Remark 3. It is worth pointing out an economic interpretation of the previously obtained results.

In the case $c_z\geq c_w$ the cost of an additional external technician is greater than the cost of the penalty. This means that, from the firm's point of view, it is better to pay the penalty than to employ an additional external technician; as a consequence the optimal value of z_i corresponds to the lower value it can assume, that is $\max\{0,M_i-\beta_x x-\lfloor\rho(M_i-m_i)\rfloor\}$.

On the other hand, in the case $c_z<c_w$ it is better for the firm to avoid penalties fulfilling all of the daily calls. In this light the firm employs all the necessary external technicians, given by $\max\{0,M_i-\beta_x x\}$.

Theorem 7 and Remark 2 allow to rewrite problem P as follows:

$$P:\begin{cases}\min\ \varphi(x)=f(x,\hat{z}(x))\\ \quad 0\leq x\leq\bar{M}\end{cases}\tag{16}$$

where $\hat{z}(x)=(\hat{z}_1(x),\ldots,\hat{z}_I(x))$ as given in (15). Just notice also that

$$\varphi(x)=Ixp+c_z\sum_{i=1}^{I}\hat{z}_i(x)+c_w w(x,\hat{z}(x))\tag{17}$$

$$=Ixp+\sum_{i=1}^{I}\psi(x,\hat{z}_i(x))\tag{18}$$

and that in the case $c_z < c_w$ it is $w(x, \hat{z}(x)) = 0$ for all $x \in [0, \bar{M}]$, while in the case $c_z \geq c_w$ it is

$$
\begin{aligned}
w(x, \hat{z}(x)) &= \sum_{i=1}^{I} \max\{0; M_i - \beta_x x - \max\{0, M_i - \beta_x x - \lfloor \rho(M_i - m_i) \rfloor\}\} \\
&= \sum_{i=1}^{I} \max\{0; \min\{M_i - \beta_x x; \lfloor \rho(M_i - m_i) \rfloor\}\}
\end{aligned}
$$

As a conclusion, problem P has become a problem of a single variable and can be solved by simply comparing the values of $\varphi(x)$ for all $x \in [0, \bar{M}]$.

4.3 Discrete Convexity of the Objective Function $\varphi(\mathbf{x})$

In the previous subsection we have shown that problem P can be easily solved, from a mathematical point of view, with a single variable discrete problem.

In order to improve the use of this problem as part of a real time system, it is important to determine the optimal solution with a CPU time as small as possible.

In this light, we now aim to study the discrete convexity of function $\varphi(x)$, in order to use the global optimality of local minima (see Corollary 3) as an efficient stopping criterion.

Theorem 8. *Consider problem P and function $\varphi(x)$ as defined in (16) and (17). Then, function $\varphi(x)$ is discrete convex.*

Proof. By means of Theorem 4 function $\varphi(x)$ is discrete convex if and only if $\Delta^2\varphi(x) \geq 0$ for all $x \in [0, \bar{M}]$. Two exhaustive cases are now going to be considered.

(Case $c_z < c_w$) Since $w(x, \hat{z}(x)) = 0$ for all $x \in [0, \bar{M}]$ it results

$$
\begin{aligned}
\Delta^2\varphi(x) &= c_z \sum_{i=1}^{I} [\hat{z}_i(x+2) + \hat{z}_i(x) - 2\hat{z}_i(x+1)] \\
&= c_z \sum_{i=1}^{I} \Delta^2 \hat{z}_i(x)
\end{aligned}
$$

By means of simple calculation, for all $i = 1, \dots, I$ we get:

$$
\Delta^2 \hat{z}_i(x) = \begin{cases} 0 & \text{if } M_i - \beta_x x \geq 2\beta_x \\ 2\beta_x - M_i + \beta_x x & \text{if } \beta_x \leq M_i - \beta_x x < 2\beta_x \\ M_i - \beta_x x & \text{if } 0 \leq M_i - \beta_x x < \beta_x \\ 0 & \text{if } M_i - \beta_x x < 0 \end{cases}
$$

so that $\Delta^2 \hat{z}_i(x) \geq 0$ for all $i = 1, \dots, I$ which implies $\Delta^2\varphi(x) \geq 0$ too.

(Case $c_z \geq c_w$) For the sake of convenience, let us introduce the following notation:

$$\hat{h}_i(x) = c_z \hat{z}_i(x) + c_w \max\{0; \min\{M_i - \beta_x x; \lfloor \rho(M_i - m_i) \rfloor\}\}$$

so that $\varphi(x) = Ixp + \sum_{i=1}^{I} \hat{h}_i(x)$ and hence $\Delta^2\varphi(x) = \sum_{i=1}^{I} \Delta^2 \hat{h}_i(x)$. Some exhaustive subcases have now to be considered for any $i = 1, \ldots, I$.

Assume $M_i - \beta_x x - \lfloor \rho(M_i - m_i) \rfloor \geq 2\beta_x$. Then, it results $\Delta^2 \hat{h}_i(x) = 0$.

Assume $\beta_x \leq M_i - \beta_x x - \lfloor \rho(M_i - m_i) \rfloor < 2\beta_x$. Then, by means of simple calculations and taking into account that $c_z \geq c_w$, we have:

$$\begin{aligned}
\Delta^2 \hat{h}_i(x) &= c_z \left[-M_i + \beta_x x + 2\beta_x + \lfloor \rho(M_i - m_i) \rfloor\right] + \\
&\quad + c_w \left[\max\{0; M_i - \beta_x x - 2\beta_x\} - \lfloor \rho(M_i - m_i) \rfloor\right] \\
&\geq c_w \left[-M_i + \beta_x x + 2\beta_x + \max\{0; M_i - \beta_x x - 2\beta_x\}\right] \\
&= c_w \max\{0; -M_i + \beta_x x + 2\beta_x\} \geq 0
\end{aligned}$$

Assume $0 \leq M_i - \beta_x x - \lfloor \rho(M_i - m_i) \rfloor < \beta_x$. Then, by means of simple calculations and taking into account that $c_z \geq c_w$, we have:

$$\begin{aligned}
\Delta^2 \hat{h}_i(x) &= c_z \left[M_i - \beta_x x - \lfloor \rho(M_i - m_i) \rfloor\right] + c_w \lfloor \rho(M_i - m_i) \rfloor + \\
&\quad + c_w \left[\max\{0; M_i - \beta_x x - 2\beta_x\} - 2\max\{0; M_i - \beta_x x - \beta_x\}\right] \\
&\geq c_w \left[M_i - \beta_x x + \max\{0; M_i - \beta_x x - 2\beta_x\}\right. \\
&\quad \left. - 2\max\{0; M_i - \beta_x x - \beta_x\}\right]
\end{aligned}$$

By means of the exhaustive cases $M_i - \beta_x x \geq 2\beta_x$, $\beta_x \leq M_i - \beta_x x < 2\beta_x$ and $M_i - \beta_x x < \beta_x$, and recalling that $M_i - \beta_x x \geq \lfloor \rho(M_i - m_i) \rfloor \geq 0$, it can then be easily verified that $\Delta^2 \hat{h}_i(x) \geq 0$.

Assume $M_i - \beta_x x - \lfloor \rho(M_i - m_i) \rfloor < 0$. Then, it results

$$\begin{aligned}
\Delta^2 \hat{h}_i(x) &= c_w \max\{0; M_i - \beta_x x - 2\beta_x\} + c_w \max\{0; M_i - \beta_x x\} + \\
&\quad - 2c_w \max\{0; M_i - \beta_x x - \beta_x\}
\end{aligned}$$

so that

$$\Delta^2 \hat{h}_i(x) = \begin{cases} 0 & \text{if } M_i - \beta_x x \geq 2\beta_x \\ c_w(2\beta_x - M_i + \beta_x x) & \text{if } \beta_x \leq M_i - \beta_x x < 2\beta_x \\ c_w(M_i - \beta_x x) & \text{if } 0 \leq M_i - \beta_x x < \beta_x \\ 0 & \text{if } M_i - \beta_x x < 0 \end{cases}$$

which implies the nonnegativity of $\Delta^2 \hat{h}_i(x)$.

As a conclusion, we have stated that $\Delta^2 \hat{h}_i(x) \geq 0$ for all $i = 1, \ldots, I$, and this implies $\Delta^2 \varphi(x) \geq 0$ too. The result is then proved.

Finally, let us conclude our study pointing out how the optimal solution can be found by using the discrete convexity of the function as an efficient stopping criterion.

Algorithm Structure

1) Determine $\bar{M}$ and let $x^* := 0$, $x' := 0$ and $local := false$;
2) While not $local$ and $x' < \bar{M}$ do
   2a) $x' := x' + 1$;
   2b) if $\varphi(x') < \varphi(x^*)$ then $x^* := x'$ else $local := true$
3) The optimal solution of problem P is $(x^*, \hat{z}(x^*))$ with optimal value $\varphi(x^*)$.

References

1. Gheysens F, Golden B, Assad A (1984) The fleet size and mix vehicle routing problem. Computers & Operations Research, 11:49-66
2. Favati P, Tardella F (1990) Convexity in nonlinear integer programming. Ricerca Operativa, 53:3-44
3. Murota K (1998) Discrete convex analysis. Mathematical Programming, 83:313-371
4. Wyatt JK (1961) Optimal Fleet Size. Operational Research Quarterly, 12:187-188
5. Yüceer Ü (2002) Discrete convexity: convexity for functions defined on discrete spaces. Discrete Applied Mathematics, 119:297-304

A Unifying Approach to Solve a Class of Parametrically-Convexifiable Problems

Riccardo Cambini and Claudio Sodini

Dept. of Statistics and Applied Mathematics, University of Pisa
Faculty of Economics, Via Cosimo Ridolfi 10, 56124 Pisa, ITALY
cambric@ec.unipi.it - csodini@ec.unipi.it

Summary. The aim of this paper is to show how a wide class of generalized quadratic programs can be solved, in a unifying framework, by means of the so-called optimal level solutions method. In other words, the problems are solved by analyzing, explicitly or implicitly, the optimal solutions of particular quadratic strictly convex parametric subproblems. In particular, it is pointed out that some of these problems share the same set of optimal level solutions. A solution algorithm is proposed and fully described. The results achieved are then deepened in the particular case of box constrained problems.

Key words: Generalized quadratic programming, fractional programming, optimal level solutions.

1 Introduction

The aim of this paper is to study and to propose a solution method for the following class of generalized quadratic problems:

$$P : \begin{cases} \inf \phi(x) = f\left(\frac{1}{2}x^TQx + q^Tx + q_0\right) g_1\left(d^Tx + d_0\right) + g_2\left(d^Tx + d_0\right) \\ x \in X = \{x \in \Re^n : Ax \geq b\} \end{cases}$$

where $A \in \Re^{m \times n}$, $q, d \in \Re^n$, $d \neq 0$, $b \in \Re^m$, $q_0, d_0 \in \Re$, $Q \in \Re^{n \times n}$ is symmetric and positive definite, $g_1, g_2 : \Omega_g \to \Re$, $f : \Omega_f \to \Re$, with g_1 positive over Ω_g and f strictly increasing over Ω_f, where

$$\Omega_g = \left\{y \in \Re : y = d^Tx + d_0, x \in X\right\} ,$$
$$\Omega_f = \left\{y \in \Re : y = \frac{1}{2}x^TQx + q^Tx + q_0, x \in X\right\} .$$

Various particular problems belonging to this class have been studied in the literature of mathematical programming and global optimization, from both a theoretic and an applicative point of view ([2, 12, 13, 14, 20]). In particular, it is worth noticing that this class covers several multiplicative, fractional, d.c. and generalized quadratic problems (see for all [4, 6, 7, 8, 11, 16, 18]) which are very used in applications, such as location models, tax programming models, portfolio theory, risk theory, Data Envelopment Analysis (see for all [1, 9, 11, 15, 16, 21]).

The solution method proposed to solve this class of problems is based on the so called "optimal level solutions" method (see [3, 4, 5, 6, 7, 8, 10, 17, 18, 19]). It is known that this is a parametric method, which finds the optimum of the problem by determining the minima of particular subproblems. In particular, the optimal solutions of these subproblems are obtained by means of a sensitivity analysis aimed to maintain the Karush-Kuhn-Tucker optimality conditions.

Applying the optimal level solutions method to problem P we obtain some strictly convex quadratic subproblems which are independent of functions f, g_1 and g_2. In other words, different problems share the same set of optimal level solutions, and this allow us to propose an unifying method to solve all of them.

In Section 2 we describe how the optimal level solutions method can be applied to problem P; in Section 3 a solution algorithm is proposed and fully described; finally, in Section 4, the obtained results are deepened on for the particular case of box constrained problems.

2 Optimal Level Solutions Approach

In this section we show how problem P can be solved by means of the optimal level solutions approach [3, 5, 6, 7, 10, 17]. With this aim, let $\xi \in \Re$ be a real parameter. The following parametric subproblem can be obtained just by adding to problem P the constraint $d^T x + d_0 = \xi$:

$$P_\xi : \begin{cases} \inf f\left(\frac{1}{2}x^T Qx + q^T x + q_0\right) g_1(\xi) + g_2(\xi) \\ x \in X_\xi = \{x \in \Re^n : Ax \geq b, d^T x + d_0 = \xi\} \end{cases}$$

The parameter ξ is said to be a *feasible level* if the set X_ξ is nonempty, that is if $\xi \in \Omega_g$. An optimal solution of problem P_ξ is called an *optimal level solution*. Since g_1 is positive over Ω_g and f is strictly increasing over Ω_f, then for any given $\xi \in \Omega_g$ the optimal solution of problem P_ξ coincides with the optimal solution of the following strictly convex quadratic problem $\overline{P}_\xi$:

$$\overline{P}_\xi : \begin{cases} \inf \frac{1}{2}x^T Qx + q^T x + q_0 \\ x \in X_\xi = \{x \in \Re^n : Ax \geq b, d^T x + d_0 = \xi\} \end{cases}$$

In this light, we say that function ϕ is *parametrically-convexifiable*.

For the sake of completeness, let us now briefly recall the optimal level solutions approach (see for example [10]). Obviously, the optimal solution of problem P is also an optimal level solution and, in particular, it is the optimal level solution with the smallest value; the idea of this approach is then to scan all the feasible levels, studying the corresponding optimal level solutions, until the minimizer of the problem is reached or a feasible halfline carrying $\phi(x)$ down to its infimum value is found.

Starting from an incumbent optimal level solution, this can be done by means of a sensitivity analysis on the parameter ξ, which allows us to move in the various steps through several optimal level solutions until the optimal solution is found.

Remark 1. Let us point out that problems $\overline{P}_\xi$ are independent of the functions f, g_1 and g_2. This means that different parametrically-convexifiable problems, either multiplicative or fractional or d.c. quadratic ones, share the same set of optimal level solutions and can then be solved by means of the same algorithm iterations. In this light, it can be said that the solution method we propose in this paper represents an unifying framework for various classes of generalized quadratic problems.

2.1 Starting Problem and Sensitivity Analysis

Let x' be the optimal solution of problem $\overline{P}_{\xi'}$, where $d^T x' + d_0 = \xi'$, and let us consider the following Karush-Kuhn-Tucker conditions for $\overline{P}_{\xi'}$:

$$\begin{cases} Qx' + q = A^T\mu + d\lambda & \\ d^T x' + d_0 = \xi' & \\ Ax' \geqq b & \textit{feasibility} \\ \mu \geqq 0 & \textit{optimality} \\ \mu^T(Ax' - b) = 0 & \textit{complementarity} \\ \lambda \in \Re,\ \mu \in \Re^m & \end{cases} \tag{1}$$

Since $\overline{P}_{\xi'}$ is a strictly convex problem, the previous system has at least one solution (μ', λ').

By means of a sort of sensitivity analysis, we now aim to study the optimal level solutions of problems $\overline{P}_{\xi'+\theta}$, $\theta \in (0, \epsilon)$ with $\epsilon > 0$ small enough. This can be done by maintaining the consistence of the Karush-Kuhn-Tucker systems corresponding to these problems.

Since the Karush-Kuhn-Tucker systems are linear whenever the complementarity conditions are implicitly handled, then the solution of the optimality conditions regarding to $\overline{P}_{\xi'+\theta}$ is of the kind:

$$x'(\theta) = x' + \theta \Delta_x \ , \ \lambda'(\theta) = \lambda' + \theta \Delta_\lambda \ , \ \mu'(\theta) = \mu' + \theta \Delta_\mu \tag{2}$$

so that it results:

$$\begin{cases} Q(x' + \theta\Delta_x) + q = A^T(\mu' + \theta\Delta_\mu) + d(\lambda' + \theta\Delta_\lambda) \\ d^T(x' + \theta\Delta_x) + d_0 = \xi' + \theta, \\ A(x' + \theta\Delta_x)\geqq b \\ (\mu' + \theta\Delta_\mu)\geqq 0 \\ (\mu'_i + \theta\Delta_{\mu_i})\,(a_i(x' + \theta\Delta_x) - b_i) = 0 \ \forall i = 1, \dots, m \\ \Delta_\lambda \in \Re, \ \Delta_\mu \in \Re^m, \ \Delta_x \in \Re^n \end{cases} \tag{3}$$

where a_i, $i = 1, \dots, m$, is the i-th row of A.

It is worth pointing out that the strict convexity of problem $\overline{P}_{\xi'+\theta}$ guarantees for any $\theta \in (0, \epsilon)$ the uniqueness of the optimal level solution $x'(\theta) = x' + \theta\Delta_x$; this implies also the following important property:

vector Δ_x is unique and different from 0.

Let us now provide an useful preliminary lemma which suggests how to study system (3). With this aim, let us define, the following sets of indices based on the binding and the nonbinding constraints:

$$B = \{i : \ a_i x' = b_i \ , \ i = 1, \dots, m\} \ , \ N = \{i : \ a_i x' > b_i \ , \ i = 1, \dots, m\} \ .$$

Lemma 1. *Let (μ', λ') be a solution of (1). Then, for $\theta \in (0, \epsilon)$ system (3) is equivalent to:*

$$\begin{cases} Q\Delta_x = A^T\Delta_\mu + d\Delta_\lambda \\ d^T\Delta_x = 1, \\ Ax' + \theta A\Delta_x\geqq b \\ \mu' + \theta\Delta_\mu\geqq 0 \\ \mu'_i = \Delta_{\mu_i} = 0 \ \forall i \in N \\ \mu'_i a_i\Delta_x = 0 \ , \ \Delta_{\mu_i} a_i \Delta_x = 0 \quad \forall i \in B \end{cases} \tag{4}$$

Proof. The first and the second equations follow directly from (1) taking into account that $\theta \neq 0$. From (1) we have also that the complementarity conditions of (3) can be rewritten as:

$$\mu'_i a_i \Delta_x + \Delta_{\mu_i}(a_i x' - b_i) + \theta\Delta_{\mu_i} a_i \Delta_x = 0 \ \forall i = 1, \dots, m \ . \tag{5}$$

For any index $i \in N$ and for $\theta > 0$ small enough it is $(a_i(x' + \theta\Delta_x) - b_i) \neq 0$, so that from (3) it results $\mu'_i + \theta\Delta_{\mu_i} = 0$. This last equation holds for any $\theta > 0$ small enough if and only if $\mu'_i = \Delta_{\mu_i} = 0$; in other words it is:

$$\mu'_i = \Delta_{\mu_i} = 0 \quad \forall i \in N$$

which also yields:

$$\Delta_{\mu_i}(a_i x' - b_i) = 0 \quad \forall i = 1, \dots, m \ .$$

This equality implies that condition (5) holds for any $\theta \in (0, \epsilon)$ if and only if for all $i = 1, \dots, m$ it is:

$$\mu'_i a_i \Delta_x = 0 \ , \ \Delta_{\mu_i} a_i \Delta_x = 0$$

and the result is proved.

Note that from the positivity of θ, the feasibility conditions and the optimality ones, we also have:

$$a_i \Delta_x \geq 0 \quad \forall i \in B \ ,$$
$$\Delta_{\mu_i} \geq 0 \quad \forall i \in B \ \text{ such that } \ \mu'_i = 0 \ .$$

As a conclusion, we can compute the values of the multipliers λ', μ', Δ_λ, Δ_μ, Δ_x by solving the following overall system (which has $2 + 2m + n$ variables):

$$\begin{cases} Qx' + q = A^T \mu' + d\lambda' \\ Q\Delta_x = A^T \Delta_\mu + d\Delta_\lambda \\ d^T \Delta_x = 1, \\ \mu'_i = \Delta_{\mu_i} = 0 \quad \forall i \in N \\ \mu'_i \geq 0 \quad \forall i \in B \\ a_i \Delta_x \geq 0 \ \forall i \in B \\ \mu'_i a_i \Delta_x = 0 \ \forall i \in B \\ \Delta_{\mu_i} a_i \Delta_x = 0 \ \forall i \in B \\ \Delta_{\mu_i} \geq 0 \quad \forall i \in B \ \text{ s.t. } \ \mu'_i = 0 \\ \lambda', \Delta_\lambda \in \Re, \ \mu', \Delta_\mu \in \Re^m, \ \Delta_x \in \Re^n \end{cases} \tag{6}$$

This system is suitable for values of $\theta \geq 0$ verifying the following conditions:

$$\textit{feasibility conditions} : Ax' + \theta A \Delta_x \geqq b \ ,$$
$$\textit{optimality conditions} : \mu' + \theta \Delta_\mu \geqq 0 \ .$$

Notice that system (6) is consistent if and only if the feasible regions $X_{\xi'+\theta}$ of problems $\overline{P}_{\xi'+\theta}$ are nonempty for $\theta > 0$ small enough.

In the case system (6) is consistent, we are finally able to determine the values of $\theta > 0$ which guarantee both the optimality and the feasibility of $x'(\theta)$. Let $N^- = \{i \in N : \ a_i \Delta_x < 0\}$ ([1]); since $Ax' \geqq b$, from the feasibility conditions we have:

$$\theta \leq \hat{F} = \begin{cases} \min_{i \in N^-} \left\{ \frac{b_i - a_i x'}{a_i \Delta_x} \right\} & \text{if } N^- \neq \emptyset \\ +\infty & \text{if } N^- = \emptyset \end{cases}$$

where $\hat{F} > 0$. On the other hand, let $B^- = \{i \in B : \ \Delta_{\mu_i} < 0\}$ (recall that $\Delta_{\mu_i} = 0 \ \forall i \in N$); from the optimality conditions we have:

$$\theta \leq \hat{O} = \begin{cases} \min_{i \in B^-} \left\{ \frac{-\mu_i}{\Delta_{\mu'_i}} \right\} & \text{if } B^- \neq \emptyset \\ +\infty & \text{if } B^- = \emptyset \end{cases}$$

where $\hat{O} > 0$ (since $\theta > 0$ then inequalities $\Delta_{\mu_i} < 0$ and $\mu'_i + \theta \Delta_{\mu_i} \geq 0$ imply $\mu'_i > 0$). Hence, $x'(\theta)$ is an optimal level solution for all θ such that:

$$0 \leq \theta \leq \theta_m = \min \left\{ \hat{F}, \hat{O} \right\}$$

where $\theta_m > 0$ (obviously, when system (6) is consistent).

[1] Since $\theta > 0$ then inequalities $a_i \Delta_x < 0$ and $a_i x' + \theta a_i \Delta_x \geqq b_i$ imply $b_i - a_i x' < 0$, that is to say that $i \in N$.

2.2 Solving the Multipliers System

The aim of this subsection is to show how system (6) can be improved in order to determine its solutions. For the sake of convenience, from now on the rows of A and the components of b and μ' will be partitioned accordingly to the set of indices B and N.

Multiplying the first and the second equations of (6) by $\Delta_x \neq 0$ and taking into account that Q is positive definite, it follows:

$$\lambda' = (Qx' + q)^T \Delta_x \quad \text{and} \quad \Delta_\lambda = \Delta_x^T Q \Delta_x > 0 \ . \tag{7}$$

Multiplying the first equation of (6) by $d \neq 0$ and after simple calculations we also get:

$$\lambda' = \frac{1}{d^T d} d^T \left(Qx' + q - A^T \mu' \right) \ .$$

Let us now define the matrix $\hat{D} = \left(I - \frac{1}{d^T d} dd^T \right)$; note that $\hat{D}$ is symmetric, singular (since $\hat{D}d = 0$) and positive semidefinite (the $n-1$ nonzero eigenvalues are all equal to 1 since $\hat{D}y = y \ \forall y \in d^\perp$). Noticing that $d\lambda' = \left(I - \hat{D} \right) \left(Qx' + q - A^T \mu' \right)$ and that $\mu'_N = 0$, we can rewrite the first equation of (6) as follows:

$$\hat{D} A_B^T \mu'_B = \hat{D} \left(Qx' + q \right) \ .$$

The solution of this system is not unique in general; in particular note that:

$$\operatorname{rank}(\hat{D} A_B^T) \leq \min\{n-1, \operatorname{rank}(A_B)\} \ .$$

For the sake of convenience, let us now define the scalar $\delta = \frac{1}{d^T Q^{-1} d} > 0$ and the symmetric matrix $\hat{Q}_d = (Q^{-1} - \delta Q^{-1} dd^T Q^{-1})$ which results to be singular (since $\hat{Q}_d d = 0$) and positive semidefinite (for Theorem 2.1 in [8] ([2])). Since Q is nonsingular then, from the second and the third equations of (6), we get:

$$\Delta_\lambda = \frac{1 - d^T Q^{-1} A^T \Delta_\mu}{d^T Q^{-1} d} = \delta - \delta d^T Q^{-1} A^T \Delta_\mu \ ,$$
$$\Delta_x = Q^{-1} A^T \Delta_\mu + Q^{-1} d \Delta_\lambda = \delta Q^{-1} d + \hat{Q}_d A^T \Delta_\mu \ .$$

As a conclusion, we have the following explicit solutions of system (6), some of them depending on Δ_{μ_B}:

[2] **Theorem 2.1 [8]** Let $Q \in \Re^{n \times n}$ be a symmetric positive definite matrix, let $k \in \Re$ and let $h \in \Re^n$. Then, the symmetric matrix $A = \left(Q + khh^T \right)$ is positive semidefinite if and only if $k \geq -\frac{1}{h^T Q^{-1} h}$.

$$\begin{aligned}\mu'_N &= 0\\ \Delta_{\mu_N} &= 0\\ \Delta_x &= \delta Q^{-1} d + \hat{Q}_d A_B^T \Delta_{\mu_B}\\ \lambda' &= (Qx' + q)^T \Delta_x\\ \Delta_\lambda &= \Delta_x^T Q \Delta_x\end{aligned}$$

Note that the uniqueness of vector Δ_x implies the uniqueness of λ' and Δ_λ.

We are now left to compute the values of vectors μ_B and Δ_{μ_B}. With this aim, for the sake of convenience, let $v_B = A_B Q^{-1} d$ and $R_B = A_B \hat{Q}_d A_B^T = (A_B Q^{-1} A_B^T - \delta v_B v_B^T)$. Matrix R_B is symmetric and positive semidefinite (due to the semipositiveness of $\hat{Q}_d$) with:

$$\text{rank}(R_B) \leq \min\{n-1, \text{rank}(A_B)\}$$

notice also that the i-th component of v_B is $v_i = a_i Q^{-1} d$ while the i-th row of R_B is given by $r_i = (a_i Q^{-1} A_B^T - \delta v_i v_B^T)$, so that $a_i \Delta_x = r_i \Delta_{\mu_B} + \delta v_i$. Vectors μ_B and Δ_{μ_B} are then solutions of the following system:

$$\begin{cases} \hat{D} A_B^T \mu_B = \hat{D}\,(Qx' + q)\\ \mu_B \geqq 0\\ R_B \Delta_{\mu_B} + \delta v_B \geqq 0\\ \mu_i (r_i \Delta_{\mu_B} + \delta v_i) = 0 \;\forall i \in B\\ \Delta_{\mu_i} (r_i \Delta_{\mu_B} + \delta v_i) = 0 \;\forall i \in B\\ \Delta_{\mu_i} \geq 0 \;\forall i \in B \text{ s.t. } \mu_i = 0 \end{cases} \tag{8}$$

Notice that the number of variables in system (8) is just $2card(B)$, where $card(B)$ is the number of elements in the set B.

2.3 Optimal Level Solutions Comparison

The optimal level solutions $x'(\theta)$ obtained by means of the sensitivity analysis can be compared just by evaluating the function $z(\theta) = \phi\,(x'(\theta))$. Defining $z' = \frac{1}{2} x'^T Q x' + q^T x' + q_0$ and recalling equations (7) it then results:

$$\frac{1}{2} x'(\theta)^T Q x'(\theta) + q^T x'(\theta) + q_0 = \frac{1}{2} \Delta_\lambda \theta^2 + \lambda' \theta + z' \ .$$

Hence, since $d^T x'(\theta) + d_0 = \xi' + \theta$, we get:

$$\begin{aligned} z(\theta) = \phi\,(x'(\theta)) &= f\left(\frac{1}{2}\Delta_\lambda \theta^2 + \lambda'\theta + z'\right) g_1\,(\xi' + \theta) + g_2\,(\xi' + \theta)\\ \frac{dz}{d\theta}(\theta) &= \frac{df}{d\theta}\left(\frac{1}{2}\Delta_\lambda \theta^2 + \lambda'\theta + z'\right)(\Delta_\lambda \theta + \lambda')\, g_1\,(\xi' + \theta) +\\ &\quad + f\left(\frac{1}{2}\Delta_\lambda \theta^2 + \lambda'\theta + z'\right) \frac{dg_1}{d\theta}\,(\xi' + \theta) + \frac{dg_2}{d\theta}\,(\xi' + \theta) \end{aligned}$$

so that, in particular:

$$\frac{dz}{d\theta}(0) = \lambda' \frac{df}{d\theta}(z') \, g_1(\xi') + f(z') \frac{dg_1}{d\theta}(\xi') + \frac{dg_2}{d\theta}(\xi') \ .$$

As it is very well known, the derivative $\frac{dz}{d\theta}(0)$ can be useful since its sign implies the local decreasing or increasing behavior of $z(\theta)$.

Level optimality is helpful also in studying local optimality, since a local minimum point in a segment of optimal level solutions is a local minimizer of the problem. This fundamental property allows to prove the following global optimality conditions in the case of a convex objective function $\phi(x)$.

Theorem 1. *Consider problem P, assume $\phi(x)$ convex and let $x'(\theta)$ be the optimal solution of problem $\overline{P}_{\xi'+\theta}$.*

i) If $\frac{dz}{d\theta}(0) > 0$ then $\phi(x') \le \phi(x)$ for all $x \in B$ such that $d^T x \ge d^T x'$.
ii) If $\theta_m < +\infty$ and $\overline{\theta} = \arg\min_{\theta\in[0,\theta_m]} \{z(\theta)\}$ is such that $0 < \overline{\theta} < \theta_m$, then $x'(\overline{\theta})$ is the optimal solution of problem P.

Proof. Since $\phi(x)$ is convex any local optimum is also global. The results then follow since a local minimum point in a segment of optimal level solutions is also a local minimizer.

3 A Solution Algorithm

In order to find a global minimum (or just the infimum) it would be necessary to solve problems $\overline{P}_\xi$ for all the feasible levels. In this section we will show that, by means of the results stated so far, this can be done algorithmically in a finite number of iterations.

The solution algorithm starts from a certain minimal level and then scans all the greater ones looking for the optimal solution, as it is pointed out in the next initialization process.

Initialization Steps
Compute, by means of two linear programs, the values ([3]):

$$\xi_{min} := \inf_{x\in X} d^T x + d_0 \quad , \quad \xi_{max} := \sup_{x\in X} d^T x + d_0 \ .$$

One of the following cases occurs.

1) If $\xi_{min} > -\infty$ then solve problem P from the starting feasible level $\xi_{start} = \xi_{min}$ up to the level $\xi_{end} = \xi_{max}$.

[3] Obviously, it may be $\xi_{min} = -\infty$ and/or $\xi_{max} = +\infty$.

2) If $\xi_{min} = -\infty$ and $\xi_{max} < +\infty$ then let $\tilde{g}_1(\xi) = g_1(-\xi)$ and $\tilde{g}_2(\xi) = g_2(-\xi)$, so that the objective function of P can be rewritten as:

$$\phi(x) = f\left(\frac{1}{2}x^T Qx + q^T x + q_0\right)\tilde{g}_1\left(-d^T x - d_0\right) + \tilde{g}_2\left(-d^T x - d_0\right) \ .$$

We can then solve problem P using $\tilde{g}_1$ and $\tilde{g}_2$ and scanning the feasible levels from the starting value $\xi_{start} = -\xi_{max} > -\infty$ up to $\xi_{end} = +\infty$.

3) If $\xi_{min} = -\infty$ and $\xi_{max} = +\infty$ then solve sequentially the next two problems from the starting level $\xi_{start} = 0$ up to the level $\xi_{end} = +\infty$:

$$P_+ : \begin{cases} \inf f(x) \\ d^T x + d_0 \geq 0 \\ x \in X \end{cases} \quad \text{and} \quad P_- : \begin{cases} \inf f(x) \\ d^T x + d_0 \leq 0 \\ x \in X \end{cases}$$

where P_- is defined using $\tilde{g}_1$ and $\tilde{g}_2$.

□

Once the starting feasible level ξ_{start} is found, the optimal solution can be searched iteratively by means of the following algorithm.

Algorithm Structure

1) Let $\xi' := \xi_{start}$; $x' := \arg\min\{\overline{P}_{\xi_{start}}\}$; $UB := \phi(x')$; $x^* := x'$; unbounded $:= false$; stop$:= false$;
2) While not stop do
   2a) With respect to ξ' and x' determine μ', λ', Δ_x, Δ_μ, Δ_λ, $\hat{F}$, $\hat{O}$; $\theta_m := \min\{\hat{F}, \hat{O}\}$;
   2b) If $\inf_{\theta\in[0,\theta_m]}\{z(\theta)\} = -\infty$ then unbounded$:= true$
   else $\overline{\theta} = \arg\min_{\theta\in[0,\theta_m]}\{z(\theta)\}$;
   2c) If unbounded$= true$ or $\{\phi(x)$ is convex and $\frac{dz}{d\theta}(0) > 0\}$
   then stop$:= true$
   else begin
   - If $z\left(\overline{\theta}\right) < UB$ then $x^* := x'\left(\overline{\theta}\right)$ and $UB := z\left(\overline{\theta}\right)$;
   - If $\xi' + \theta_m \geq \xi_{end}$ or $\{\phi(x)$ is convex and $0 < \overline{\theta} < \theta_m\}$
     then stop$:= true$
     else $x' := x' + \theta_m \Delta_x$; $\xi' := \xi' + \theta_m$;
   end;
3) If unbounded$= true$ then $\inf_{x\in X}\phi(x) = -\infty$ else x^* is the optimal solution for problem P.

Variable UB gives in the various iterations an upper bound for the optimal value with respect to the levels $\xi > \xi'$, while x^* is the best optimal level solution with respect to the levels $\xi \leq \xi'$. Let us also point out that:

- in *Step 1)* we have to determine the optimal solution of the strictly convex quadratic problem $\overline{P}_{\xi_{start}}$; actually, this is the only quadratic problem which needs to be solved within the solution algorithm;

- in *Step 2a)* the multipliers have to be determined by solving a system whose dimension has been reduced as much as possible (see Subsection 2.2 and system (8)); notice that these multipliers do not depend on the chosen functions f, g_1 and g_2; in the next section we will show that this step can be improved in the case of box constrained problems;
- in *Step 2b)* we have to determine the minimum of $z(\theta)$ for $\theta \in [0, \theta_m]$; notice that $z(\theta)$ is a single variable function and that its minimum over the segment $[0, \theta_m]$ can be computed with various numerical methods; notice also that *Step 2b)* is the only step which depends on the chosen functions f, g_1 and g_2;
- finally, it is worth noticing that for particular classes of functions this solution algorithm can be improved and detailed; in other words, for particular functions f, g_1 and g_2, the algorithm can be optimized for convex functions ϕ, and/or the multipliers in *Step 2a)* and the value of $\overline{\theta}$ in *Step 2b)* can be determined analytically (see for all [4, 6, 7, 8, 18]).

Once *Step 2b)* is implemented, the correctness of the proposed algorithm follows since all the feasible levels are scanned and the optimal solution, if it exists, is also an optimal level solution. As regards to the convergence (finiteness) of the procedure, note that in every iteration the set of binding constraints B changes; note also that the level is increased from ξ' to $\xi' + \theta_m$ so that it is not possible to obtain again an already used set of binding constraints B; the convergence then follows since we have a finite number of sets of binding constraints.

In particular, if $\theta_m = +\infty$ an halfline of optimal level solutions is found and the algorithm stops. Consider now the case $\theta_m < +\infty$; if $\theta_m = \hat{F}$ then at least one non binding constraint enters the set B; if $\theta_m = \hat{O}$ then at least one of the positive multipliers corresponding to a binding constraints vanishes, so that the related constraint will leave the set B in the following iteration.

4 Box Constrained Case

The aim of this section is to deepen the results stated so far in the particular case of box constrained problems:

$$P : \begin{cases} \inf \phi(x) = f\left(\frac{1}{2}x^T Q x + q^T x + q_0\right) g_1\left(d^T x + d_0\right) + g_2\left(d^T x + d_0\right) \\ \qquad x \in X^B = \{x \in \Re^n : l \leqq x \leqq u\} \end{cases}$$

where $l, u, d \in \Re^n$, $d \geq 0$ ([4]). Obviously, all the other hypotheses required in Section 1 are assumed too. By means of the general approach described in Section 2 we have:

[4] Notice that the $d \geq 0$ assumption is not restrictive, since it can be obtained by means of a trivial change of the variables x_i corresponding to the components $d_i < 0$.

$$\overline{P}_\xi : \begin{cases} \min \frac{1}{2}x^T Qx + q^T x + q_0 \\ x \in X_\xi^B = \{x \in \Re^n : l \leqq x \leqq u, d^T x + d_0 = \xi\} \end{cases}$$

Note that the feasible region X_ξ^B is no more given by box constraints.

Clearly, this class of box constrained problems can be solved by means of the solution algorithm described in Section 3. With this aim, notice that it results $\xi_{start} = \xi_{min} = d^T l + d_0$ and $\xi_{end} = \xi_{max} = d^T u + d_0$, and that the only strictly convex quadratic problem which has to be explicitly solved in *Step 1)* is:

$$\overline{P}_{\xi_{start}} : \begin{cases} \min \frac{1}{2}x^T Qx + q^T x + q_0 \\ x_i = l_i \quad \forall i = 1, \ldots, n \text{ such that } d_i > 0 \\ l_i \leq x_i \leq u_i \quad \forall i = 1, \ldots, n \text{ such that } d_i = 0 \end{cases}$$

In the rest of this section we will point out how the solution method can be improved in the case of box constrained problems, in particular with respect to the calculus of the multipliers in *Step 2a)*.

4.1 Incumbent Problem

Let x' be the optimal solution of problem $\overline{P}_{\xi'}$, with $\xi' = d^T x' + d_0 \in [\xi_{min}, \xi_{max}]$, and let us define, for the sake of convenience, the following partition $L \cup U \cup N \cup Z$ of the set of indices $\{1, \ldots, n\}$:

$$L = \{i : \ l_i = x'_i < u_i\} \quad , \quad N = \{i : \ l_i < x'_i < u_i\} \ ,$$
$$U = \{i : \ l_i < x'_i = u_i\} \quad , \quad E = \{i : \ l_i = x'_i = u_i\} \ .$$

Since $\overline{P}_{\xi'}$ is a strictly convex problem, x' is its unique optimal solution if and only if the following Karush-Kuhn-Tucker conditions hold ([5]):

$$\begin{cases} Qx' + q = \lambda d + \alpha - \beta & \\ d^T x' + d_0 = \xi', & \\ l \leqq x' \leqq u & \textit{feasibility} \\ \alpha \geqq 0, \ \beta \geqq 0, & \textit{optimality} \\ \alpha^T (x' - l) = 0, \ \beta^T (u - x') = 0 & \textit{complementarity} \\ \lambda \in \Re, \ \alpha, \beta \in \Re^n & \end{cases} \tag{9}$$

Denoting with Q_i the i-th row of Q, we can rewrite these Karush-Kuhn-Tucker conditions as follows:

[5] If $l \not< u$, that is $l_i = u_i$ for some indices i, the Karush-Kuhn-Tucker conditions are sufficient but not necessary since no constraint qualification conditions are verified. These indices will be handled implicitly in the rest of the paper by properly choosing the values of the multipliers.

$$\begin{cases} \alpha_i = 0,\ \beta_i = 0,\ Q_i x' + q_i = 0 & \forall i \in N \text{ s.t. } d_i = 0 \\ \alpha_i = 0,\ \beta_i = 0,\ \lambda = \frac{1}{d_i}\left(Q_i x' + q_i\right) & \forall i \in N \text{ s.t. } d_i \neq 0 \\ \beta_i = 0,\ \alpha_i = Q_i x' + q_i - \lambda d_i \geq 0 & \forall i \in L \\ \alpha_i = 0,\ \beta_i = \lambda d_i - Q_i x' - q_i \geq 0 & \forall i \in U \\ \alpha_i = \max\{0, Q_i x' + q_i - \lambda d_i\} \geq 0 & \forall i \in E \\ \beta_i = \max\{0, \lambda d_i - Q_i x' - q_i\} \geq 0 & \forall i \in E \\ d^T x + d_0 = \xi'\ ,\ l \leqq x \leqq u \end{cases}$$

Let $Z = \{i :\ d_i \neq 0\}$. Since $d \geq 0$ it results:

$$\begin{cases} \lambda = \frac{1}{d_i}\left(Q_i x' + q_i\right) \ \forall i \in N \cap Z \\ \lambda \leq \frac{1}{d_i}\left(Q_i x' + q_i\right) \ \forall i \in L \cap Z \\ \lambda \geq \frac{1}{d_i}\left(Q_i x' + q_i\right) \ \forall i \in U \cap Z \end{cases}$$

Given the optimal level solution x' for problem $P_{\xi'}$ the multipliers $\lambda', \alpha', \beta'$ can then be computed as follows. First, notice that when $(L \cup N \cup U) \cap Z = \emptyset$ then the linear function $d^T x + d_0$ is constant on the box feasible region, that is to say that the problem admits one unique feasible level and is then trivial.

Assuming $(L \cup N \cup U) \cap Z \neq \emptyset$, we can determine the value of λ' as described below:

$$\lambda' = \begin{cases} \frac{Q_i x' + q_i}{d_i},\ \text{for any } i \in N \cap Z & \text{if } N \cap Z \neq \emptyset \\ \min_{i \in L \cap Z} \left\{ \frac{Q_i x' + q_i}{d_i} \right\} & \text{if } N \cap Z = \emptyset \text{ and } L \cap Z \neq \emptyset \\ \max_{i \in U \cap Z} \left\{ \frac{Q_i x' + q_i}{d_i} \right\} & \text{if } N \cap Z = \emptyset \text{ and } L \cap Z = \emptyset \end{cases} \tag{10}$$

Then, the components of α' and β' can be obtained as follows:

$$\alpha'_i = \begin{cases} 0 & \forall i \in N \cup U \\ Q_i x' + q_i - \lambda' d_i & \forall i \in L \\ \max\{0, Q_i x' + q_i - \lambda' d_i\} & \forall i \in E \end{cases} \tag{11}$$

$$\beta'_i = \begin{cases} 0 & \forall i \in L \cup N \\ \lambda' d_i - Q_i x' - q_i & \forall i \in U \\ \max\{0, \lambda' d_i - Q_i x' - q_i\} & \forall i \in E \end{cases} \tag{12}$$

Let us remark that, unlike the general case of Subsection 2.1, we have been able to determine explicitly the values of all the multipliers of the Karush-Kuhn-Tucker conditions regarding to $\overline{P}_\xi$.

4.2 Sensitivity Analysis

In the light of the optimal level solution parametrical approach we now have to study the optimal solution of problem $\overline{P}_{\xi'+\theta}$, with $\theta > 0$. In order to avoid trivialities, we can assume $\xi' < \xi_{max}$. Since the Karush-Kuhn-Tucker system is linear whenever the complementarity conditions are implicitly handled, then the solution of the optimality conditions regarding to $\overline{P}_{\xi'+\theta}$ results:

$$x'(\theta) = x' + \theta\Delta_x \ , \ \lambda'(\theta) = \lambda' + \theta\Delta_\lambda \ ,$$
$$\alpha'(\theta) = \alpha' + \theta\Delta_\alpha \ , \ \beta'(\theta) = \beta' + \theta\Delta_\beta \ ,$$

so that it follows:

$$\begin{cases} Q(x' + \theta\Delta_x) + q = (\alpha' + \theta\Delta_\alpha) - (\beta' + \theta\Delta_\beta) + d(\lambda' + \theta\Delta_\lambda) \\ d^T(x' + \theta\Delta_x) + d_0 = \xi' + \theta \\ l \leqq x' + \theta\Delta_x \leqq u \\ \alpha' + \theta\Delta_\alpha \geqq 0, \ \beta' + \theta\Delta_\beta \geqq 0 \\ (\alpha' + \theta\Delta_\alpha)^T(x' + \theta\Delta_x - l) = 0, \ (\beta' + \theta\Delta_\beta)^T(u - x' - \theta\Delta_x) = 0 \end{cases} \tag{13}$$

Since x', λ', α' and β' are known, we are left to compute $\Delta_x, \Delta_\lambda, \Delta_\alpha, \Delta_\beta$. With this aim, let us provide the following lemma.

Lemma 2. *Let $(\lambda', \alpha', \beta')$ be a solution of (9). Then, for $\theta \in (0, \epsilon)$ system (13) is equivalent to:*

$$\begin{cases} Q\Delta_x = \Delta_\alpha - \Delta_\beta + d\Delta_\lambda \\ d^T\Delta_x = 1 \\ l \leqq x' + \theta\Delta_x \leqq u \\ \alpha' + \theta\Delta_\alpha \geqq 0, \ \beta' + \theta\Delta_\beta \geqq 0 \\ \Delta_{x_i} = 0 \ \forall i \in E \\ \Delta_{\alpha_i} = 0 \ \forall i \in N \cup U \\ \Delta_{\beta_i} = 0 \ \forall i \in L \cup N \\ \alpha'_i\Delta_{x_i} = 0 \ , \ \Delta_{\alpha_i}\Delta_{x_i} = 0 \quad \forall i \in L \\ \beta'_i\Delta_{x_i} = 0 \ , \ \Delta_{\beta_i}\Delta_{x_i} = 0 \quad \forall i \in U \end{cases} \tag{14}$$

Proof. The first and the second equations follow directly from (9) taking into account that $\theta \neq 0$, while $\Delta_{x_i} = 0 \ \forall i \in E$ follows directly from the definition of E. From (9) we have also that the complementarity conditions of (13) can be rewritten as:

$$\Delta_{\alpha_i}(x'_i - l_i) + \alpha'_i\Delta_{x_i} + \theta\Delta_{\alpha_i}\Delta_{x_i} = 0 \quad \forall i = 1, \dots, n \ ,$$
$$\Delta_{\beta_i}(u_i - x'_i) - \beta'_i\Delta_{x_i} - \theta\Delta_{\beta_i}\Delta_{x_i} = 0 \quad \forall i = 1, \dots, n \ .$$

Since $\theta \in (0, \epsilon)$ these conditions hold if and only if for all $i = 1, \dots, n$:

$$\Delta_{\alpha_i}\Delta_{x_i} = 0 \quad , \quad \Delta_{\beta_i}\Delta_{x_i} = 0 \ , \tag{15}$$
$$\Delta_{\alpha_i}(x'_i - l_i) + \alpha'_i\Delta_{x_i} = 0 \quad , \quad \Delta_{\beta_i}(u_i - x'_i) - \beta'_i\Delta_{x_i} = 0 \ . \tag{16}$$

Noticing that $x'_i + \theta\Delta_{x_i} < u_i$ for all $i \in L \cup N$ and for $\theta > 0$ small enough, from the complementarity conditions $(\beta'_i + \theta\Delta_{\beta_i})(u_i - x'_i - \theta\Delta_{x_i}) = 0$ it yields $\beta'_i + \theta\Delta_{\beta_i} = 0$; analogously, we also have $\alpha'_i + \theta\Delta_{\alpha_i} = 0$ for all $i \in U \cup N$. Since $\theta > 0$, for (11) and (12) these conditions imply:

$$\Delta_{\alpha_i} = 0 \ \forall i \in N \cup U \quad , \quad \Delta_{\beta_i} = 0 \ \forall i \in L \cup N \ ,$$

so that:

$$\Delta_{\alpha_i}(x'_i - l_i) = \Delta_{\beta_i}(u_i - x'_i) = 0 \quad \forall i = 1, \dots, n$$

and the result is proved.

Note that from the first and the second equations of (14) and from the positive definiteness of Q we obtain again $\Delta_x \neq 0$ and:

$$\Delta_\lambda = \Delta_x^T Q \Delta_x > 0 \ ,$$

while from (11), (12) and (14) we have:

$$\Delta_{\alpha_i} \geq 0 \ \forall i \in L \cup E \text{ such that } \alpha'_i = 0 \ ,$$
$$\Delta_{\beta_i} \geq 0 \ \forall i \in U \cup E \text{ such that } \beta'_i = 0 \ .$$

From the two last conditions of (14) it yields:

$$\Delta_{x_i} = 0 \ \forall i \in L \ s.t. \ \alpha'_i > 0 \ , \ \forall i \in U \ s.t. \ \beta'_i > 0 \ .$$

As a conclusion, we have the following explicit solution, depending on Δ_x, of the multipliers in (14):

$$\Delta_\lambda = \Delta_x^T Q \Delta_x$$
$$\Delta_{\alpha_i} = \begin{cases} 0 & \forall i \in N \cup U \\ Q_i \Delta_x - d_i \Delta_\lambda & \forall i \in L \\ \max\{0, Q_i \Delta_x - d_i \Delta_\lambda\} & \forall i \in E \end{cases}$$
$$\Delta_{\beta_i} = \begin{cases} 0 & \forall i \in L \cup N \\ d_i \Delta_\lambda - Q_i \Delta_x & \forall i \in U \\ \max\{0, d_i \Delta_\lambda - Q_i \Delta_x\} & \forall i \in E \end{cases}$$

In order to determine vector Δ_x it is worth using the partitions $L = L_p \cup L_0$ and $U = U_p \cup U_0$ defined as follows:

$$L_p = \{i \in L : \ \alpha'_i > 0\} \ , \ L_0 = \{i \in L : \ \alpha'_i = 0\} \ ,$$
$$U_p = \{i \in U : \ \beta'_i > 0\} \ , \ U_0 = \{i \in U : \ \beta'_i = 0\} \ .$$

Vector Δ_x is then the unique solution (recall that $\overline{P}_{\xi'+\theta}$ is a strictly convex problem) of the following system:

$$\begin{cases} Q_i \Delta_x = d_i \Delta_x^T Q \Delta_x \ \forall i \in N \\ d^T \Delta_x = 1 \\ \Delta_{x_i} = 0 \ \forall i \in L_p \ , \ \forall i \in U_p \ , \ \forall i \in E \\ (Q_i \Delta_x - d_i \Delta_x^T Q \Delta_x) \Delta_{x_i} = 0 \quad \forall i \in L_0 \cup U_0 \\ Q_i \Delta_x \geq d_i \Delta_x^T Q \Delta_x \ , \ \Delta_{x_i} \geq 0 \quad \forall i \in L_0 \\ d_i \Delta_x^T Q \Delta_x \geq Q_i \Delta_x \ , \ \Delta_{x_i} \leq 0 \quad \forall i \in U_0 \end{cases} \tag{17}$$

which is suitable for values of $\theta \geq 0$ which verify the following conditions:

$$\textit{feasibility conditions} : l \leqq (x' + \theta \Delta_x) \leqq u \ ,$$
$$\textit{optimality conditions} : \alpha'_i + \theta \Delta_{\alpha_i} \geq 0 \ \forall i \in L, \ \beta'_i + \theta \Delta_{\beta_i} \geq 0 \ \forall i \in U \ .$$

Notice that only the components Δ_{x_i} such that $i \in L_0 \cup N \cup U_0$ are left to be determined in (17). Notice also that the assumption $\xi' < \xi_{max}$ implies

$L \cup N \neq \emptyset$. We are finally able to determine the values of $\theta > 0$ which guarantee both the optimality and the feasibility of $x'(\theta)$. From the feasibility conditions we have:

$$\theta \leq \hat{F} = \min \left\{ \min_{i \in L_0 \cup N:\ \Delta_{x_i} > 0} \left\{ \frac{u_i - x'_i}{\Delta_{x_i}} \right\}, \min_{i \in N \cup U_0:\ \Delta_{x_i} < 0} \left\{ \frac{l_i - x'_i}{\Delta_{x_i}} \right\} \right\} .$$

Let us recall that whenever $\xi' < \xi_{max}$ then $\Delta_x \neq 0$, $L \cup N \neq \emptyset$ and hence $\hat{F} > 0$. On the other hand, denoting $L_p^- = \{i \in L_p : \ \Delta_{\alpha_i} < 0\}$ and $U_p^- = \{i \in U_p : \ \Delta_{\beta_i} < 0\}$, from the optimality conditions we have:

$$\theta \leq \hat{O} = \begin{cases} \min \left\{ \min_{i \in L_p^-} \left\{ \frac{\alpha'_i}{-\Delta_{\alpha_i}} \right\}, \min_{i \in U_p^-} \left\{ \frac{\beta'_i}{-\Delta_{\beta_i}} \right\} \right\} & \text{if } L_p^- \cup U_p^- \neq \emptyset \\ +\infty & \text{if } L_p^- \cup U_p^- = \emptyset \end{cases}$$

so that $\hat{O} > 0$.

As a consequence, $x'(\theta)$ is an optimal level solution for all θ such that:

$$0 \leq \theta \leq \theta_m = \min \left\{ \hat{F}, \hat{O} \right\}$$

where $\theta_m > 0$ whenever $\xi' < \xi_{max}$.

4.3 Box Constraints and Diagonal Matrix Q

In the case matrix Q is diagonal several further improvements can be done to the solution method. In particular, it is possible to explicitly determine all of the multipliers in the Karush-Kuhn-Tucker system. Notice that the particular case of $f(y) = y$, $g_1(y) = 1$ and $g_2(y) = \frac{1}{2}ky^2$ has been already studied in [8]. The following results can be proved analogously to the ones in [8]. First of all, notice that the parametric subproblem $\overline{P}_\xi$ becomes:

$$\overline{P}_\xi : \begin{cases} \min \frac{1}{2} x^T D x + q^T x + q_0 \\ x \in X_\xi^B = \{x \in \Re^n : l \leqq x \leqq u, d^T x + d_0 = \xi\} \end{cases}$$

where $D = \text{diag}(\delta_1, \ldots, \delta_n) \in \Re^{n \times n}$, $\delta_i > 0 \ \forall i = 1, \ldots, n$. As a preliminary result, it is worth pointing out that it is possible to determine explicitly the optimal value for all the variables x_i such that $d_i = 0$.

Theorem 2. *Consider the subproblems $\overline{P}_\xi$, with $\xi \in [\xi_{min}, \xi_{max}]$. Then, for all indices $i = 1, \ldots, n$ such that $d_i = 0$ the optimal level solution is reached at*

$$x_i^* = \begin{cases} l_i & \text{if} \quad -\frac{c_i}{\delta_i} \leq l_i \\ u_i & \text{if} \quad -\frac{c_i}{\delta_i} \geq u_i \\ -\frac{c_i}{\delta_i} & \text{if} \quad l_i < -\frac{c_i}{\delta_i} < u_i \end{cases}$$

As a consequence, the feasible region can be reduced *a priori*, without loosing the optimal solution, by means of the following commands:

- if $-\frac{c_i}{\delta_i} \le l_i$ then set $u_i := l_i$,
- if $-\frac{c_i}{\delta_i} \ge u_i$ then set $l_i := u_i$,
- if $l_i < -\frac{c_i}{\delta_i} < u_i$ then set $l_i := -\frac{c_i}{\delta_i}$ and $u_i := -\frac{c_i}{\delta_i}$.

From now on we can then assume that:

$$i \in E \text{ for all } i = 1, \dots, n \text{ such that } d_i = 0. \tag{18}$$

where $L \cup U \cup N \cup E = \{1, \dots, n\}$ with:

$$L = \{i : \; l_i = x'_i < u_i\} \quad , \quad N = \{i : \; l_i < x'_i < u_i\} \; ,$$
$$U = \{i : \; l_i < x'_i = u_i\} \quad , \quad E = \{i : \; l_i = x'_i = u_i\} \; .$$

Notice that assumption (18) implies also $X^B_{\xi_{min}} = \{l\}$ and $X^B_{\xi_{max}} = \{u\}$, so that there is no need to solve the starting quadratic problem $\overline{P}_{\xi_{start}}$ in *Step 1)* since we can simply choose $x' := l$.

By means of assumption (18) and the results stated in the previous subsections, the following explicit solutions of the Karush-Kuhn-Tucker systems can be determined:

$$\lambda' = \begin{cases} \frac{\delta_i x'_i + q_i}{d_i}, \text{ for any } i \in N & \text{if } N \neq \emptyset \\ \min_{i \in L} \left\{ \frac{\delta_i l_i + q_i}{d_i} \right\} & \text{if } N = \emptyset \text{ and } L \neq \emptyset \\ \max_{i \in U} \left\{ \frac{\delta_i u_i + q_i}{d_i} \right\} & \text{if } N = \emptyset \text{ and } L = \emptyset \end{cases}$$

$$\alpha'_i = \begin{cases} 0 & \forall i \in N \cup U \\ \delta_i l_i + q_i - \lambda' d_i & \forall i \in L \\ \max\{0, \delta_i l_i + q_i - \lambda' d_i\} & \forall i \in E \end{cases}$$

$$\beta'_i = \begin{cases} 0 & \forall i \in L \cup N \\ \lambda' d_i - \delta_i u_i - q_i & \forall i \in U \\ \max\{0, \lambda' d_i - \delta_i u_i - q_i\} & \forall i \in E \end{cases}$$

By defining the following further partition of indices $L = L^+ \cup L^0$:

$$L^+ = \{i \in L : \; \alpha'_i > 0\} \quad , \quad L^0 = \{i \in L : \; \alpha'_i = 0\} \; ,$$

we also have that:

$$\Delta_\lambda = \frac{1}{\sum_{i \in L^0 \cup N} \frac{1}{\delta_i} d_i^2} > 0$$

$$\Delta_{x_i} = \begin{cases} 0 & \text{if } i \in L^+ \cup U \cup E \\ \Delta_\lambda \frac{d_i}{\delta_i} > 0 & \text{if } i \in L^0 \cup N \end{cases}$$

$$\Delta_{\alpha_i} = \begin{cases} 0 & \text{if } i \notin L^+ \\ -\Delta_\lambda d_i < 0 & \text{if } i \in L^+ \end{cases}$$

$$\Delta_{\beta_i} = \begin{cases} 0 & \text{if } i \in L \cup N \\ \Delta_\lambda d_i \ge 0 & \text{if } i \in U \cup E \end{cases}$$

Finally, notice that it is:

$$\hat{F} = \begin{cases} \min_{i \in L^0 \cup N} \left\{ \frac{u_i - x'_i}{\Delta_{x_i}} \right\} & \text{if } L^0 \cup N \neq \emptyset \\ 0 & \text{if } L^0 \cup N = \emptyset \end{cases}$$

$$\hat{O} = \begin{cases} \min_{i \in L^+} \left\{ \frac{-\alpha'_i}{\Delta_{\alpha_i}} \right\} & \text{if } L^+ \neq \emptyset \\ +\infty & \text{if } L^+ = \emptyset \end{cases}$$

where $\theta_m > 0$ if and only if $x' \neq u$.

As a conclusion, let us point out that:

- in *Step 2a)* all of the parameters of the solution algorithm can be computed explicitly without the need of solving any further system;
- since $x'(\theta)$ and $\alpha'(\theta)$ are, respectively, increasing and decreasing with respect to θ (this follows from the nonnegativity of Δ_x and the nonpositivity of Δ_α) then, it can be proved that the algorithm stops after no more than $2n - 1$ iterations.

Acknowledgements

Careful reviews by the anonymous referees are gratefully acknowledged.

References

1. Barros AI (1998) Discrete and Fractional Programming Techniques for Location Models. Kluwer Academic Publishers, Dordrecht
2. Bomze IM, Csendes T, Horst R (eds) (1997) Developments in Global Optimization. Kluwer Academic Publishers, Dordrecht
3. Cambini A, Martein L (1986) A modified version of Martos' algorithm. Methods of Operation Research 53:33–44
4. Cambini A, Martein L, Sodini C (1983) An algorithm for two particular nonlinear fractional programs. Methods of Operations Research 45:61–70
5. Cambini R (1994) A class of non-linear programs: theoretical and algorithmic results. In: Komlósi S, Rapcsák T, Schaible S (eds) Generalized Convexity. Springer-Verlag, Berlin. Lecture Notes in Economics and Mathematical Systems 405:294–310
6. Cambini R, Sodini C (2002) A finite algorithm for a particular d.c. quadratic programming problem. Annals of Operations Research 117:33–49
7. Cambini R, Sodini C (2003) A finite algorithm for a class of nonlinear multiplicative programs. Journal of Global Optimization 26:279–296
8. Cambini R, Sodini C (2005) A sequential method for a class of box constrained quadratic programming problems. Report 272 Department of Statistics and Applied Mathematics
9. Cooper WW, Seiford LM, Zhu J (eds) (2004) Handbook on Data Envelopment Analysis. Kluwer Academic Publishers, Boston
10. Ellero A (1996) The optimal level solutions method. Journal of Information & Optimization Sciences 17:355–372

11. Frenk JBG, Schaible S (2005) Fractional programming. In: Handbook of Generalized Convexity and Generalized Monotonicity. Springer, New York. Nonconvex Optimization and Its Applications 76:335–386
12. Horst R, Pardalos PM (eds) (1995) Handbook of Global Optimization. Kluwer Academic Publishers, Dordrecht
13. Horst R, Tuy H (1996) Global Optimization: Deterministic Approaches, 3rd rev. Springer, Berlin
14. Horst R, Pardalos PM, Thoai NV (2001) Introduction to Global Optimization, 2nd ed. Kluwer Academic Publishers, Dordrecht
15. Mjelde KM (1983) Methods of the Allocation of Limited Resources. Wiley, New York
16. Schaible S (1995) Fractional programming. In: Horst R, Pardalos PM (eds) Handbook of Global Optimization. Kluwer Academic Publishers, Dordrecht. Nonconvex Optimization and Its Applications 2:495–608
17. Schaible S, Sodini C (1995) A finite algorithm for generalized linear multiplicative programming. Journal of Optimization Theory and Applications 87:441–455
18. Sodini C (1986) Minimizing the sum of a linear function and the square root of a convex quadratic form. Methods of Operations Research 53:171–182
19. Sodini C (1990) Equivalence and parametric analysis in linear fractional programming. In: Cambini A, Castagnoli E, Martein L, Mazzoleni P, Schaible S (eds) Generalized Convexity and Fractional Programming with Economic Applications. Springer-Verlag, Berlin. Lecture Notes in Economics and Mathematical Systems 345:143–154
20. Tuy H (1998) Convex Analysis and Global Optimization. Kluwer Academic Publishers, Dordrecht
21. Yitzakhi S (1982) A tax programming model. Journal of Public Economics 19:107–120

Mathematical Programming with (Φ, ρ)-invexity

Giuseppe Caristi[1], Massimiliano Ferrara[2], and Anton Stefanescu[3]

[1] University of Messina, Faculty of Economics gcaristi@dipmat.unime.it
[2] University of Messina, Faculty of Economics mferrara@unime.it
[3] University of Bucharest, Faculty of Mathematics and Computer Science, astefanescu@fmi.unibuc.ro

Summary. We introduce new invexity-type properties for differentiable functions, generalizing $(F, \rho)-$ convexity. Optimality conditions for nonlinear programming problems are established under such assumptions, extending previously known results. Wolfe and Mond-Weir duals are also considered, and we obtain direct and converse duality theorems.

Key words: Nonlinear programming, generalized invexity, duality.

1 Introduction

The theory of mathematical programming has grown remarkably after generalized convexity has been used in the settings of optimality conditions and duality theory. In 1981, Hanson [3] showed that both weak duality and Kuhn-Tucker sufficiency for optimum hold when convexity was replaced by a weaker condition. This condition, called invexity by Craven [1], was further studied for more general problems and was a source of a vast literature.

After the works of Hanson and Craven, other types of differentiable functions have been introduced with the intent of generalizing invex functions from different points of view. Hanson and Mond [4] introduced the concept of $F-$convexity and Jeyakumar [2] generalized Vial's $\rho-$convexity ([7]) introducing the concept of $\rho-$invexity. The concept of generalized $(F, \rho)-$convexity, introduced by Preda [6] is in turn an extension of the above properties and was used by several authors to obtain relevant results.

The $(F, \rho)-$convexity is now generalized to $(\Phi, \rho)-$invexity, and we will show that the main theoretical results of mathematical programming hold under this new condition.

2 $(\Phi, \rho)-$invexity

We begin by introducing a consistent notation for vector inequalities and for derivative operators.

In the following, $\mathbb{R}^n$ denotes the $n-$ dimensional Euclidean space. If $x, y \in \mathbb{R}^n$, then $x \geq y$ means $x_i \geq y_i$ for all $i = 1, 2, ..., n$, while $x > y$ means $x_i > y_i$ for all $i = 1, 2, ..., n$. An element of $\mathbb{R}^{n+1}$ may be regarded as (t, r) with $t \in \mathbb{R}^n$ and $r \in \mathbb{R}$.

Let $\varphi : D \subseteq \mathbb{R}^n \mapsto \mathbb{R}$ be a differentiable (twice differentiable) function of the independent variable x, and $a \in D$.

We will denote by $\nabla_x \varphi|_{x=a}$ the gradient of φ at the point a, and $\nabla^2_{xx}\varphi|_{x=a}$ stands for the matrix formed by the second order derivatives of φ. When any confusion is avoided, we will omit the subscript, writing simply $\nabla\varphi(a)$, respectively, $\nabla^2\varphi(a)$.

In the next definitions, ρ is a real number and Φ is a real-valued function defined on $D \times D \times \mathbb{R}^{n+1}$, such that $\Phi(x, a, .)$ is convex on $\mathbb{R}^{n+1}$ and $\Phi(x, a, (0, r)) \geq 0$ for every $(x, a) \in D \times D$ and $r \in \mathbb{R}_+$.

Definition 1. *We say that φ is $(\Phi, \rho) - invex$ at a with respect to $X \subseteq D$, if*

$$\varphi(x) - \varphi(a) \geq \Phi(x, a, (\nabla\varphi(a), \rho)), \forall x \in X \tag{1}$$

φ is $(\Phi, \rho) - invex$ on D if it is $(\Phi, \rho) - invex$ at a, for every $a \in D$.

Remark 1. If φ_1 is $(\Phi, \rho_1)-invex$ and φ_2 is $(\Phi, \rho_2)-invex$,then $\lambda\varphi_1+(1-\lambda)\varphi_2$ is $(\Phi, \lambda\rho_1 + (1-\lambda)\rho_2) - invex$, whenever $\lambda \in [0, 1]$. In particular, if φ_1 and φ_2 are $(\Phi, \rho)-invex$ with respect to the same Φ and ρ, then so is $\lambda\varphi_1+(1-\lambda)\varphi_2$.

The following two definitions generalizes $(\Phi, \rho)-$invexity.

Definition 2. *We say that φ is pseudo $(\Phi, \rho) - invex$ at a with respect to X, if whenever $\Phi(x, a, (\nabla\varphi(a), \rho)) \geq 0$ for some $x \in X$,then $\varphi(x) - \varphi(a) \geq 0$.*

Definition 3. *We say that φ is quasi $(\Phi, \rho) - invex$ at a with respect to X, if whenever $\varphi(x) - \varphi(a) \leq 0$ for some $x \in X$, then $\Phi(x, a, (\nabla\varphi(a), \rho)) \leq 0$,*

Remark 2. For $\Phi(x, a, (y, r)) = F(x, a, y) + rd^2(x, a)$, where $F(x, a, .)$ is sublinear on $\mathbb{R}^n$, the definition of $(\Phi, \rho)-$invexity reduces to the definition of $(F, \rho)-$convexity introduced by Preda [6], which in turn generalizes the concepts of $F-$convexity ([2]) and $\rho-$invexity ([7]).

More comments on the relationships between $(\Phi, \rho)-$invexity and invexity and their earlier extensions are in the next two sections

3 Optimality Conditions

The typical mathematical programming problem to be considered here is:

$$(P): \ \inf\{f(x) \mid x \in X_0, \ g_j(x) \leq 0, \ j = 1, 2, ..., m\}$$

where X_0 is a nonvoid open subset of $\mathbb{R}^n$, $f : X_0 \mapsto \mathbb{R}$, $g_j : X_0 \mapsto \mathbb{R}$, $j = 1, 2, ..., m$.

Let X be the set of all feasible solutions of (P);

$$X = \{x \in X_0, \ g_j(x) \leq 0, \ j = 1, 2, ..., m\}$$

Everywhere in this paper f and g_j, $j = 1, 2, ..., m$ are assumed to be differentiable on X_0, and we will refer to a Kuhn-Tucker point of (P) according to the usual definition.

Definition 4. *$(a, v) \in X \times \mathbb{R}^m_+$ is said to be a Kuhn-Tucker point of the problem (P) if:*

$$\nabla f(a) + \sum_{j=1}^{m} v_j \nabla g_j(a) = 0 \tag{2}$$

$$\sum_{j=1}^{m} v_j g_j(a) = 0 \tag{3}$$

Denoting by $J(a) = \{j \in \{1, 2, ..., m\} \mid g_j(a) = 0\}$, then summation in (2) and (3) is over $J(a)$.

First, we use $(\Phi, \rho)-$invexity to prove the sufficiency of Kuhn-Tucker conditions for the optimality in (P).

Everywhere in the following, we will assume invexity with respect to the set X of the feasible solutions of (P), but for the sake of simplicity we will omit to mention X.

Theorem 1. *Let (a, v) be a Kuhn-Tucker point of (P). If f is pseudo $(\Phi, \rho_0)-$invex at a, and for each $j \in J(a)$, g_j is quasi $(\Phi, \rho_j)-$invex at a, for some ρ_0, ρ_j, $j \in J(a)$ such that $\rho_0 + \sum_{j \in J(a)} v_j \rho_j \geq 0$, then a is an optimum solution of (P).*

Proof. Set $\lambda_0 = 1/(1 + \sum_{j=1}^{m} v_j)$, $\lambda_j = \lambda_0 v_j$, $j = 1, 2, ..., m$. Obviously,

$$\sum_{j=0}^{m} \lambda_j \rho_j = \lambda_0 \rho_0 + \sum_{j \in J(a)} \lambda_j \rho_j \geq 0$$

and .

$$\lambda_0 \nabla f(a) + \sum_{j \in J(a)} \lambda_j \nabla g_j(a) = 0$$

Then, it follows from the definition of Φ that

$$0 \leq \Phi(x, a, (\lambda_0 \nabla f(a) + \sum_{j \in J(a)} \lambda_j \nabla g_j(a), \lambda_0 \rho_0 + \sum_{j \in J(a)} \lambda_j \rho_j)) \leq$$

$$\lambda_0 \Phi(x, a, (\nabla f(a), \rho_0)) + \sum_{j \in J(a)} \lambda_j \Phi(x, a, (\nabla g_j(a), \rho_j))$$

for every $x \in \mathbb{R}^n$.

Now, let $x \in X$ be a feasible solution. Since $g_j(x) - g_j(a) \leq 0$ and g_j is quasi $(\Phi, \rho_j)-$invex , it results that $\Phi(x, a, (\nabla g_j(a), \rho_j)) \leq 0$, for each $j \in J(a)$. Hence, the above inequalities imply $\Phi(x, a, (\nabla f(a), \rho_0)) \geq 0$, and the pseudo $(\Phi, \rho_0)-$invexity of f implies $f(x) - f(a) \geq 0$.

With this theorem we have established that a sufficient condition for any Kuhn-Tucker point to be a minimum solution of (P) is that there exists a function Φ with the properties specified in Section 2, and a set of positive real numbers ρ_0, $\rho_1, ..., \rho_m$ such that for every $x, a \in X$,

$$\Phi(x, a, (\nabla g_j(a), \rho_j)) \leq 0, \forall j \in J(a) \tag{A$_1$}$$

$$\Phi(x, a, (\nabla f(a), \rho_0)) \geq 0 \Rightarrow f(x) - f(a) \geq 0 \tag{A$_2$}$$

Martin [5] obtained first a necessary and sufficient condition for the sufficiency of Kuhn-Tucker conditions in terms of modified invexity. He have established that any Kuhn-Tucker point is a minimum solution of (P) if and only if there exists a function $\eta : X_0 \times X_0 \mapsto \mathbb{R}$ such that for every $x, a \in X$,

$$\langle \eta(x, a), \nabla g_j(a) \rangle \leq 0, \forall j \in J(a) \tag{B$_1$}$$

$$f(x) - f(a) \geq \langle \eta(x, a), \nabla f(a) \rangle \tag{B$_2$}$$

Obviously, our conditions (A_1), (A_2) are weaker than those of Martin, and are satisfied whenever Martin's condition are satisfied, if Φ is defined by:

$$\Phi(x, a, (y, r)) = \langle \eta(x, a), y \rangle, \forall (x, a) \in X_0 \times X_0, y \in \mathbb{R}^n, r \in \mathbb{R}$$

On the other hand, if a is a Kuhn-Tucker point of (P) and (A_1), (A_2) hold for all $x \in X$, then a is a minimum point of f on X and $(B_1), (B_2)$ are also trivially satisfied for $\eta = 0$.

Likewise all earlier generalizations of the invexity, (Φ, ρ)-invexity reduces to invexity when it is used to establish optimality conditions. However, (Φ, ρ)-invexity enlarges the set of scale functions which can be used for proving the sufficiency of Kuhn-Tucker conditions. Moreover, as the example of the next section shows, the (Φ, ρ)-invexity could be strictly weaker that the invexity when duality conditions are checked.

Remark 3. Unlike Martin's conditions, where the properties of all functions involved in the problem, (f and g_j), are defined in respect to the same scale function η, our conditions are allowed to be satisfied for different scale functions. In fact, considering different values of ρ, f and each g_j should satisfy different invexity conditions. Thus in the definitions of Section 2, ρ should be interpreted as a parameter, and Φ generates a family of functions, one for each value of ρ. Similar situation appears in the case of $(F, \rho)-$convexity (or, $\rho-$invexity), but in that case the sign of ρ determines explicitly the properties of the function subjected to such condition. As we can observe in the proof of Theorem 1 (and in all results bellow), all that we need is that $\Phi(.,.,(0,r))$ is non-negative for some values of r. We have asked this condition to be satisfied whenever $r \geq 0$, but this is a convention which can be replaced by any other one.

Now, we will establish the necessity of Kuhn-Tucker conditions, under $(\Phi, \rho)-$invexity.

Theorem 2. *Let a be an optimum solution of (P). Suppose that Slater's constraint qualification holds for restrictions in $J(a)$ (i.e. there exists $x^* \in X_0$ such that $g_j(x^*) < 0$, for all $j \in J(a)$). If, for each $j \in J(a)$, g_j is $(\Phi, \rho_j)-$invex at a for some $\rho_j \geq 0$, then there exists $v \in \mathbb{R}^m_+$ such that (a, v) is a Kuhn-Tucker point of (P).*

Proof. Since f and g_j are differentiable, then there exist Fritz-John multipliers $\mu \in \mathbb{R}_+$ and $\lambda \in \mathbb{R}^m_+$ such that:

$$\mu \nabla f(a) + \sum_{j=1}^{m} \lambda_j \nabla g_j(a) = 0 \tag{4}$$

$$\sum_{j=1}^{m} \lambda_j g_j(a) = 0 \tag{5}$$

$$\mu + \sum_{j=1}^{m} \lambda_j > 0 \tag{6}$$

All that we need is to prove that $\mu > 0$.

Suppose, by way of contradiction, that $\mu = 0$. Then, $\sum_{j \in J(a)} \lambda_j > 0$ from (6), and we can define $\mu_j = \lambda_j / \sum_{j \in J(a)} \lambda_j$. Obviously, $\sum_{j \in J(a)} \mu_j g_j(a) = 0$ and $\sum_{j \in J(a)} \mu_j \rho_j \geq 0$.

Hence, since each g_j is $(\Phi, \rho_j)-$invex,

$$0 \leq \Phi(x^*, a, (\sum_{j \in J(a)} \mu_j \nabla g_j(a), \sum_{j \in J(a)} \mu_j \rho_j)) \leq$$

$$\sum_{j \in J(a)} \mu_j \Phi(x^*, a, (\nabla g_j(a), \rho_j)) \leq \sum_{j \in J(a)} \mu_j (g_j(x^*) - g_j(a))$$

But $\sum_{j\in J(a)} \mu_j g_j(a) = 0$ by (5), so that $\sum_{j\in J(a)} \mu_j(g_j(x^*) - g_j(a)) < 0$, contradicting the above inequalities.

4 Wolfe Type Duality

Let us consider the Wolfe dual of (P) :

$$(WD) : \sup\{f(y) + \sum_{j=1}^{m} v_j g_j(y) \mid y \in X_0, v \in \mathbb{R}_+^m, \nabla f(y) + \sum_{j=1}^{m} v_j \nabla g_j(y) = 0\}$$

The main duality results also hold under our invexity type conditions. We establish first, the general duality property.

Theorem 3. *Let (y, v) be a feasible solution of (WD). If f is $(\Phi, \rho_0)-$ invex at y, each g_j is (Φ, ρ_j) -invex at y, and $\rho_0 + \sum_{j=1}^{m} v_j \rho_j \geq 0$, then*

$$f(x) \geq f(y) + \sum_{j=1}^{m} v_j g_j(y) \tag{7}$$

for every feasible solution $x \in X$ of (P).

Proof. Likewise in the proof of Theorem 1, setting $\lambda_0 = 1/(1 + \sum_{j=1}^{m} v_j)$, $\lambda_j = \lambda_0 v_j$, $j = 1, 2, ..., m$, it follows that

$$0 \leq \lambda_0 \Phi(x, y, (\nabla f(y), \rho_0)) + \sum_{j=1}^{m} \lambda_j \Phi(x, y, (\nabla g_j(y), \rho_j)) \tag{8}$$

Further, since f is $(\Phi, \rho_0)-$ invex and g_j is (Φ, ρ_j) -invex, it results:

$$0 \leq \lambda_0(f(x) - f(y)) + \sum_{j=1}^{m} \lambda_j(g_j(x) - g_j(y))$$

Hence,

$$f(x) + \sum_{j=1}^{m} v_j g_j(x) \geq f(y)) + \sum_{j=1}^{m} v_j g_j(y)$$

and, since $x \in X$, the inequality (7) holds.

Corollary 1. *If the equality holds in (7), then x is optimal for (P). Moreover, if f is $(\Phi, \rho_0)-$ invex on X_0, each g_j is (Φ, ρ_j) -invex on X_0, and the equality holds in (7), then (y, v) is also optimal for (WD).*

The example below shows that the weak duality holds when assumptions of the previous theorem are verified, even if the usual invexity conditions fail. Thus, (Φ, ρ_j) -invexity is actually strictly weaker than the invexity. For the sake of simplicity, we have considered a pathological optimization problem where X is a singleton.

Example 1. Let us consider the problem (P) defined in $\mathbb{R}^2$ by the objective function

$$f(x) = -(x_1 + 4 \times 10^{-3})(x_2 + 4 \times 10^{-3})$$

and the three restriction functions:

$$\begin{aligned} g_1(x) &= (x_1 + 4.5 \times 10^{-3})(x_2 + 5.5 \times 10^{-3}) - 35.75 \times 10^{-6} \\ g_2(x) &= (x_1 + 5.5 \times 10^{-3})(x_2 + 4.5 \times 10^{-3}) - 35.75 \times 10^{-6} \\ g_3(x) &= (x_1 - 5 \times 10^{-3})^2 + (x_2 - 5 \times 10^{-3})^2 - 32 \times 10^{-6} \end{aligned}$$

It is easy to check that $X = \{x^0\}$, where $x^0 = (10^{-3}, 10^{-3})$, and (a, ν) is a feasible solution of the dual, where $a = (2\times10^{-3}, 2\times10^{-3})$, and $\nu = (\frac{9}{14}, \frac{9}{14}, \frac{1}{2})$.

Let Φ be defined on $\mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}^3$ by

$$\Phi(x, a, (t, r)) = r + 1 + (t_1^2 + t_2^2) - \sqrt{(t_1 + t_2)^2 + 1}$$

Φ has all properties required in Definition 1, f is $(\Phi, \rho_0)-$invex and each g_j is $(\Phi, \rho_j)-$invex at a, with respect to X, for $\rho_0 = 10.9 \times 10^{-6}$, $\rho_1 = \rho_2 = -13.6 \times 10^{-6}$ and $\rho_3 = 13.9 \times 10^{-6}$. Since $\rho_0 + \sum_{j=1}^{3} \nu_j \rho_j \geq 0$, all assumptions of the theorem are satisfied.

Hence, inequality (7) holds at (a, ν).

To illustrate the elementary but not easy calculation needed for the above assertion, consider the case of f. By definition, f is $(\Phi, \rho_0)-$invex at a iff

$$\rho_0 + 1 + 72 \times 10^{-6} - \sqrt{144 \times 10^{-6} + 1} \leq 11 \times 10^{-6}$$

Since $\sqrt{144 \times 10^{-6} + 1} \geq 71.9 \times 10^{-6} + 1$, it follows that $\rho_0 = 10.9 \times 10^{-6}$ satisfies this inequality.

On the other hand, the functions involved in this problem are not invex at a, for any scale function η. For, if such a function exists then the following inequalities should be simultaneously satisfied:

$$\begin{aligned} \left\langle \eta(x^0, a), \nabla f(a) \right\rangle &\leq f(x^0) - f(a) \\ \left\langle \eta(x^0, a), \nabla g_1(a) \right\rangle &\leq g_1(x^0) - g_1(a) \\ \left\langle \eta(x^0, a), \nabla g_2(a) \right\rangle &\leq g_2(x^0) - g_2(a) \end{aligned}$$

Since $-\nabla f(a) = \frac{6}{14}(\nabla g_1(a) + \nabla g_2(a))$, $f(x^0) - f(a) = 11 \times 10^{-6}$, $g_1(x^0) - g_1(a) = g_2(x^0) - g_2(a) = -13 \times 10^{-6}$, we arrive to the impossible inequality $-11 \leq \frac{6}{14}(-26)$.

Remark 4. Following the same line as in the above, it is easy to verify that the functions f, g_1 and g_2 are not $(F,\rho)-$ convex at a, for positive values of ρ. Obviously, so are f, g_1, g_2 and g_3, but our example is not conclusive for the $(F,\rho)-$ convexity when the values of ρ are restricted only to the inequality $\rho_0 + \sum_{j=1}^{3} \nu_j \rho_j \geq 0$.

Now, let us establish a direct duality result.

Theorem 4. *Let a be an optimum solution of (P). Assume that Slater's constraint qualification holds. If f is $(\Phi,\rho_0)-$ invex and each g_j is (Φ,ρ_j) -invex on X_0, for some $\rho_0, \rho_1, ..., \rho_m \geq 0$, then there exists $v \in \mathbb{R}_+^m$, such that (a,v) is an optimum solution of (WD).*

Proof. According with Theorem 2, there exists $v \in \mathbb{R}_+^m$ such that (a,v) is a Kuhn-Tucker point of (P). Therefore, (a,v) is a feasible solution of (WD) and $\sum_{j=1}^{m} v_j g_j(a) = 0$. Then, as it was stated in Corollary 1, (a,v) is optimal for (WD).

The converse duality theorem can be also proved under $(\Phi,\rho)-$ invexity.

Let us denote, as usual, by L the Lagrangian of (P), $L(y,v) = f(y) + \sum_{j=1}^{m} g_j(y)$.

Theorem 5. *Let (y^*, v^*) be an optimum solution of (WD). Suppose that f and g_j are twice differentiable and* $\det(\nabla_{yy}^2 L(y,v^*)|_{y=y^*}) \neq 0$. *If f is $(\Phi,\rho_0)-$ invex on X_0, each g_j is (Φ,ρ_j) -invex on X_0, and $\rho_0 + \sum_{j=1}^{m} v_j^* \rho_j \geq 0$, then y^* is an optimum solution of (P).*

Proof. Since (y^*, v^*) is a feasible solution of (WD), it is a solution of the system:

$$\nabla_y L(y,v) = 0, \ (y,v) \in X_0 \times \mathbb{R}_+^m$$

Since $\det(\nabla_{yy}^2 L(y,v^*)|_{y=y^*}) \neq 0$, this system can be explicitly solved with respect to y, in some neighborhood of v^*. Therefore, there exists the open neighborhood $V_{v^*} \subseteq \mathbb{R}_+^m$of v^* and the continuous function $y : V_{v^*} \to X_0$ such that $y(v^*) = y^*$ and

$$\nabla_y L(y(v),v) = 0, \forall v \in V_{v^*} \tag{9}$$

Particularly, this means that $(y(v),v)$ is a feasible solution of (WD), for every $v \in V_{v^*}$. Now, since (y^*,v^*) is optimal for (WD), it follows that v^* maximizes $L(y(v),v)$ on V_{v^*}. Then,

$$\nabla_v L(y(v),v)|_{v=v^*} \leq 0 \tag{10}$$

and

$$\langle v^*, \nabla_v L(y(v),v)|_{v=v^*} \rangle = 0 \tag{11}$$

But,

$$\nabla_v L(y(v), v) = \left\langle \nabla_y L(y, v)|_{y=y(v)}, \nabla_v y(v) \right\rangle + \nabla_v L(y, v)|_{y=y(v)}$$

and then, it follows from (9) that

$$\nabla_v L(y(v), v)|_{v=v^*} = \nabla_v L(y^*, v^*)$$

Thus, by (10), y^* should satisfy the inequalities $g_j(y^*) \leq 0$, $j = 1, 2, ..., m$, and by (11) it should satisfy the equality $\sum_{j=1}^{m} v_j^* g_j(y^*) = 0$. Subsequently, y^* is a feasible solution of (P), and $f(y^*) = L(y^*, v^*)$. Then, Theorem 3, (Corollary 1), shows that y^* is optimal for (P).

5 Mond - Weir Duality

Consider now the Mond-Weir dual of (P).

$$(MWD) : \sup\{f(y)|y \in X_0,\ v \in \mathbb{R}_+^m,\ \nabla f(y) + \textstyle\sum_{j=1}^{m} v_j \nabla g_j(y) = 0, \sum_{j=1}^{m} v_j g_j(y) = 0\}$$

Theorem 6. *Let* (y, v) *be a feasible solution of* (MWD). *If* f *is pseudo* $(\Phi, \rho_0)-$ *invex at* y, *each* g_j *is* (Φ, ρ_j) *-invex at* y, *and* $\rho_0 + \sum_{j=1}^{m} v_j \rho_j \geq 0$, *then*

$$f(x) \geq f(y) \tag{12}$$

for all $x \in X$.

Proof. As in the above, inequality (8) follows from the properties of Φ. Since g_j is (Φ, ρ_j) -invex at y, and x and (y, v) are feasible it results:

$$\sum_{j=1}^{m} \lambda_j \Phi(x, y, (\nabla g_j(y), \rho_j)) \leq \sum_{j=1}^{m} \lambda_j (g_j(x) - g_j(y) = \sum_{j=1}^{m} \lambda_j g_j(x) \leq 0$$

Thus, $\Phi(x, y, (\nabla f(y), \rho_0)) \geq 0$, and then the pseudo $(\Phi, \rho_0)-$invexity of f, give us $f(x) - f(y) \geq 0$.

Corollary 2. *If the equality holds in* (12), *then* x *is optimal for* (P). *Moreover, if* f *is pseudo* $(\Phi, \rho_0)-$ *invex on* X_0, *each* g_j *is* (Φ, ρ_j) *-invex on* X_0, *and the equality holds in* (12), *then* (y, v) *is also optimal for* (MWD).

With a slight modification of the proof of Theorem 4, we state the following result.

Theorem 7. *Let* a *be an optimum solution of* (P). *Assume that Slater's constraint qualification holds. If* f *is* $(\Phi, \rho_0)-$ *invex and each* g_j *is* (Φ, ρ_j) *-invex on* X_0, *for some* $\rho_0, \rho_1, ..., \rho_m \geq 0$, *then there exists* $v \in \mathbb{R}_+^m$, *such that* (a, v) *is an optimum solution of* (WD).

References

1. Craven BD (1981) Invex functions and constrained local minima. Bull. Austral. Math. Soc. 24: 357-366
2. Jeyakumar Y (1985) Strong and weak invexity in mathematical programming. In: Methods of Operations Research, Vol. 55: 109-125
3. Hanson MA (1981) On sufficiency of Kuhn Tucker conditions. J. Math. Anal. Appl. 30: 545-550
4. Hanson MA, Mond B (1982) Further generalization of convexity in mathematical programming, J. Inform. Optim. Sci. 3: 22-35
5. Martin DH (1985) The essence of invexity. J. Opt. Theory. Appl. 47: 65-76
6. Preda V (1992) On efficiency and duality for multiobjective programs, J. Math. Anal. Appl. 166: 365-377
7. Vial JP (1983) Strong and weak convexity of sets and functions, Math. Oper. Res. 8: 231-259
8. Weir T, Mond B (1989) Generalized convexity and duality in multiple objective programming, Bull. Austral. Math. Soc. 39: 287-299

Some Classes of Pseudoconvex Fractional Functions via the Charnes-Cooper Transformation

Laura Carosi[1] and Laura Martein[2]

[1] Department of Statistics and Applied Mathematics, University of Pisa Italy
lcarosi@ec.unipi.it
[2] Department of Statistics and Applied Mathematics, University of Pisa Italy
lmartein@ec.unipi.it

Summary. Using a very recent approach based on the Charnes-Cooper trasformation we characterize the pseudoconvexity of the sum between a quadratic fractional function and a linear one. Furthemore we prove that the ratio between a quadratic fractional function and the cube of an affine one is pseudoconvex if and only if the product between a quadratic fractional function and an affine one is pseudoconvex and we provide a sort of canonical form for this latter class of functions. Benefiting by the new results we are able to characterize the pseudoconvexity of the ratio between a quadratic fractional function and the cube of an affine one.

Key words: Pseudoconvexity, fractional programming, quadratic programming.

1 Introduction

Since the early sixties, the strict relationship between generalized convexity and fractional programming has been highlighted and from the beginning, fractional programming has benefited from advances in generalized convexity, and vice versa (see for instance [14, 16]). Generalized fractional programming and in particular quadratic and multiplicative fractional programming are extremely important even for their numerous applications such as Data Envelopment Analysis, tax programming, risk and portfolio theory, logistics and location theory (see for instance [2, 3, 11, 12, 15]). Among the different classes of generalized convex functions, the pseudoconvex one occupies a leading position in optimization for its good properties. Nevertheless pseudoconvex functions have no algebraic structure and this lack of structure causes many difficulties to establish whether a function is pseudoconvex or not.

There are several characterizations for continuously differentiable function and for twice differentiable functions [1, 13]. Since these conditions are not very easy to be checked, some more operative characterizations, dealing with quadratic fractional functions, have been recently proposed [4, 8]. According with a very recent approach the pseudoconvexity of fractional functions is studied by means of the generalized Charnes-Cooper transformation. It is shown [6, 7] that this transformation maintains the pseudoconvexity so that the fundamental idea behind this study is the following: if an unknown class of functions can be transformed in a class of pseudoconvex functions, it is possible to determine necessary and sufficient conditions guaranteeing the pseudoconvexity of the unknown class of functions.

Following this idea, we prove that the sum between a quadratic fractional function and a linear one is pseudoconvex if and only if a suitable quadratic fractional function is pseudoconvex. Therefore, using the known results for this latter class of functions we establish a new characterization and we give a simple algorithm in order to test the pseudoconvexity for the sum between a quadratic fractional function and a linear one.

Furthermore we address our attention to the pseudoconvexity of the ratio between a quadratic function and the power p of an affine one. Since the cases $p = 1$ and $p = 2$ have been handled in [8, 9] we deal with the case $p = 3$. Performing the Charnes-Cooper transformation, we prove that this class of functions is pseudoconvex if and only if the product between the quadratic function and a suitable affine one is pseudoconvex. As far as we know, even for this latter class of functions there are no easy to be checked conditions for testing the pseudoconvexity. Consequently, we first characterize the pseudoconvexity for the product between a quadratic function and an affine one: more precisely we prove that a function belonging to this class is pseudoconvex if and only if it has a suitable canonical form. The obtained result allows to provide a new characterization for the ratio between a quadratic function and the cube of an affine one.

2 Preliminary Results and Notations

Throughout the paper we will use the following notations and properties.

- A is a $n \times n$ symmetric matrix such that $A \neq [0]$ where $[0]$ is the null matrix;
- $\nu_-(A)$ ($\nu_+(A)$) denotes the number of negative (positive) eigenvalues of a matrix A;
- $\ker A$ denotes the kernel of A i.e., $\ker A = \{v : Av = 0\}$;
- $\dim W$ denotes the dimension of the vector space W;
- $\mathrm{Im}A$ denotes the set $\mathrm{Im}A = \{z = Av,\ v \in \Re^s\}$;
- $v^\perp$ denotes the orthogonal space to a vector v i.e., $v^\perp = \{w : v^T w = 0\}$.

For the sake of completeness we recall the definition of pseudoconvex functions and the related properties we are going to use in the next section (for further details see for instance [1]).

Definition 1. *Let f be a differentiable function on the open and convex set $C \subseteq \Re^n$. f is pseudoconvex if for $x, y \in C$*

$$f(y) < f(x) \text{ implies that } \nabla f(x)^T (y - x) < 0.$$

- f is pseudoconvex if and only if for every $x_0, v \in \Re^n$ the restriction of f on the line $x = x_0 + tv$, $t \in \Re$, is pseudoconvex.
- Let $C \subseteq \Re^n$ an open and convex set f is pseudoconvex if and only if $\forall x \in C, \forall v \in \Re^n \backslash \{0\}$, such that $\nabla f(x)^T v = 0$ the function $\varphi(t) = f(x+tv)$ attains a local minimum at $t = 0$.
- Let $C \subseteq \Re^n$ an open and convex set and let f be a twice continuously differentiable. f is pseudoconvex if and only if $\forall x_0 \in C$, $\forall v \in \Re^n \setminus \{0\}$, such that $\nabla f(x_0)^T v = 0$ either $v^T H(x_0) v > 0$ or $v^T H(x_0) v = 0$ and the function $\varphi(t) = f(x^0 + tv)$ attains a local minimum at $t = 0$.

Consider the Charnes-Cooper transformation [10]

$$y(x) = \frac{x}{b^T x + b_0} \tag{1}$$

defined on the set $S = \{x \in \Re^n : b^T x + b_0 > 0\}$ where $b \in \Re^n$ and $b_0 \in \Re$, $b_0 \neq 0$. It is well known that this map is a diffeomorphism and its inverse is

$$x(y) = \frac{b_0 y}{1 - b^T y} \tag{2}$$

defined on the set $S^* = \{y \in \Re^n : \frac{b_0}{1-b^T y} > 0\}$. As it is shown in [6, 7] the Charnes-Cooper transformation preserves the pseudoconvexity of f. More precisely the following theorem holds.

Theorem 1. *Let f be a differentiable function defined on $\Re^n$ and let $\psi(y)$ be the function obtained by applying the inverse of the Charnes-Cooper transformation (2) to $f(x)$.*
Function $f(x)$ is pseudoconvex on S if and only if function $\psi(y)$ is pseudoconvex on S^*.

In some cases, the study of the pseudoconvexity of the transformed function $\psi(y)$ may be easier than the study of the pseudoconvexity of f. Therefore, thanks to the previous theorem, by means of the results on $\psi(y)$ we can characterize the pseudoconvexity of f in terms of its initial data. Following this approach in the next section we aim to study the pseudoconvexity of some classes of generalized quadratic fractional functions.

The following Lemma will be also useful.

Lemma 1. *Consider a non-null symmetric matrix A of order n and a non-null vector $a \in \Re^n$. Then there exists $d \in \Re^n$ such that $d^T A d \neq 0$ and $a^T d \neq 0$.*

Proof. Suppose on the contrary for every $d \in \Re^n$ $a^T d \neq 0$ implies $d^T A d = 0$. Since $a \neq 0$, setting $d = a$ we get $\|a\|^2 \neq 0$ and hence $a^T A a = 0$. Take $x = ta + w$, $w \in a^{\perp}$,$t \in \Re$; we have

$$\frac{1}{2} x^T A x = \frac{1}{2}(ta + w)^T A (ta + w) = t a^T A w + \frac{1}{2} w^T A w.$$

Since $a^T x = t\|a\|^2 \neq 0$ for every $t \neq 0$ it results $t a^T A w + \frac{1}{2} w^T A w = 0$ for every $t \neq 0$. Then necessarily we have $a^T A w = 0$ and $w^T A w = 0$ for every $w \in a^{\perp}$. From the second equality it follows $Aw = ka$ and since $a^T A w = k\|a\|^2 = 0$ we obtain $k = 0$, so that $Aw = 0$ for every $w \in a^{\perp}$. Taking into account that $A \neq [0]$, we get $A = \lambda a a^T$ and then $a^T A a = \lambda \|a\|^2 \neq 0$ which is a contradiction.

3 New Classes of Pseudoconvex Fractional Functions

3.1 Pseudoconvexity of the Sum Between a Quadratic Fractional Function and a Linear One

Consider the following function

$$f(x) = \frac{\frac{1}{2} x^T A x}{b^T x + b_0} + p^T x \tag{3}$$

on the halfspace $S = \{x \in \Re^n : b^T x + b_0 > 0\}$, $b_0 \neq 0$. Performing the Charnes-Cooper transformation (2) we obtain the following function defined on the halfspace $S^* = \{y \in \Re^n : \frac{b_0}{1 - b^T y} > 0\}$

$$g(y) = \frac{\frac{b_0^2}{2(1-b^T y)^2} y^T A y}{\frac{b_0}{1-b^T y}} + \frac{b_0}{1 - b^T y} p^T y = \frac{b_0}{1 - b^T y}\left(\frac{1}{2} y^T A y + p^T y\right)$$

that is, setting $c = -\frac{b}{b_0}$, $c_0 = \frac{1}{b_0}$

$$g(y) = \frac{\frac{1}{2} y^T A y + p^T y}{c^T y + c_0},\ y \in S^*. \tag{4}$$

From Theorem 1, the pseudoconvexity of f on S is equivalent to the pseudoconvexity of g on S^*. A characterization of the pseudoconvexity for such a class of functions is given in [4]. More precisely the following theorem holds.

Theorem 2. *Consider function* $g(y) = \frac{\frac{1}{2} y^T A y + p^T y}{c^T y + c_0}$ *on the halfspace* $S^* = \{y \in \Re^n : c^T y + c_0 > 0\}$, $c_0 \neq 0$. *g is pseudoconvex if and only if one of the following conditions holds:*

a) $\nu_-(A) = 0$,

b) $\nu_-(A) = 1$, $\exists \overline{x}, \overline{y} \in \Re^n$ *such that* $A\overline{x} = p$ *and* $A\overline{y} = c$, $c^T\overline{y} = 0$, $c^T\overline{x} = c_0$ *and* $p^T\overline{x} \leq 0$;
c) $\nu_-(A) = 1$, $\exists \overline{x}, \overline{y} \in \Re^n$ *such that* $A\overline{x} = p$ *and* $A\overline{y} = c$, $c^T\overline{y} < 0$ *and* $\frac{\Delta}{4} = (c_0 - c^T\overline{x})^2 - c^T\overline{y}(p^T\overline{x}) \leq 0$.

Thanks to the Charnes-Cooper transformation, Theorem 2 allows us to characterize the pseudoconvexity of f in term of its initial data.

Theorem 3. *Consider function* $f(x) = \frac{\frac{1}{2}x^TAx}{b^Tx+b_0} + p^Tx$ *on the set* $S = \{x \in \Re^n : b^Tx + b_0 > 0\}$, $b_0 \neq 0$. f *is pseudoconvex if and only if one of the following conditions holds:*

a) $\nu_-(A) = 0$,
b) $\nu_-(A) = 1$, $\exists \overline{x}, \overline{z} \in \Re^n$ *such that* $A\overline{x} = p$ *and* $A\overline{z} = b$, $b^T\overline{z} = 0$, $b^T\overline{x} = -1$ *and* $p^T\overline{x} \leq 0$;
c) $\nu_-(A) = 1$, $\exists \overline{x}, \overline{z} \in \Re^n$ *such that* $A\overline{x} = p$ *and* $A\overline{z} = b$, $b^T\overline{z} < 0$ *and* $\frac{\Delta}{4} = (1 + b^T\overline{x})^2 - b^T\overline{z}(p^T\overline{x}) \leq 0$.

Proof. From Theorem 1 f is pseudoconvex on S if and only if g is pseudoconvex on S^* and so if and only if one of conditions a), b), c) in Theorem 2 holds. Recalling that $b = -\frac{c}{c_0}$, $b_0 = \frac{1}{c_0}$, by means of simple calculations it can be proved that conditions a), b), c) are equivalent to the corresponding ones given in Theorem 2.

The following example shows that function $f(x)$ in (3) can be pseudoconvex even if the fractional quadratic function is not pseudoconvex.

Example 1. Consider function $f(x,y) = \frac{x^2-y^2}{-x+y+2} + x + y$, that is $A = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$, $p = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $b = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$. It is easy to verify that $\frac{x^2-y^2}{-x+y+2}$ is not pseudoconvex on S. On the other hand $v_-(A) = v_+(A) = 1$, $\bar{z} = \left(-\frac{1}{2}-, \frac{1}{2}\right)^T$, $\bar{x} = \left(\frac{1}{2}, -\frac{1}{2}\right)^T$ and so $b^T\bar{z} = 0$, $b^T\bar{x} = -1$ and $p^T\bar{x} = 0$. Hence condition b) in Theorem 3 holds and f is pseudoconvex on S.

According with the previous result, we suggest the following algorithm for testing the pseudoconvexity of the sum between a quadratic fractional function and an affine one.

ALGORITHM

STEP 1

Calculate the eigenvalues of A. If $\nu_-(A) > 1$, STOP : f is not pseudoconvex. If $\nu_-(A) = 0$, STOP : f is pseudoconvex; otherwise go to STEP 2.

STEP 2

Solve the linear systems $Ax = p$ and $Az = b$. If one of these systems has no solutions STOP: f is not pseudoconvex; otherwise go to STEP 3.

STEP 3

Let $\bar{z}$ such that $A\bar{z} = b$. Calculate $b^T\bar{z}$. If $b^T\bar{z} > 0$ STOP : f is not pseudoconvex. If $b^T\bar{z} = 0$ go to STEP 4, otherwise go to STEP 5.

STEP 4

Let $\bar{x}$ such that $A\bar{x} = p$. Calculate $b^T\bar{x}$. If $b^T\bar{x} \neq -1$ STOP: f is not pseudoconvex, otherwise calculate $p^T\bar{x}$. If $p^T\bar{x} > 0$ STOP: f is not pseudoconvex otherwise STOP: f is pseudoconvex.

STEP 5

Let $\bar{x}$ such that $A\bar{x} = p$. Calculate $\frac{\Delta}{4} = (1 + b^T\bar{x})^2 - b^T\bar{z}(p^T\bar{x})$. If $\Delta > 0$ STOP : f is not pseudoconvex otherwise f is pseudoconvex.

3.2 Pseudoconvexity of the Ratio Between a Quadratic Function and the Cube of an Affine One

Consider now the following function

$$h(x) = \frac{1}{2}\frac{x^T Ax}{(b^T x + b_0)^p}\ x \in S$$

where $p \in \mathbf{N}\backslash\{0\}$, $b_0 \neq 0$.
Performing the transformation (2) we get

$$g\left(y\right) = \frac{1}{2}\frac{b_0^2}{\left(1 - b^T y\right)^2} y^T Ay \frac{1}{\left(\frac{b_0}{1 - b^T y} b^T y + b_0\right)^p} = \frac{1}{2} y^T Ay \frac{\left(1 - b^T y\right)^{p-2}}{b_0^{p-2}}. \quad (5)$$

When $p = 1$ and $p = 2$, the pseudoconvexity of the function $h(x)$ has been completely characterized in [7, 9]. In this section we aim to study the case $p = 3$, that is

$$h(x) = \frac{1}{2}\frac{x^T Ax}{\left(b^T x + b_0\right)^3}. \quad (6)$$

Setting $p = 3$, $a = -\frac{b}{b_0}$ and $a_0 = -\frac{1}{b_0}$ in (5) we obtain

$$g\left(y\right) = \frac{1}{2} y^T Ay \left(a^T y - a_0\right). \quad (7)$$

In order to study the pseudoconvexity of h, we first deal with the pseudoconvexity of its transformed function g. In this light, the next subsection is devoted to the study of the pseudoconvexity of the product between a quadratic function and a linear one. The obtained results will allow us to characterize the pseudoconvexity of function (6).

Pseudoconvexity of the Product Between a Quadratic Function and a Linear One

Let us consider the following function

$$f(x) = \frac{1}{2}x^T Ax\left(a^T x - a_0\right). \tag{8}$$

Taking into account Lemma 1, we can easily prove that f in (8) is not pseudoconvex on $\Re^n$. More precisely take a vector $d \in \Re^n$ such that $d^T Ad \neq 0$ and $a^T d \neq 0$; the restriction of f along the line $x = td$ is $\varphi(t) = f(td) = \frac{1}{2}\left(t^2 d^T Ad\right)\left(ta^T d - a_0\right)$ and $\varphi'(t) = \frac{1}{2}3t^2\left(a^T d\right)\left(d^T Ad\right) - 2a_0 d^T Adt$. $\varphi(t)$ has two distinct critical points so that it is not pseudoconvex and hence f is not pseudoconvex on $\Re^n$. Due to this, we study the pseudoconvexity of f on the halfspace $S^* = \{x \in \Re^n : a^T x - a_0 > 0\}$.

Preliminary and useful computations are the following

$$\nabla f(x) = Ax\left(a^T x - a_0\right) + \frac{1}{2}x^T Axa \tag{9}$$

$$H(x) = A\left(a^T x - a_0\right) + 2Axa^T$$

$$\varphi(t) = f(x_0 + td) = \frac{1}{2}\left(x_0{}^T Ax_0 + 2x_0^T Adt + t^2 d^T Ad\right)\left(\alpha + ta^T d\right) \tag{10}$$

$$\varphi''(t) = 3t\left(a^T d\right)\left(d^T Ad\right) + 2\left(\alpha d^T Ad + 2\left(a^T d\right)\left(d^T Ax_0\right)\right). \tag{11}$$

Moreover for every $d \in (\nabla f(x))^{\perp}$ we get

$$\varphi'(t) = \frac{3}{2}t^2\left(a^T d\right)\left(d^T Ad\right) + \left(\alpha d^T Ad + 2\left(d^T Ax_0\right)\left(a^T d\right)\right)t \tag{12}$$

where $\alpha = a^T x_0 - a_0$. Before presenting a complete characterization of the pseudoconvexity of f, we state the following necessary conditions.

Theorem 4. *Consider function f in (8). If f is pseudoconvex on $S^* = \{x \in \Re^n : a^T x - a_0 > 0\}$ then*
i) $a_0 \geq 0$.
ii) A is not indefinite.

Proof. i) Suppose $a_0 < 0$. From Lemma 1 there exists $u \in \Re^n$ such that $u^T Au \neq 0$ and $a^T u \neq 0$. Consider the line $x = tu$, $t \in \Re$. It results $\varphi(t) = f(tu) = \frac{1}{2}t^2 u^T Au\left(ta^T u - a_0\right)$, $\varphi'(t) = \frac{1}{2}u^T Au\left(3ta^T u - 2a_0\right)$, $\varphi''(t) = u^T Au\left(3ta^T u - a_0\right)$. $\varphi(t)$ has two distinct critical points $t_1 = 0$ and $t_2 = \frac{2}{3}\frac{a_0}{a^T u}$ with $\varphi''(t_1) = -u^T Aua_0$, $\varphi''(t_2) = u^T Aua_0$. Since t_1 and t_2 are both feasible, $\varphi(t)$ has a feasible maximum point and so it is not pseudoconvex. Consequently f is not pseudoconvex and this is a contradiction.
ii) By contradiction suppose that A is indefinite and take a unit norm eigenvector u associated with a negative eigenvalue λ. We first show that $a^T u \neq 0$; suppose on the contrary that $a^T u = 0$ and take $x = ka + tu$, $k, t \in \Re$. Since $\left(a^T x - a_0\right) = \left(k\|a\|^2 - a_0\right)$, for a sufficiently big k we get $x \in S^*$ for every $t \in \Re$. The restriction of f along the line $x = ka + tu$, $t \in \Re$ is the following

$$\varphi(t) = \left(\frac{1}{2}\lambda t^2 + \frac{1}{2}k^2 a^T Aa\right)\left(k\|a\|^2 - a_0\right).$$

Since $\lambda < 0$, $\varphi(t)$ has a feasible maximum point and so it is not pseudoconvex. Therefore f is not pseudoconvex and this is a contradiction.
Without any loss of generality we can assume $a^T u > 0$. Let v be an eigenvector associated with a positive eigenvalue μ, such that $\|v\| = 1$, $u^T v = 0$. We are going to prove that $a^T v = 0$. Suppose on the contrary that $a^T v \neq 0$; take $k \in \Re$ such that $x_0 = kv \in S^*$, that is $\alpha = ka^T v - a_0 > 0$ and consider $x = x_0 + tu = kv + tu$. Observe that $x \in S^*$ for every $t > -\frac{\alpha}{a^T u}$ and that the restriction of f along the line $x = kv + tu$, $t \in \Re$ is the following

$$\varphi(t) = \left(\frac{1}{2}\lambda t^2 + \frac{1}{2}\mu k^2\right)\left(ta^T u + \alpha\right)$$

so that

$$\varphi'(t) = \frac{3}{2}\lambda a^T u t^2 + \lambda \alpha t + \frac{1}{2}\mu k^2 a^T u \ .$$

Since $\Delta = \alpha^2\lambda^2 - 3\lambda\left(a^T u\right)^2 \mu k^2 > 0$ and $\frac{3}{2}\lambda a^T u < 0$, then $\varphi(t)$ has a feasible maximum point at $t_1 = -\frac{\alpha}{3a^T u} - \frac{\sqrt{\Delta}}{3\lambda a^T u}$ and so φ and f are not pseudoconvex, which is a contradiction. Consequently $a^T v = 0$.
At last consider $x = t\left(u - \sqrt{\frac{|\lambda|}{\mu}}v\right) + kv$, $k, t \in \Re$. It results $x \in S^*$ for $t > \frac{a_0}{a^T u}$ and for every $k \in \Re$; the restriction of f along the line $x = t\left(u - \sqrt{\frac{|\lambda|}{\mu}}v\right) + kv$ is the following

$$\varphi(t) = -ka^T u\mu\sqrt{\frac{|\lambda|}{\mu}}t^2 + \left(\frac{1}{2}k^2\mu a^T u + a_0 k\sqrt{\frac{|\lambda|}{\mu}}\mu\right)t - \frac{1}{2}k^2\mu a_0$$

and hence

$$\varphi'(t) = ka^T u\mu\left(-2\sqrt{\frac{|\lambda|}{\mu}}t + \frac{1}{2}k + \frac{a_0}{a^T u}\sqrt{\frac{|\lambda|}{\mu}}\right).$$

Consequently $\varphi(t)$ as a critical point at $t_1 = \frac{k}{4\sqrt{\frac{|\lambda|}{\mu}}} + \frac{a_0}{2a^T u}$. For $k > \frac{2a_0}{a^T u}\sqrt{\frac{|\lambda|}{\mu}}$, $t_1 \in S^*$ and it is a feasible maximum point for $\varphi(t)$. This implies φ and f are not pseudoconvex, which is a contradiction.

Theorem 5. *Consider function f in (8). If f is pseudoconvex on S^* then*
i) $x^T Ax \geq 0$ for every $x \in a^\perp$.
ii) $v_-(A) \leq 1$.

Proof. i) Suppose there exists $d \in a^\perp$ such that $d^T Ad < 0$. Take $x_0 \in S^*$ and the line $x = x_0 + td$, $t \in \Re$. Observe that $x \in S^*$ for every $t \in \Re$ and since $d^T Ad < 0$ the restriction

$$\varphi(t) = f(x_0 + td) = \frac{1}{2}\left(t^2 d^T Ad + 2d^T Ax_0 t + x_0^T Ax_0\right)\left(a^T x_0 - a_0\right)$$

has a feasible maximum point. Therefore $\varphi(t)$ is not pseudoconvex and this is a contradiction.
ii) Suppose by contradiction that $v_-(A) > 1$ and let u, v be two orthogonal eigenvectors of A associated with two distinct negative eigenvalues λ_1, λ_2. Since $\dim\{u, v\} = 2$ and $\dim a^\perp = n-1$, there exists $d = \alpha u + \beta v$ such that $d \in a^\perp$. Consider $x = x_0 + td$ $t \in \Re$ and the corresponding restriction $\varphi(t)$ of f. By means of simple calculations we get

$$\varphi(t) = \frac{1}{2}(\lambda_1 t^2\alpha^2 \|u\|^2 + \lambda_2 t^2 \beta^2 \|u\|^2 + 2(\alpha u + \beta v)^T A x_0 t + x_0^T A x_0)(a^T x_0 - a_0).$$

Since $\lambda_1, \lambda_2 < 0$, $\varphi(t)$ is not pseudoconvex and this is a contradiction.

The following theorem presents a complete characterization of the pseudoconvexity of f.

Theorem 6. *Consider function f in (8). f is pseudoconvex on S^* if and only if f is of the following form*

$$f(x) = \frac{1}{2}\lambda \left(a^T x\right)^2 \left(a^T x - a_0\right) \text{ where } a_0 \geq 0,\ \lambda \in \Re. \tag{13}$$

Proof. $\Longrightarrow$From Theorem 4, $a_0 \geq 0$ and A can not be indefinite. We are left to deal with the case A is semidefinite. We first assume that A is negative semidefinite. From ii) of Theorem 5, if follows that A has exactly one negative eigenvalues and so A can be rewritten as $A = \mu u u^T$ with $\mu < 0$. From i) of Theorem 5 $d^T A d = 0$ for every $d \in a^\perp$ and so we necessarily have that $u = ka$, i.e., a is an eigenvector of A associated with the negative eigenvalue μ. Therefore $f(x) = \frac{1}{2}\lambda \left(a^T x\right)^2 \left(a^T x - a_0\right)$ where $\lambda = k^2\mu < 0$.
Finally consider the case A positive semidefinite. Let be $x_0 \in S^*$ such that $\nabla f(x_0) \neq 0$. Since A is semidefinite positive, $Ax_0 = 0$ if and only if $x_0^T A x_0 = 0$ and so from (9) it follows that $x_0^T A x_0 \neq 0$. We are going to prove that $a^T d = 0$ for every $d \in (\nabla f(x_0))^\perp$. Suppose on the contrary there exists $d \in (\nabla f(x_0))^\perp$ such that $a^T d \neq 0$. Without any loss of generality we can assume $a^T d > 0$. It results $d^T \nabla f(x_0) = 0$ if and only if $d^T A x_0 \alpha + \frac{1}{2} x_0^T A x_0 d^T a = 0$ where $\alpha = a^T x_0 - a_0$. Consider $x = x_0 + td$ where $t > -\frac{\alpha}{a^T d}$, i.e., $x \in S^*$; the corresponding restriction of f is

$$\varphi(t) = f(x_0 + td) = \frac{1}{2}\left(t^2 d^T A d + 2 d^T A x_0 t + x_0^T A x_0\right)\left(t a^T d + \alpha\right)$$

and from (12)

$$\varphi'(t) = \frac{3}{2} t^2 \left(a^T d\right)\left(d^T A d\right) + \left(\alpha d^T A d + 2\left(d^T A x_0\right)\left(a^T d\right)\right) t\ .$$

Observe that $d^T A d \neq 0$; in fact if $d^T A d = 0$ then $Ad = 0$ and hence $d^T A x_0 = 0$. This can not be true since $d^T \nabla f(x_0) = 0$ and $x_0^T A x_0 \neq 0$.

Since $d^T Ad > 0$, $\varphi'(t) = 0$ for $t_1 = 0$ and $t_2 = -\frac{\alpha d^T Ad + 2\left(d^T Ax_0\right)\left(a^T d\right)}{\frac{3}{2}d^T Ad(a^T d)}$. Obviously $t_1 \in S^*$ and t_2 is feasible if and only if

$$-\frac{\alpha d^T Ad + 2\left(d^T Ax_0\right)\left(a^T d\right)}{\frac{3}{2}d^T Ad\left(a^T d\right)} > -\frac{\alpha}{a^T d}$$

that is

$$\frac{1}{2}\alpha d^T Ad - 2\left(d^T Ax_0\right)\left(a^T d\right) > 0. \tag{14}$$

Since $d^T \nabla f(x_0) = 0$ we have $d^T Ax_0 = -\frac{1}{2\alpha}x_0^T Ax_0 d^T a$ and so condition (14) becomes

$$\frac{1}{2}\alpha d^T Ad + \frac{x_0^T Ax_0}{\alpha}\left(a^T d\right)^2 > 0$$

which is always verified because A is positive semidefinite. Therefore $\varphi(t)$ has a feasible maximum point and this contradicts the pseudoconvexity of f.

Since $a^T d = 0$ for every $d \in \left(\nabla f(x_0)\right)^{\perp}$, $\nabla f(x_0)$ is proportional to a; from (9) it results that for every $x \in S^*$ with $\nabla f(x_0) \neq 0$ we have $Ax_0 = ha$, for some $h \in \Re$. We are going to prove that a is an eigenvector of A associated with a positive eigenvalue λ and that λ is the unique positive eigenvalue of A. Consider $x_0 = k_1 a$ and observe that $x_0 \in S^*$ if and only if $k_1 > \frac{a_0}{\|a\|^2}$. It follows that $Ax_0 = k_1 ha$ and hence $Aa = A\frac{x_0}{k_1} = \frac{h}{k_1}a = \lambda a$. Let $\mu > 0$ be a positive eigenvalue of A, with $\mu \neq \lambda$ and let u be a corresponding eigenvector such that $u \in a^{\perp}$. Take $x_0 = k_1 a + u$ with $k_1 > \frac{a_0}{\|a\|^2}$, i.e. $x_0 \in S^*$. It is easy to verify that $\nabla f(x_0) \neq 0$, so that there exists $\bar{h}$ such that $Ax_0 = \bar{h}a$. On the other hand, $Ax_0 = \lambda k_1 a + \mu u$ and therefore $\bar{h}a = \lambda k_1 a + \mu u$ that is $\left(\bar{h} - \lambda k_1\right)a = \mu u$ which contradicts $u \in a^{\perp}$. Consequently a is an eigenvector associated with the unique positive eigenvalue λ and hence f is of the form in (13).

$\Longleftarrow$ It results $\nabla f(x) = \lambda(\frac{3}{2}(a^T x)^2 - (a^T x)a_0)a$, $H(x) = \lambda(3a^T x - a_0)aa^T$. Since $a_0 \geq 0$, then the critical points of f do not belong to S^* and so it remains to prove that for every $d \in (\nabla f(x))^{\perp}$ we get $d^T H(x)d \geq 0$. Since $d \in (\nabla f(x))^{\perp}$ if and only if $d \in a^{\perp}$ we get $d^T H(x)d = \lambda\left(3a^T x - a_0\right)d^T aa^T d = 0$ and the proof is complete.

Remark 1. It is worth noticing that f in (13) is also pseudoconcave on S^*; in fact the critical points do not belong to S^* and for every $d \in (\nabla f(x))^{\perp}$ we get $d^T H(x)d = 0$. Therefore, the second order characterization for the pseudoconcave function is verified and so f is pseudolinear.

Recalling that function f is the Charnes-Cooper transformed function of h, by means of the previous result we can characterize the pseudoconvexity of function h.

Theorem 7. *Consider function* $h(x) = \frac{1}{2}\frac{x^T Ax}{(b^T x+b_0)^3}$. *h is pseudoconvex on* $S = \{x \in \Re^n : b^T x + b_0 > 0\}$ *if and only if h is of the following form*

$$h(x) = \frac{1}{2}\frac{\mu\left(b^T x\right)^2}{\left(b^T x + b_0\right)^3} \text{ where } b_0 < 0,\ \mu \in \Re.$$

Theorem 8. *Consider the function* $h(x) = \frac{1}{2}\frac{\mu\left(b^T x\right)^2}{(b^T x+b_0)^3}$ *where* $b_0 < 0,\ x \in S = \{x \in \Re^n : b^T x + b_0 > 0\}$. *Then* $h(x)$ *is convex if* $\mu > 0$ *and it is concave if* $\mu < 0$.

Proof. Consider the function $\rho(z) = \frac{z}{z+b_0} = 1 + \left(\frac{-b_0}{z+b_0}\right)$ defined on $z + b_0 > 0$. Since $b_0 < 0$, ρ is convex and hence the function $\rho(b^T x)$ is convex on S. Moreover it results $\frac{1}{\mu}h(x) = \frac{(\rho(b^T x))^2}{b^T x+b_0}$ which is the ratio of a squared convex function and an affine positive one; such a kind of function is convex (see for instance [1]), so that the function h is convex if $\mu > 0$ and concave if $\mu < 0$.

4 Concluding Remarks

In this paper we have characterized the pseudoconvexity of two new classes of generalized fractional functions using the Charnes-Cooper transformation. The problem of characterizing pseudoconvex functions is not yet sufficently studied in the literature because of its difficulty. We hope that the given approach, applied also in [6, 7, 9], provides further developments in this direction.

References

1. Avriel M, Diewert WE, Schaible S, Zang I (1988) Generalized Concavity. Plenum Press, New York
2. Bajalinov EB (2004) Linear-Fractional Programming: Theory, Methods, Applications, and Software. Kluwer Academic Publishers, Boston MA
3. Barros AI (1998) Discrete and Fractional Programming Techniques for Location Models. Kluwer Academic Publishers, Dordrecht
4. Cambini A, Crouzeix JP, Martein L (2002) On the pseudoconvexity of a quadratic fractional function. Optimization 51:667-687
5. Cambini A, Martein L (1986) A modified version of Martos' algorithm. Methods of Operation Research 53:33-44
6. Cambini A, Martein L, Schaible S. (2005.a) On the pseudoconvexity of the sum of two linear fractional functions. In: Eberherd A, Hadjisavvas N, Luc DT (eds) Generalized convexity, generalized monotonicity and applications:161-172, Springer, New York NY
7. Cambini A, Martein L, Schaible S. (2005.b) Pseudoconvexity under the Charnes-Cooper transformation. Pacific Journal of Optimization 2:1-11

8. Cambini R, Carosi L (2002) On generalized convexity of quadratic fractional functions. Rendiconti del Circolo Matematico di Palermo Series II Suppl.70:155-176
9. Carosi L, Martein L (2004) On the pseudoconvexity and pseudolinearity of some classes of fractional functions. Report n. 247, Department of Statistics and Applied Mathematics, University of Pisa (forthcoming in Optimization)
10. Charnes A, Cooper WW (1962) Programming with linear fractional functionals. Naval Research Logistics Quarterly 9:181-186
11. Charnes A, Cooper WW, Rhodes E (1978) Measuring the efficiency of decision making units. European Journal of Operational Research 2:429-444.
12. Charnes A, Cooper WW, Lewin AY, Seiford LM (eds) (1995) Data Envelopment Analysis: Theory, Methodology and Applications. Kluwer Academic Publishers, Boston MA
13. Crouzeix JP (1998) Characterizations of Generalized Convexity and Monotonicity, a survey. In: Crouzeix JP, Martinez-Legaz JE, Volle M (eds) Generalized Convexity, Generalized Monotonicity:237-256. Kluwer Academic Publishers, London
14. Frenk JBG, Schaible S (2005) Fractional programming. In: Hadjisavvas N, Komlòsi S, Schaible S (eds) Handbook of generalized convexity and generalized monotonicity:335-386. Springer, New York
15. Mjelde KM (1978) Allocation of resources according to a fractional objective. European Journal of Operational Research 2:116-124
16. Schaible S (1995) Fractional Programming. In Horst R, PM Pardalos (eds) Handbook of Global Optimization:495-608. Kluwer Academic Publishers, Dordrecht

Equilibrium Problems Via the Palais-Smale Condition

Ouayl Chadli[1], Zaki Chbani[2], and Hassan Riahi[2]

[1] Department of Economics, Ibn Zohr University, B.P. 8658, Agadir 80000, Morocco. ochadli@esta.ac.ma
[2] Department of Mathematics, Cadi Ayyad University, Marrakesh 40000, Morocco.

Summary. Inspired by some results from nonsmooth critical point theory, we propose in this paper to study equilibrium problems by means of a general Palais-Smale condition adapted to bifunctions. We introduce the notion of critical points for equilibrium problems and we give some existence results for (EP) with lack of compacity.

Key words: Equilibrium problems, critical points, Palais-Smale condition, Ekeland variational principle.

1 Introduction and Motivation

Let X be a real normed space, $K \subset X$ be a nonempty convex set, $D \subset X$ is an open set containing K and $\Phi : K \times D \longrightarrow I\!R$ be a given function satisfying $\Phi(x, x) = 0$ for all $x \in K$. We consider the following Equilibrium Problem

$$\text{(EP)} \qquad \text{Find } \overline{x} \in K \text{ such that } \Phi(\overline{x}, y) \geq 0 \quad \forall y \in K$$

From its formulation, equilibrium problems theory has emerged as an interesting branch of applicable mathematics. This theory provides a general and convenient format to write and investigate many problems and becomes a rich source of inspiration and motivation for the study of a large number of problems arising in economics, optimization, and operation research in a general and unified way , see [1, 2]. There are a substantial number of papers on existence results for solving equilibrium problems based on different relaxed monotonicity notions and various coercivity assumptions, see [3, 4, 5, 6]. In a noncoercive framework, equilibrium problems have been studied by using arguments from the recession analysis, see [7, 8, 9, 10].

In this paper, we study the existence of solutions for equilibrium problems by using an approach inspired from the nonsmooth critical point theory and an adapted Palais-Smale condition for bifunctions. Our approach is of two types, the first one is by using some kind of monotonicity assumption and a recent result by Bianchi-Kassay-Pini [11] on an Ekeland variational principle for equilibrium problems. The second approach is without any monotonicity assumption and which will be developed in the last section of this paper. First of all, we need to introduce the concept of critical points for equilibrium problems and its connection with the solutions of equilibrium problems. In the following, we give a motivation for introducing this notion for (EP).

Motivation

If we assume that the function $y \mapsto \Phi(x,y)$ is convex for all $x \in K$ and $\overline{x}$ is a solution of (EP), then

$$\frac{\Phi(\overline{x},\overline{x}+t(y-\overline{x}))-\Phi(\overline{x},\overline{x})}{t} \geq 0 \quad \forall t \in (0,1],\ y \in K$$

Taking the limit when $t \to 0^+$, we get

$$\Phi^0(\overline{x},\overline{x})(h) \geq 0 \quad \forall h \in K-\overline{x},$$

where Φ^0 denotes the directional derivative of Φ with respect to the second variable.
On the other hand, suppose $\overline{x},\overline{y} \in K$ are such that $\Phi(\overline{x},\overline{y}) \geq 0$, Φ is quasi-convex with respect to the second argument and

$$\Phi^0(\overline{x},\overline{y})(h) > 0 \quad \forall h \in K-\overline{y} \quad \text{with } h \neq 0. \tag{1}$$

Consider $t \in (0,1]$ and $y \in K$ with $y \neq \overline{y}$. Since $\Phi(\overline{x},\cdot)$ is quasiconvex, then

$$\Phi(\overline{x},\overline{y}+t(y-\overline{y})) \leq \max\{\Phi(\overline{x},y),\Phi(\overline{x},\overline{y})\}$$

Therefore

$$\frac{\Phi(\overline{x},\overline{y}+t(y-\overline{y}))-\Phi(\overline{x},\overline{y})}{t} \leq \frac{1}{t}\max\{\Phi(\overline{x},y)-\Phi(\overline{x},\overline{y}),0\}.$$

Since $\lim\limits_{t \longrightarrow 0^+} \dfrac{\Phi(\overline{x},\overline{y}+t(y-\overline{y}))-\Phi(\overline{x},\overline{y})}{t} = \Phi^0(\overline{x},\overline{y})(y-\overline{y})$, then for each $\varepsilon > 0$, with $2\varepsilon < \Phi^0(\overline{x},\overline{y})(y-\overline{y})$, there exists $\eta > 0$ such that for $0 < t < \eta$

$$\frac{\Phi(\overline{x},\overline{y}+t(y-\overline{y}))-\Phi(\overline{x},\overline{y})}{t} > \Phi^0(\overline{x},\overline{y})(y-\overline{y})-\varepsilon.$$

It follows that $\dfrac{1}{t}\max\{\Phi(\overline{x},y)-\Phi(\overline{x},\overline{y}),0\} > \varepsilon$ for $0 < t < \eta$. Hence

$$\max\{\Phi(\overline{x},y)-\Phi(\overline{x},\overline{y}),0\} > t[\Phi^0(\overline{x},\overline{y})(y-\overline{y})-\varepsilon] > 0$$

Therefore $\Phi(\overline{x},y)-\Phi(\overline{x},\overline{y}) > t[\Phi^0(\overline{x},\overline{y})(y-\overline{y})-\varepsilon] > 0$. We conclude $\Phi(\overline{x},y) > \Phi(\overline{x},\overline{y}) \geq 0 \ \ \forall y \in K$ with $y \neq \overline{y}$. Since $\Phi(\overline{x},\overline{x}) = 0$, it follows that $\overline{y} = \overline{x}$ and $\Phi(\overline{x},y) \geq 0$ for all $y \in K$. Hence $\overline{x}$ is a solution of (EP).
Note that if we assume Φ is convex with respect to the second argument, then the strict inequality in relation (1) can be replaced by a large one.

Motivated by the arguments above, we shall introduce in the next section the notions of critical points and strict critical points for equilibrium problems.

2 Notion of Critical Points for Equilibrium Problems

Definition 1. *The bifunction Φ is said to be locally Lipschitz with respect to the second variable if $\forall y \in K$ there exists $L_y > 0$ and a neighborhood $U_y \subset D$ of y such that*

$$|\Phi(x, y') - \Phi(x, y'')| \leq L_y \|y' - y''\|$$

for all $y', y'' \in U_y$ and $x \in K \cap U_y$.

The family of all bifunctions $\Phi : K \times D \longrightarrow I\!R$ with the above property will be denoted by $\mathcal{L}ip_{loc}(K)$.

Remark 1. Let Φ a bifunction defined by $\Phi(x, y) = \langle T(x), y - x \rangle$, where T is a nonlinear operator. Then Definition 1 is satisfied when T is locally bounded.

Definition 2. *Let $\Phi \in \mathcal{L}ip_{loc}(K)$. For $x \in K$ and $h \in X$, the generalized Clarke-type derivative of Φ at the point (x, x) with respect to the second variable in the direction h is defined by*

$$\Phi^0(x, x)(h) = \limsup_{\substack{t \longrightarrow 0^+ \\ (u,v) \longrightarrow (x,x) \\ v \in K}} \frac{\Phi(u, v + th) - \Phi(u, v)}{t}$$

In the next lemma, some properties of Φ^0 that we will need in the sequel are presented.

Lemma 1. *Let $\Phi \in \mathcal{L}ip_{loc}(K)$. Then*

(i) For each $x \in K$, the function $h \mapsto \Phi^0(x, x)(h)$is sublinear Lipschitz continuous on X;
(ii) The function $(x, x, h) \mapsto \Phi^0(x, x)(h)$ is upper semicontinuous on $K \times K \times X$.

Proof. Let $x \in K$ fixed. One can verify easily that the function $h \mapsto \Phi^0(x, x)(h)$ is positively homogeneous. We need only to show the subadditivity of the function $h \mapsto \Phi^0(x, x)(h)$. To this aim, let $h, w \in X$ then

$$\begin{aligned}
\Phi^0(x,x)(h+w) &= \limsup_{\substack{t \longrightarrow 0^+ \\ (u,v) \longrightarrow (x,x) \\ v \in K}} \frac{\Phi(u, v+th+tw) - \Phi(u,v)}{t} \\
&\leq \limsup_{\substack{t \longrightarrow 0^+ \\ (u,v) \longrightarrow (x,x) \\ v \in K}} \frac{\Phi(u, v+th+tw) - \Phi(u,v+tw)}{t} \\
&\quad + \limsup_{\substack{t \longrightarrow 0^+ \\ (u,v) \longrightarrow (x,x) \\ v \in K}} \frac{\Phi(u, v+tw) - \Phi(u,v)}{t} \\
&= \Phi^0(x,x)(h) + \Phi^0(x,x)(w)
\end{aligned}$$

On the other hand, from the local Lipshitz property of Φ, one has

$$\Phi^0(x,x)(h) \leq L_y \|h\| \quad \forall h \in X.$$

Taking account of the subadditivity property, one has

$$|\Phi^0(x,x)(h) - \Phi^0(x,x)(w)| \leq L_y \|h - w\| \quad \forall h, w \in X.$$

Which completes the proof of (i).
Now to prove (ii), let $x \in K$, $h \in X$ and consider $\{x_n\}$, $\{h_n\}$ sequences in K, respectively in X, such that $x_n \to x$ and $h_n \to h$. From the definition of Φ^0, there exist sequences $t_n > 0$, $\tilde{x}_n \in K$, $\tilde{y}_n \in K$ such that

$$t_n < \frac{1}{n}, \quad \|\tilde{x}_n - x_n\| < \frac{1}{n}, \quad \|\tilde{y}_n - x_n\| < \frac{1}{n}$$

and $\Phi^0(x_n, x_n)(h_n) - \frac{1}{n} \leq \dfrac{\Phi(\tilde{x}_n, \tilde{y}_n + t_n h_n) - \Phi(\tilde{x}_n, \tilde{y}_n)}{t_n}$.
For $n \in I\!N$ sufficiently large, one has $\tilde{y}_n + t_n h_n \in U_y$. Hence

$$|\Phi(\tilde{x}_n, \tilde{y}_n + t_n h_n) - \Phi(\tilde{x}_n, \tilde{y}_n + t_n h)| \leq L_y \|h_n - h\|.$$

It follows

$$\Phi^0(x_n, y_n)(h_n) - \frac{1}{n} - L_y \|h_n - h\| \leq \frac{\Phi(\tilde{x}_n, \tilde{y}_n + t_n h) - \Phi(\tilde{x}_n, \tilde{y}_n)}{t_n}.$$

By passing to the limsup when $n \to +\infty$ in the above inequality, one obtain

$$\limsup_{n \to +\infty} \Phi^0(x_n, x_n)(h_n) \leq \limsup_{n \to +\infty} \frac{\Phi(\tilde{x}_n, \tilde{y}_n + t_n h) - \Phi(\tilde{x}_n, \tilde{y}_n)}{t_n} \leq \Phi^0(x,x)(h).$$

Which completes the proof of (ii). □

Now we introduce the notion of critical points for equilibrium problems.

Definition 3. *$\overline{x} \in K$ is said to be a critical point (resp. strict critical point) of Φ, if*

$$\Phi^0(\overline{x}, \overline{x})(h) \geq 0 \quad \forall h \in K - \overline{x}$$

$$(\text{ resp. } \Phi^0(\overline{x}, \overline{x})(h) > 0 \quad \forall h \in K - \overline{x} \quad \text{with} \quad h \neq 0).$$

Proposition 1. *Let $\Phi \in \mathcal{L}ip_{loc}(K)$ and $\overline{x}$ a strict critical point of Φ. Assume that*

(i) $\forall x \in K$ fixed, $y \mapsto \Phi(x, y)$ is quasiconvex in K;
(ii) $\forall y \in K$ fixed, $x \mapsto \Phi(x, y)$ is continuous on K.

Then $\overline{x}$ is a solution of (EP).

Proof. Since $\overline{x}$ is a strict critical point of Φ, then

$$\Phi^0(\overline{x}, \overline{x})(h) > 0 \quad \forall h \in K - \overline{x} \quad \text{with} \quad h \neq 0.$$

Let $\{t_n\}$, $\{x_n\}$ and $\{y_n\}$ be sequences such that $t_n \searrow 0^+$, $x_n \to \overline{x}$, $y_n \to \overline{x}$ and

$$\Phi^0(\overline{x}, \overline{x})(h) = \lim_{n \longrightarrow +\infty} \frac{\Phi(x_n, y_n + t_n h) - \Phi(x_n, y_n)}{t_n} > 0$$

Since $y \mapsto \Phi(x, y)$ is quasiconvex, then

$$\Phi(x_n, y_n + t_n h) \leq \max\{\Phi(x_n, y_n + h), \Phi(x_n, y_n)\}$$

Hence

$$\frac{\Phi(x_n, y_n + t_n h) - \Phi(x_n, y_n)}{t_n} \leq \frac{1}{t_n} \max\{\Phi(x_n, y_n + h) - \Phi(x_n, y_n), 0\}. \quad (2)$$

On the other hand, let $\varepsilon > 0$ with $\varepsilon < \Phi^0(\overline{x}, \overline{x})(h)$, then there exists $N \in \mathbb{N}$ such that $\forall n > N$ one has

$$\frac{\Phi(x_n, y_n + t_n h) - \Phi(x_n, y_n)}{t_n} > \Phi^0(\overline{x}, \overline{x})(h) - \varepsilon$$

Therefore by taking account of (2), one deduces

$$\max\{\Phi(x_n, y_n + h) - \Phi(x_n, y_n), 0\} > t_n(\Phi^0(\overline{x}, \overline{x})(h) - \varepsilon) > 0$$

Hence $\Phi(x_n, y_n + h) - \Phi(x_n, y_n) > 0$ for all $n > N$. Therefore

$$\limsup_{n \longrightarrow +\infty} \Phi(x_n, y_n + h) \geq \limsup_{n \longrightarrow +\infty} \Phi(x_n, y_n).$$

Taking account of (ii), one deduces

$$\Phi(\overline{x}, \overline{x} + h) \geq \Phi(\overline{x}, \overline{x}) = 0.$$

Hence, $\overline{x}$ is a solution of (EP). $\square$

Proposition 2. *Let $\Phi \in \mathcal{L}ip_{loc}(K)$ and $\overline{x}$ a critical point of Φ. Assume that*

(i) $\forall x \in K$ fixed, $y \mapsto \Phi(x,y)$ is convex in K;
(ii) $\forall y \in K$ fixed, $x \mapsto \Phi(x,y)$ is continuous on K.

Then $\overline{x}$ is a solution of (EP).

Proof. Since $\overline{x}$ is a critical point of Φ, then $\Phi^0(\overline{x},\overline{x})(h) \geq 0 \ \ \forall h \in K - \overline{x}$. On the other hand since $y \mapsto \Phi(x,y)$ is convex, then

$$\Phi(u, v+th) \leq t\Phi(u, v+h) + (1-t)\Phi(u,v).$$

Hence

$$\frac{\Phi(u,v+th) - \Phi(u,v)}{t} \leq \Phi(u,v+h) - \Phi(u,v).$$

Consequently

$$\limsup_{(u,v)\longrightarrow(\overline{x},\overline{y})} [\Phi(u,v+h) - \Phi(u,v)] \geq \Phi^0(\overline{x},\overline{x})(h) \geq 0.$$

Taking account of (ii), one deduces

$$\Phi(\overline{x}, \overline{x}+h) \geq \Phi(\overline{x},\overline{x}) = 0 \ \ \forall h \in K - \overline{x}. \quad \square$$

Definition 4. *[13] Let $f : K \longrightarrow I\!R$ be a locally Lipschitz function. Then f is called pseudoconvex, if for every $x, y \in K$ the following implication holds*

$$f^0(x)(y-x) \geq 0 \quad \Longrightarrow \quad \forall z \in [x,y] \quad f(z) \leq f(y).$$

Inspired by the definition above, we shall introduce in the next definition the notion of pseudoconvexity to the class of bifunctions $\Phi \in \mathcal{L}ip_{loc}(K)$.

Definition 5. *Let $\Phi \in \mathcal{L}ip_{loc}(K)$, Φ is said to be pseudoconvex if $\forall x, y \in K$ the following implication holds*

$$\Phi^0(x,x)(y-x) \geq 0 \quad \Longrightarrow \quad \forall t \in [0,1] \quad \Phi(x, tx + (1-t)y) \leq \Phi(x,y).$$

Remark 2. If $\Phi(x,y) = f(y) - f(x)$ where $f : K \longrightarrow I\!R$ is a locally Lipschitz function, then one can easily verify that f is pseudoconvex is equivalent to Φ is pseudoconvex.

Proposition 3. *Let $\Phi \in \mathcal{L}ip_{loc}(K)$ and $\overline{x}$ a critical point of Φ. Assume that Φ is pseudonconvex, then $\overline{x}$ is a solution of (EP).*

Proof. Since $\overline{x}$ a critical point of Φ, then $\Phi^0(\overline{x},\overline{x})(y-\overline{x}) \geq 0$ for all $y \in K$. On the other hand, the bifunction Φ is pseudonconvex, then for $y \in K$ one has for all $t \in [0,1]$,

$$\Phi(\overline{x}, ty + (1-t)\overline{x}) \leq \Phi(\overline{x}, y). \tag{3}$$

Set $t = 0$ in relation (3), one obtain

$$0 = \Phi(\overline{x},\overline{x}) \leq \Phi(\overline{x}, y).$$

Hence $\overline{x}$ is a solution of (EP). $\square$

Definition 6. *[14] Let $f : X \longrightarrow I\!R$ be a locally Lipschitz functional. We say that f satisfies the Palais-Smale condition if for every sequence $\{u_n\} \subset X$ such that $\{f(u_n)\}$ is bounded and*

$$f^0(u_n)(v - u_n) \geq -\varepsilon_n \|v - u_n\|, \text{ for all } v \in X,$$

for a sequence $\{\varepsilon_n\} \subset I\!R^+$ with $\lim_{n \longrightarrow +\infty} \varepsilon_n = 0$, $\{u_n\}$ contains a convergent subsequence, where f^0 is the Clarke derivative.

An extension of the Palais-Smale condition to the case of bifunctions can be given by the following definition

Definition 7. *Let $\Phi \in \mathcal{L}ip_{loc}(K)$, we say that Φ satisfies the Palais-Smale condition if for every sequence $\{u_n\} \subset K$ such that the sequence $\{\inf_{v \in K} \Phi(u_n, v)\}$ is bounded from below and*

$$\Phi^0(u_n, u_n)(h) \geq -\varepsilon_n \|h\|, \text{ for all } h \in K - u_n,$$

for a sequence $\{\varepsilon_n\} \subset I\!R^+$ with $\lim_{n \longrightarrow +\infty} \varepsilon_n = 0$, $\{u_n\}$ contains a convergent subsequence.

Recall the following coercivity definition

Definition 8. *[1] A bifunction $\Phi : K \times K \longrightarrow I\!R$ is said to be coercive if there exists $a \in K$ such that $\Phi(u, a) \longrightarrow -\infty$ when $\|u - a\| \longrightarrow +\infty$.*

One has the following property

Proposition 4. *Let $\Phi \in \mathcal{L}ip_{loc}(K)$, if Φ is coercive then it satisfies the Palais-Smale condition.*

Proof. Suppose the contrary, then there exists a sequence $\{u_n\} \subset K$ such that $\{\inf_{v \in K} \Phi(u_n, v)\}$ is bounded from below and

$$\Phi^0(u_n, u_n)(h) \geq -\varepsilon_n \|h\|, \text{ for all } h \in K - u_n,$$

for $\{\varepsilon_n\} \subset I\!R^+$ with $\lim_{n \longrightarrow +\infty} \varepsilon_n = 0$, and $\{u_n\}$ has no convergent subsequence. Since Φ is coercive, then $\Phi(u_n, a) \longrightarrow -\infty$ when $n \longrightarrow +\infty$. On the other hand, since

$$\inf_{v \in K} \Phi(u_n, v) \leq \Phi(u_n, a)$$

then $\inf_{v \in K} \Phi(u_n, v) \longrightarrow -\infty$ when $n \longrightarrow +\infty$. Which contradicts the fact that $\{\inf_{v \in K} \Phi(u_n, v)\}$ is bounded from below. □

3 Existence Results for (EP) - The Monotone Case

In this section we shall study the existence of critical points and solution points for equilibrium problems. Our approach will be based on a recent result by Bianchi-Kassay-Pini [11] on an extension of Ekeland's variational principle to the setting of equilibrium problems.

Lemma 2 (Ekeland). *Assume that f is a proper lower semicontinous function on a Banach space X. Suppose that $\varepsilon > 0$ and that $f(x_0) < \inf_{x \in X} f(x) + \varepsilon$. Then for any λ with $0 < \lambda < 1$ there exists $z \in dom(f)$ such that*

(i) $\lambda \|z - x_0\| \le f(x_0) - f(z)$;
(ii) $\|z - x_0\| < \varepsilon/\lambda$;
(iii) $f(z) < f(x) + \lambda \|x - z\|$ *whenever* $x \neq z$.

Lemma 3. *[11] Assume that $K \subset X$ be a closed set, where X is a normed space and $\Phi : K \times K \to I\!R$. Suppose that Φ satisfies the following assumptions*

(i) for $x \in K$ fixed, the function $\Phi(x, \cdot)$ is lower bounded and lower semicontinuous;
(ii) for every $x \in K$, $\Phi(x, x) = 0$;
(iii) for every $x, y, z \in K$, $\Phi(x, y) \le \Phi(x, z) + \Phi(z, y)$.

Then, for every $\varepsilon > 0$ and for every $x_0 \in K$, there exists $\overline{x} \in K$ such that

$$\begin{array}{ll} (a) & \Phi(x_0, \overline{x}) + \varepsilon \|x_0 - \overline{x}\| \le 0, \\ (b) & \Phi(\overline{x}, x) + \varepsilon \|\overline{x} - x\| > 0, \quad \forall x \in K, \ x \neq \overline{x}. \end{array}$$

The main result of this section is the following

Theorem 1. *Let $\Phi \in \mathcal{L}ip_{loc}(K)$, where K is a closed convex subset of a normed space X. Suppose that Φ satisfies the following assumptions*

(i) for $x \in K$ fixed, the function $\Phi(x, \cdot)$ is lower bounded;
(ii) for all $x \in K$, $\Phi(x, x) = 0$;
(iii) there exists $x_0 \in K$ such that $\Phi(x_0, \cdot)$ is bounded above;
(iv) for every $x, y, z \in K$, $\Phi(x, y) \le \Phi(x, z) + \Phi(z, y)$;
(v) Φ satisfies the Palais-Smale condition.

Then, Φ has a critical point $\overline{x}$. Furthermore, if Φ is assumed to be pseudoconvex then $\overline{x}$ is a solution of (EP).

Proof. Let $\{\varepsilon_n\}$ be a sequence of positive numbers such that $\varepsilon_n \to 0^+$. From Lemma 3, one has for $\varepsilon_n > 0$ there exists $x_{\varepsilon_n} \in K$ such that

$$\Phi(x_{\varepsilon_n}, x) + \varepsilon_n \|x_{\varepsilon_n} - x\| > 0, \quad \forall x \in K, \ x \neq x_{\varepsilon_n}.$$

Hence for $x = x_{\varepsilon_n} + th$ with $h \in K - x_{\varepsilon_n}$ and $t > 0$, one has

$$\Phi(x_{\varepsilon_n}, x_{\varepsilon_n} + th) - \Phi(x_{\varepsilon_n}, x_{\varepsilon_n}) > -\varepsilon_n t \|h\|$$

Therefore

$$\frac{\Phi(x_{\varepsilon_n}, x_{\varepsilon_n} + th) - \Phi(x_{\varepsilon_n}, x_{\varepsilon_n})}{t} > -\varepsilon_n \|h\|.$$

By passing to the limit in the above relation when t goes to 0^+, one obtain

$$\Phi^0(x_{\varepsilon_n}, x_{\varepsilon_n})(h) \geq -\varepsilon_n \|h\| \quad \forall h \in K - x_{\varepsilon_n}. \tag{4}$$

On the other hand, from assumptions (iii) there exists $\alpha \in I\!\!R$ such that

$$\Phi(x_0, x_{\varepsilon_n}) \leq \alpha \quad \forall n \in I\!\!N.$$

Let $x \in K$, then from (iv) one has

$$\begin{aligned} \Phi(x_0, x) &\leq \Phi(x_0, x_{\varepsilon_n}) + \Phi(x_{\varepsilon_n}, x) \\ &\leq \alpha + \Phi(x_{\varepsilon_n}, x) \end{aligned}$$

Hence, $\Phi(x_{\varepsilon_n}, x) \geq \Phi(x_0, x) - \alpha$ for all $x \in K$. Taking account of (i), one deduces that $\{\inf_{x \in K} \Phi(x_{\varepsilon_n}, x)\}$ is bounded from below.
Then from the Palais-Smale condition, one deduces that $x_{\varepsilon_n} \to \overline{x} \in K$. By passing to the limit in relation (4) and from Lemma 1, one deduces

$$\Phi^0(\overline{x}, \overline{x})(h) \geq 0 \quad \forall h \in K - \overline{x}$$

and since $\Phi(\overline{x}, \overline{x}) = 0$, one conclude that $\overline{x}$ is a critical point of Φ. □

Remark 3.

1- Condition (iii) in Lemma 3, as mentioned in [11], implies the cyclic monotonicity of $-\Phi$, i.e. for every $x_1, \cdots, x_n \in K$ we have $\sum_{i=1}^{n} \Phi(x_i, x_{i+1}) \geq 0$. Therefore, Theorem 1 can be seen as an existence result for critical points for equilibrium problems in a monotone framework.

2- Condition (iii) in Lemma 3 and Theorem 1 has been initially introduced by Blum-Oettli [1, Theorem 3] for studying equilibrium problems in complete metric spaces.

3- If the bifunction Φ is of the form $\Phi(x, y) = \varphi(y) - \varphi(x)$ then condition (iii) in Lemma 3 and Theorem 1 is easily satisfied. We point out that there is other types of bifunctions which are not of the previous form and satisfy condition (iii). For example (see [11]), take Φ defined as the following

$$\Phi(x, y) = \begin{cases} e^{-\|x-y\|} + \varphi(y) - \varphi(x) & \text{if } \ x \neq y \\ 0 \quad \text{if } \ x = y \end{cases}$$

then Φ is lower bounded and lower semicontinuous.

4- Let Φ defined by $\Phi(x, y) = \langle T(x), y - x \rangle$ where $T : K \to X^*$ is an operator, X^* is the topological dual of X. Then condition (iv) can be written

$$\forall x, y, z \in K, \ \langle T(z) - T(x), z - y \rangle \leq 0.$$

Note that if $\Phi(x, y) = \langle T(x), y - x \rangle$ satisfies condition (iv), then it satisfies condition $(*)$ in Brézis-Haraux [12] with $A = -T$.

4 Existence Results for (EP) - The Nonmonotone Case

In this section we shall investigate under which conditions we can have the existence of a solution point for equilibrium problems when K is not necessarily compact and without a monotonicity assumption. In our approach, we follow the method used in [15] which was devoted to establish some local minimax theorems without compactness for functionals which are of class C^1.

In the following, we shall suppose that the bifunction Φ is locally Lipshitz with respect to the second argument. For $x, y \in K$, the Generalized Clark-type derivative of Φ at the point (x, y) with respect to the second argument in the direction h is defined by

$$\Phi^0(x,y)(h) = \limsup_{\substack{t \longrightarrow 0^+ \\ (u,v) \longrightarrow (x,y) \\ v \in K}} \frac{\Phi(u, v+th) - \Phi(u,v)}{t}$$

In this context, the definition of the Palais-Smale condition will be as the following

Definition 9. *We say that Φ satisfies the Generalized Palais-Smale condition if for every sequence $\{(u_n, v_n)\} \subset K \times K$ such that $\Phi(u_n, v_n) = \sup_{u \in K} \Phi(u, v_n)$, the sequence $\{\Phi(u_n, v_n)\}$ is bounded and*

$$\Phi^0(u_n, v_n)(h) \geq -\varepsilon_n \|h\|, \text{ for all } h \in K - v_n,$$

for a sequence $\{\varepsilon_n\} \subset I\!R^+$ with $\lim_{n \longrightarrow +\infty} \varepsilon_n = 0$, then $\{(u_n, v_n)\}$ contains a convergent subsequence.

We will need some preliminary results which are listed below.

Lemma 4. *[1] Let E be a convex, compact set, and F be a convex set. Let $p : E \times F \longrightarrow I\!R$ be quasiconcave and upper semicontinuous in the first argument, and convex in the second argument. Assume that*

$$\max_{\xi \in E} p(\xi, y) \geq 0 \quad \forall y \in F.$$

Then there exists $\overline{\xi} \in E$ such that $p(\overline{\xi}, y) \geq 0$ for all $y \in F$.

We shall need the following two lemmas

Lemma 5. *Let K be a closed convex set of a Banach space X and let $f : K \longrightarrow I\!R$ be a locally Lipschitz function bounded from below and satisfies the Palais-Smale condition. Then the set*

$$\mathcal{S} = \{\overline{x} \in K : f(\overline{x}) = \min_{x \in K} f(x)\}$$

is nonempty and compact. Moreover, if f is quasiconvex then $\mathcal{S}$ is convex. If $\mathcal{N}$ is an open set containing $\mathcal{S}$ and $\partial\mathcal{N}$ its boundary, then

$$\inf_{x\in K\cap\partial\mathcal{N}} f(x) > \inf_{x\in K} f(x).$$

Proof. From the Ekeland variational principle, one has for $\lambda_n > 0$ with $\lambda_n \longrightarrow 0^+$, there exists $x_n \in K$ such that

$$\lambda_n\|x - x_n\| + f(x) > f(x_n) \qquad \forall x \neq x_n. \tag{5}$$

For $h \in X$ and $t > 0$, set $x = x_n + th$. Then from (5), we have

$$\frac{f(x_n + th) - f(x_n)}{t} > -\lambda_n\|h\|.$$

Hence

$$f^0(x_n)(h) \geq \limsup_{t\longrightarrow 0^+} \frac{f(x_n + th) - f(x_n)}{t} > -\lambda_n\|h\|.$$

Taking account of the Palais-Smale condition, one deduces that $\{x_n\}$ has a convergent subsequence $x_{n_k} \longrightarrow \overline{x}$. From (5), we have $\overline{x} \in \mathcal{S}$.
Now let $\{\overline{x}_n\}$ be a sequence in $\mathcal{S}$. For $\delta_n \searrow 0^+$ and $x \neq \overline{x}_n$, one has

$$f(\overline{x}_n) < f(x) + \delta_n\|x - \overline{x}_n\| \qquad \forall x \neq \overline{x}_n$$

It follows

$$f^0(\overline{x}_n)(h) \geq \limsup_{t\longrightarrow 0^+} \frac{f(\overline{x}_n + th) - f(\overline{x}_n)}{t} > -\lambda_n\|h\|.$$

Hence from the Palais-Smale condition, $\{\overline{x}_n\}$ has a convergent subsequence $\overline{x}_{n_k} \longrightarrow \overline{x}$ and one can easily see that $\overline{x} \in \mathcal{S}$.
To see that $\mathcal{S}$ is convex, consider $\overline{x}, \overline{y} \in \mathcal{S}$, $\lambda \in [0,1]$ and let us show that $\lambda\overline{x} + (1-\lambda)\overline{y} \in \mathcal{S}$. Since $\overline{x}, \overline{y} \in \mathcal{S}$, then $f(\overline{x}) = f(\overline{y}) = \min_{x\in K} f(x)$. On the other hand from the quasiconvexity of f, one has

$$f(\lambda\overline{x} + (1-\lambda)\overline{y}) \leq \max\{f(\overline{x}), f(\overline{y})\} = \min_{x\in K} f(x) \leq f(\lambda\overline{x} + (1-\lambda)\overline{y}).$$

Hence $\lambda\overline{x} + (1-\lambda)\overline{y} \in \mathcal{S}$.
Now, let $\mathcal{N}$ be an open set such that $\mathcal{S} \subset \mathcal{N}$. Let us denote by $\alpha = dist(\mathcal{S}, \partial\mathcal{N})$, exists since $\mathcal{S}$ is compact. By contradiction suppose

$$\inf_{x\in K\cap\partial\mathcal{N}} f(x) = \inf_{x\in K} f(x).$$

Let $\{x_n\} \subset K \cap \partial\mathcal{N}$ be a sequence such that $\lim_{n\longrightarrow+\infty} f(x_n) = \inf_{x\in K} f(x)$ and let $\varepsilon_n > 0$ such that

$$f(x_n) < \inf_{x\in K} f(x) + \varepsilon_n.$$

Then by the Ekeland variational principle with $\lambda_n = \frac{2}{\alpha}\varepsilon_n$, there exists $w_n \in K$ such that $\Phi(w_n) \leq \Phi(x_n)$, $\|w_n - x_n\| \leq \frac{\alpha}{2}$ and

$$f(w_n) - \frac{2}{\alpha}\varepsilon_n\|x - w_n\| < f(x) \qquad \forall x \neq w_n.$$

Therefore, $f^0(w_n)(h) \geq -\frac{2}{\alpha}\varepsilon_n\|h\|$. Hence from the Palais-Smale condition, one deduces $w_{n_k} \longrightarrow \overline{w}$ for a subsequence. Since $f(w_{n_k}) \leq f(x_{n_k})$, then

$$f(\overline{w}) = \lim f(w_{n_k}) \leq \lim f(x_{n_k}) = \inf_{x\in K} f(x).$$

Hence $\overline{w} \in \mathcal{S}$. On the other hand

$$dist(\overline{w}, \partial\mathcal{N}) \leq \|\overline{w} - x_n\| \leq \|\overline{w} - w_n\| + \|w_n - x_n\| \leq \|\overline{w} - w_n\| + \frac{\alpha}{2}.$$

Therefore $dist(\overline{w}, \partial\mathcal{N}) \leq \frac{\alpha}{2}$ and hence $dist(\mathcal{S}, \partial\mathcal{N}) \leq \frac{\alpha}{2}$. Which is absurd. □

Lemma 6. *Suppose that for all $y \in K$, the function $x \mapsto -\Phi(x, y)$ is bounded below, locally Lipschitz, quasiconvex and satisfies the Palais-Smale condition. Then the set-valued mapping $\mathcal{S} : K \longrightarrow 2^K$ defined by*

$$\mathcal{S}(y) = \{\overline{x} \in K : \Phi(\overline{x}, y) = \sup_{x\in K} \Phi(x, y)\}$$

is upper semicontinuous with nonempty compact convex values.

Proof. From Lemma 5, $\mathcal{S}(y)$ is nonempty, compact and convex set for each $y \in K$. It remains to show that the set-valued mapping $\mathcal{S}$ is upper semicontinuous. We need to show that for $y_0 \in K$ and $\mathcal{N}$ an open set containing $\mathcal{S}(y_0)$, there is an open neighborhood $U(y_0)$ of y_0 such that $\forall y \in U(y_0)$ one has $\mathcal{S}(y) \subset \mathcal{N}$. By contradiction, suppose there exists $y_0 \in K$ and an open set $\mathcal{N}$ such that $\forall n \in I\!N^*$, $\exists y_n \in B(y_0, \frac{1}{n})$, $\exists x_n \in \mathcal{S}(y_n)$ with $x_n \notin \mathcal{N}$. Let $x_0 \in \mathcal{S}(y_0) \subset \mathcal{N}$. Since $x_n \notin \mathcal{N}$, then there exists $\alpha_n \in (0, 1)$ such that $w_n = (1-\alpha_n)x_0+\alpha_n x_n \in \partial\mathcal{N}$ the boundary of $\mathcal{N}$. One has

$$\sup_{n\in I\!N} \Phi(w_n, y_0) \leq \sup_{x\in\partial\mathcal{N}\cap K} \Phi(x, y_0) < \sup_{x\in K} \Phi(x, y_0) = \Phi(x_0, y_0). \tag{6}$$

On the other hand from the local Lipschitz property of Φ, there exists $L_y > 0$ such that

$$\Phi(w_n, y_n) - \Phi(w_n, y_0) \leq L_y\|y_n - y_0\|.$$

From the quasiconcavity of $x \mapsto -\Phi(x, y)$, one has

$$\Phi(w_n, y_n) = \Phi((1-\alpha_n)x_0+\alpha_n x_n, y_n) \geq \min\{\Phi(x_0, y_n), \Phi(x_n, y_n)\} \geq \Phi(x_0, y_n).$$

Therefore

$$\Phi(x_0, y_n) \leq \Phi(w_n, y_0) + L_y\|y_n - y_0\|.$$

Hence

$$\Phi(x_0, y_0) \leq \liminf \Phi(x_0, y_n) \leq \liminf \Phi(w_n, y_0) \leq \sup_{n\in I\!N} \Phi(w_n, y_0) < \Phi(x_0, y_0),$$

which is absurd, hence $\mathcal{S}$ is upper semicontinuous. □

Remark 4. Note that if Φ satisfies the assumptions in Lemma 6. Then the set-valued mapping $\mathcal{S}$ is closed graph, i.e. if $\{x_n\}$ and $\{y_n\}$ are sequences such that $x_n \in \mathcal{S}(y_n)$ with $x_n \longrightarrow \overline{x}$ and $y_n \longrightarrow \overline{y}$, then $\overline{x} \in \mathcal{S}(\overline{y})$.

Lemma 7. *Let X and Y be two metric spaces and $\mathcal{S} : X \to 2^Y$ be a multivalued mapping upper semicontinuous with nonempty compact values. Let $\{x_n\} \subset X$ be a sequence such that $x_n \to \overline{x}$. For $n \in I\!N$, consider $y_n \in \mathcal{S}(x_n)$. Then the sequence $\{y_n\}$ has a convergent subsequence to a point $\overline{y} \in \mathcal{S}(\overline{x})$.*

Proof. For $p \in I\!N^*$, consider the open set $\mathcal{N}_p$ defined by $\mathcal{N}_p = \{y \in Y : dist(y, \mathcal{S}(\overline{x})) < \frac{1}{p}\}$. Since $\mathcal{S}(\overline{x}) \subset \mathcal{N}_p$, $\mathcal{S}$ is upper semicontinuous and $x_n \to \overline{x}$, then there exists $N(p) \in I\!N$ such that for all $n \geq N(p)$ one has $\mathcal{S}(x_n) \subset \mathcal{N}_p$. Hence there exists $z_p \in \mathcal{S}(\overline{x})$ such that $dist(y_{N(p)}, z_p) < \frac{1}{p}$. Since $\mathcal{S}(\overline{x})$ is compact, then the sequence $\{z_p\}$ has a convergent subsequence to $\overline{y} \in \mathcal{S}(\overline{x})$. From the triangular inequality

$$dist(y_{N(p)}, \overline{y}) \leq dist(y_{N(p)}, z_p) + dist(z_p, \overline{y}),$$

one deduces $y_{N(p)} \to \overline{y}$. Which completes the proof. □

Definition 10. *[16] The bifunction Φ is said to be regular at $(\overline{x}, \overline{y})$ if*

$$\Phi^0(\overline{x}, \overline{y})(h) = \limsup_{\substack{t \longrightarrow 0^+ \\ (u,v) \longrightarrow (x,y) \\ v \in K}} \frac{\Phi(u, v+th) - \Phi(u,v)}{t} = \limsup_{t \longrightarrow 0^+} \frac{\Phi(\overline{x}, \overline{y}+th) - \Phi(\overline{x}, \overline{y})}{t}.$$

Now we can state our main result on existence of solution points for equilibrium problems in the noncompact case.

Theorem 2. *Suppose that Φ is regular and the following assumptions hold*

(i) there exists $x_0 \in K$ such that the function $y \mapsto \Phi(x_0, y)$is bounded below;
(ii) for $y \in K$ fixed, the function $x \mapsto -\Phi(x, y)$ is bounded from below, locally Lipschitz, quasiconvex and satisfies the Palais-Smale condition;
(iii) Φ satisfies the Generalized Palais-Smale condition.

Then there exists $\overline{x}, \overline{y} \in K$ such that $\Phi^0(\overline{x}, \overline{y})(h) \geq 0 \;\; \forall h \in K - \overline{y}$. Moreover, if $\forall x \in K$ the function $y \mapsto \Phi(x, y)$ is pseudoconvex then $\overline{x}$ is a solution to (EP).

Proof. Consider the function $V : K \longrightarrow K$ defined by $V(y) = \sup_{x \in K} \Phi(x, y)$. One can easily see that V is lower semicontinuous and bounded from below. From the Ekeland variational principle, one has $\forall \varepsilon > 0$ there exists $y_\varepsilon \in K$ such that

$$V(y_\varepsilon) \leq \inf_{y \in K} V(y) + \varepsilon, \tag{7}$$

$$V(y_\varepsilon) < V(y) + \varepsilon\|y - y_\varepsilon\| \quad \forall y \neq y_\varepsilon. \tag{8}$$

Consider $h \in K - y_\varepsilon$ and let $y = y_\varepsilon + th$, with $t > 0$. Then from relation (8), one has

$$\frac{V(y_\varepsilon + th) - V(y_\varepsilon)}{t} > -\varepsilon\|h\|.$$

Hence

$$\liminf_{t \longrightarrow 0^+} \frac{V(y_\varepsilon + th) - V(y_\varepsilon)}{t} > -\varepsilon\|h\|.$$

Let $\{t_n\}$ be a sequence of positive numbers such that $t_n \searrow 0$, then

$$\lim_{n \longrightarrow +\infty} \frac{V(y_\varepsilon + t_n h) - V(y_\varepsilon)}{t_n} \geq \liminf_{t \longrightarrow 0^+} \frac{V(y_\varepsilon + th) - V(y_\varepsilon)}{t} > -\varepsilon\|h\|. \tag{9}$$

On the other hand, by Lemma 5, $\mathcal{S}(y_\varepsilon + t_n h) \neq \emptyset$ and hence there exists $x_n \in K$ such that $V(y_\varepsilon + t_n h) = \Phi(x_n, y_\varepsilon + t_n h)$. Since $V(y_\varepsilon) \geq \Phi(x_n, y_\varepsilon)$, then from relation (9) one deduces

$$\liminf_{n \longrightarrow +\infty} \frac{\Phi(x_n, y_\varepsilon + t_n h) - \Phi(x_n, y_\varepsilon)}{t_n} > -\varepsilon\|h\|$$

For $\{\varepsilon_n\} \subset I\!R^+$ such that $\varepsilon_n \searrow 0^+$, one has

$$\frac{\Phi(x_n, y_\varepsilon + t_n h) - \Phi(x_n, y_\varepsilon)}{t_n} < \Phi^0(x_n, y_\varepsilon)(h) + \varepsilon_n$$

Hence

$$\liminf_{n \longrightarrow +\infty} [\Phi^0(x_n, y_\varepsilon)(h) + \varepsilon_n] \geq -\varepsilon\|h\|. \tag{10}$$

On the other hand, $x_n \in \mathcal{S}(y_\varepsilon + t_n h)$ and since $\mathcal{S}$ is upper semicontinuous with nonempty compact values then from Lemma 7 one deduces that $\{x_n\}$ has a convergent subsequence $x_{n_k} \longrightarrow x \in \mathcal{S}(y_\varepsilon)$. Therefore, from relation (10) and Lemma 1, one deduces

$$\Phi^0(x, y_\varepsilon)(h) \geq -\varepsilon\|h\|.$$

We have then established the following

$$\forall h \in X, \ \ \exists x \in K \quad \text{such that} \quad \Phi^0(x, y_\varepsilon)(h) \geq -\varepsilon\|h\|.$$

Now consider the following bifunction $\varphi_\varepsilon : \mathcal{S}(y_\varepsilon) \times (K - y_\varepsilon) \longrightarrow I\!R$ defined by

$$\varphi_\varepsilon(x, h) = \Phi^0(x, y_\varepsilon)(h) + \varepsilon\|h\|$$

The set $\mathcal{S}(y_\varepsilon)$ is compact and $(K - y_\varepsilon)$ is convex. It is easy to see that $\varphi_\varepsilon(x, \cdot)$ is convex, let us show that $\varphi_\varepsilon(\cdot, h)$ is quasiconcave. To this aim, let $\alpha \in [0, 1]$ and $x, y \in \mathcal{S}(y_\varepsilon)$. We need only to show that

$$\Phi^0(\alpha x + (1 - \alpha)y, y_\varepsilon)(h) \geq \min\{\Phi^0(x, y_\varepsilon)(h), \Phi^0(y, y_\varepsilon)(h)\}.$$

Since Φ is regular, then

$$\begin{aligned}&\Phi^0(\alpha x+(1-\alpha)y,y_\varepsilon)(h)\\&=\limsup_{t\to 0^+}\frac{\Phi(\alpha x+(1-\alpha)y,y_\varepsilon+th)-\Phi(\alpha x+(1-\alpha)y,y_\varepsilon)}{t}.\end{aligned}$$

Since $\Phi(\cdot,y_\varepsilon+th)$ is quasiconcave, then

$$\Phi(\alpha x+(1-\alpha)y,y_\varepsilon+th)\geq\min\{\Phi(x,y_\varepsilon+th),\Phi(y,y_\varepsilon+th)\}.$$

On the other hand, since $\mathcal{S}(y_\varepsilon)$ is convex then $\alpha x+(1-\alpha)y\in\mathcal{S}(y_\varepsilon)$ and

$$\Phi(\alpha x+(1-\alpha)y,y_\varepsilon)=\Phi(x,y_\varepsilon)=\Phi(y,y_\varepsilon).$$

Therefore

$$\frac{\Phi(\alpha x+(1-\alpha)y,y_\varepsilon+th)-\Phi(\alpha x+(1-\alpha)y,y_\varepsilon)}{t}\geq$$
$$\min\{\frac{\Phi(x,y_\varepsilon+th)-\Phi(x,y_\varepsilon)}{t},\frac{\Phi(y,y_\varepsilon+th)-\Phi(y,y_\varepsilon)}{t}\}.$$

Hence by passing to the limit when $t\to 0^+$, one deduces that

$$\Phi^0(\alpha x+(1-\alpha)y,y_\varepsilon)(h)\geq\min\{\Phi^0(x,y_\varepsilon)(h),\Phi^0(y,y_\varepsilon)(h)\}.$$

Consequently, φ_ε satisfies assumptions of Lemma 4. Then

$$\exists x_\varepsilon\in\mathcal{S}(y_\varepsilon)\text{ such that }\Phi^0(x_\varepsilon,y_\varepsilon)(h)+\varepsilon\|h\|\geq 0\quad\forall h\in K-y_\varepsilon.$$

For $\varepsilon_n\to 0^+$, one has

$$\Phi^0(x_{\varepsilon_n},y_{\varepsilon_n})(h)\geq-\varepsilon_n\|h\| \tag{11}$$

with $x_{\varepsilon_n}\in\mathcal{S}(y_{\varepsilon_n})$ and $h\in K-y_{\varepsilon_n}$. Hence, from the generalized Palais-Smale condition, $x_{\varepsilon_n}\longrightarrow\overline{x}$ and $y_{\varepsilon_n}\longrightarrow\overline{y}$ for subsequences. Moreover $\overline{x}\in\mathcal{S}(\overline{y})$. Passing to the limit in (11), one deduces

$$\Phi^0(\overline{x},\overline{y})(h)\geq 0\quad\forall h\in K-\overline{y}.$$

On the other hand $\overline{x}\in\mathcal{S}(\overline{y})$, then $\Phi(\overline{x},\overline{y})=\sup_{x\in K}\Phi(x,\overline{y})\geq\Phi(\overline{y},\overline{y})\geq 0$. Since the function $y\mapsto\Phi(\overline{x},y)$ is pseudoconvex, then one can easily verify that $\Phi(\overline{x},y)\geq 0$ for all $y\in K$. $\square$

Remark 5. Under the assumptions of Theorem 2, it has been established the following

$$\exists\overline{x},\overline{y}\in K\text{ such that }\Phi(\overline{x},\overline{y})\geq 0\text{ and }\Phi^0(\overline{x},\overline{y})(h)\geq 0\ \ \forall h\in K-\overline{y}. \tag{12}$$

Note that if $\overline{y}=\overline{x}$, then $\overline{x}$ is a critical point of the bifunction Φ according to Definition 3. Hence, relation (12) could be a more general formulation of a critical point for a bifunction.

Example

For an illustration of our approach, we give the following example. Let A be a bounded closed convex set of $\mathbb{R}^m$ containing the origin and let $\Omega \subset \mathbb{R}^m$ be a bounded open set. We consider the subset K of $W^{1,0}(\Omega, \mathbb{R}^m)$ defined by

$$K = \{x \in W^{1,0}(\Omega, \mathbb{R}^m) : x(\xi) \in A \;\; a.e. \quad \text{on } \Omega\}.$$

Consider $\varphi \in C^1(K \times K, \mathbb{R})$ and the bifunction Φ defined on $K \times K$ as the following

$$\Phi(x,y) = \int_\Omega [\frac{1}{2}|\nabla y|^2 - \frac{1}{2}|\nabla x|^2 + \varphi(x,y)] d\xi.$$

We shall verify that for $y \in K$ fixed, the function $x \mapsto -\Phi(x,y)$ satisfies the Palais-Smale condition and the bifunction Φ satisfies the generalized Palais-Smale condition.
To this aim, let $y_0 \in K$ fixed and $\{x_n\}$ be a sequence in K such that $\{-\Phi(x_n, y_0)\}$ is bounded and

$$(-\Phi)^0(x_n, y_0)(h) \geq -\varepsilon_n \|h\| \quad \text{for all} \quad h \in K - x_n.$$

Hence, the sequence

$$\int_\Omega \frac{1}{2}|\nabla x_n|^2 - \varphi(x_n, y_0)\; d\xi \quad \text{is bounded} \tag{13}$$

$$\int_\Omega \nabla x_n \cdot \nabla h - \varphi'_x(x_n, y_0)(h)\; d\xi \geq -\varepsilon_n \|h\| \quad \text{for all} \;\; h \in K - x_n. \tag{14}$$

From relation (13), one deduces that $\{x_n\}$ is bounded in $W^{1,0}(\Omega, \mathbb{R}^m)$. Hence, for a subsequence one has $x_{n_k} \rightharpoonup \overline{x}$ for the weak topology on $W^{1,0}(\Omega, \mathbb{R}^m)$ and $x_{n_k} \to \overline{x}$ a.e. on Ω and thus $\overline{x} \in K$. Let $h_0 \in K - \overline{x}$, we shall prove that

$$\int_\Omega \nabla \overline{x} \cdot \nabla h_0 - \varphi'_x(\overline{x}, y_0)(h_0)\; d\xi \geq 0.$$

To this aim, set $z_{n_k} = \overline{x} + h_0 - x_{n_k}$. Then

$$\begin{aligned}
&\int_\Omega \nabla \overline{x} \cdot \nabla h_0 \; d\xi = \int_\Omega \nabla(\overline{x} - x_{n_k}) \cdot \nabla h_0 \; d\xi + \int_\Omega \nabla x_{n_k} \cdot \nabla h_0 \; d\xi \\
&= \int_\Omega \nabla(\overline{x} - x_{n_k}) \cdot \nabla h_0 \; d\xi + \int_\Omega \nabla x_{n_k} \cdot \nabla z_{n_k} \; d\xi + \int_\Omega \nabla x_{n_k} \cdot \nabla(x_{n_k} - \overline{x}) \; d\xi \\
&\geq \int_\Omega \nabla x_{n_k} \cdot \nabla z_{n_k} \; d\xi + \theta(x_{n_k} - \overline{x})
\end{aligned}$$

where $\theta(\overline{x} - x_{n_k}) \to 0$ since $x_{n_k} \rightharpoonup \overline{x}$ for the weak topology on $W^{1,0}(\Omega, \mathbb{R}^m)$.
On the other hand

$$\int_\Omega \varphi'_x(\overline{x}, y_0)(h_0)\; d\xi = \int_\Omega \varphi'_x(x_{n_k}, y_0)(h_0)\; d\xi + \theta(x_{n_k} - \overline{x}).$$

Hence

$$\begin{aligned}
&\int_\Omega \nabla\overline{x}\cdot\nabla h_0 - \varphi'_x(\overline{x},y_0)(h_0)\ d\xi \\
&\geq \int_\Omega \nabla x_{n_k}\cdot\nabla z_{n_k}\ d\xi - \int_\Omega \varphi'_x(x_{n_k},y_0)(h_0)\ d\xi + \theta(x_{n_k}-\overline{x}) \\
&\geq \int_\Omega \nabla x_{n_k}\cdot\nabla z_{n_k}\ d\xi - \int_\Omega \varphi'_x(x_{n_k},y_0)(z_{n_k}+(\overline{x}-x_{n_k}))\ d\xi + \theta(x_{n_k}-\overline{x}) \\
&\geq \int_\Omega \nabla x_{n_k}\cdot\nabla z_{n_k}\ d\xi - \int_\Omega \varphi'_x(x_{n_k},y_0)(z_{n_k})\ d\xi + \theta(x_{n_k}-\overline{x}) \\
&\geq -\varepsilon_{n_k}\|z_{n_k}\| + \theta(x_{n_k}-\overline{x})
\end{aligned}$$

It follows that

$$\int_\Omega \nabla\overline{x}\cdot\nabla h_0 - \varphi'_x(\overline{x},y_0)(h_0)\ d\xi \geq 0. \tag{15}$$

Set $h=\overline{x}-x_{n_k}$ in relation (14) and $h_0=x_{n_k}-\overline{x}$ in relation (15). Then

$$\begin{aligned}
&\int_\Omega \nabla x_n\cdot\nabla(\overline{x}-x_{n_k}) - \varphi'_x(x_n,y_0)(\overline{x}-x_{n_k})\ d\xi \geq -\varepsilon_{n_k}\|\overline{x}-x_{n_k}\| \\
&\int_\Omega \nabla\overline{x}\cdot\nabla(x_{n_k}-\overline{x}) - \varphi'_x(\overline{x},y_0)(x_{n_k}-\overline{x})\ d\xi \geq 0.
\end{aligned}$$

By adding the above two relations, one obtain

$$\begin{aligned}
-\int_\Omega |\nabla(\overline{x}-x_{n_k})|^2 - \int_\Omega [\varphi'_x(\overline{x},y_0)(\overline{x}-x_{n_k}) + \varphi'_x(x_n,y_0)(\overline{x}-x_{n_k})]\ d\xi \\
\geq -\varepsilon_{n_k}\|\overline{x}-x_{n_k}\|.
\end{aligned}$$

Therefore

$$\int_\Omega |\nabla(\overline{x}-x_{n_k})|^2\ d\xi \leq \varepsilon_{n_k}\|\overline{x}-x_{n_k}\| + \theta(x_{n_k}-\overline{x}).$$

It follows that $x_{n_k}\to\overline{x}$ in $W^{1,0}(\Omega,\mathbb{R}^m)$.

Now, let $\{(x_n,y_n)\}$ be a sequence in $K\times K$ such that $\Phi(x_n,y_n)=\sup_{x\in K}\Phi(x,y_n)$, $\{\Phi(x_n,y_n)\}$ is bounded and $\Phi^0(x_n,y_n)(h)\geq -\varepsilon_n\|h\|$ for all $h\in K-y_n$. It follows that the sequence

$$\int_\Omega [\frac{1}{2}|\nabla y_n|^2 - \frac{1}{2}|\nabla x_n|^2 + \varphi(x_n,y_n)]d\xi \quad \text{is bounded} \tag{16}$$

$$\int_\Omega \nabla y_n\cdot\nabla h + \varphi'_y(x_n,y_n)(h)\ d\xi \geq -\varepsilon_n\|h\| \quad \text{for all} \quad h\in K-y_n. \tag{17}$$

On the other hand, since $\Phi(x_n,y_n)\geq\Phi(x,y_n)$ for all $x\in K$, then

$$\varphi'_x(x_n,y_n)(h)\leq 0 \quad \text{for all} \quad h\in K-x_n.$$

Hence

$$-\int_{\Omega} \nabla x_n \cdot \nabla h + \varphi'_x(x_n, y_n)(h)\ d\xi \leq 0 \quad \text{for all} \quad h \in K - x_n. \tag{18}$$

For $x_0 \in K$ fixed, set $h = x_0 - x_n$, then

$$-\int_{\Omega} \nabla x_n \cdot \nabla(x_0 - x_n) + \varphi'_x(x_n, y_n)(x_0 - x_n)\ d\xi \leq 0.$$

It follows

$$\int_{\Omega} -|\nabla x_n|^2 + \varphi'_x(x_n, y_n)(x_0 - x_n)\ d\xi - \int_{\Omega} \nabla x_n \cdot \nabla x_0\ d\xi \leq 0.$$

Thus

$$\int_{\Omega} -|\nabla x_n|^2\ d\xi \leq \|x_n\|\ \|x_0\| + c(\|x_n\| + \|x_0\|)$$

where c is a positive constant. It follows that $\{x_n\}$ is bounded in $W^{1,0}(\Omega, I\!R^m)$. From relation (16), one deduces that $\{y_n\}$ is bounded in $W^{1,0}(\Omega, I\!R^m)$. Hence, for a subsequence one has $x_{n_k} \rightharpoonup \overline{x}$ and $y_{n_k} \rightharpoonup \overline{y}$ for the weak topology on $W^{1,0}(\Omega, I\!R^m)$. Thus $x_{n_k} \to \overline{x}$ and $y_{n_k} \to \overline{y}$ a.e. on Ω.
Similarly as before, we can show that $\overline{x}, \overline{y} \in K$ and

$$\int_{\Omega} \nabla \overline{y} \cdot \nabla h + \varphi'_y(\overline{x}, \overline{y})(h)\ d\xi \geq 0 \quad \text{for all} \ \ h \in K - \overline{y}. \tag{19}$$

From relations (17) and (19), one can verify

$$\|y_{n_k} - \overline{y}\|^2 \leq -\varepsilon_{n_k} \|y_{n_k} - \overline{y}\| + \theta(y_{n_k} - \overline{y}).$$

Hence $y_{n_k} \to \overline{y}$ in $W^{1,0}(\Omega, I\!R^m)$. Similarly, we have relation (15) and therefore one can deduce $x_{n_k} \to \overline{x}$ in $W^{1,0}(\Omega, I\!R^m)$.

Acknowledgements

The authors would like to express their sincere thanks to the referees, whose careful reading greatly improved this article.

References

1. Blum E, Oettli W (1994) From optimization and variational inequalities to equilibrium problems. Math. Student 63:123–145
2. Konnov I (2005) Generalized monotone equilibrium problems and variational inequalities. In: Hadjisavvas N, Komlósi S, Schaible S (eds) Handbook of generalized convexity and generalized mononicity. Springer, Berlin Heidelberg New York 559-618

3. Bianchi M, Schaible S (1996) Generalized monotone bifunctions and equilibrium problems. J. Optim. Theory Appl. 90:31–43
4. Hadjisavvas N, Schaible S (1998) Quasimonotonicity and pseudomonotonicity in variational inequalities and equilibrium problems. In: Crouzeix J-P, Martinez-Legaz J-E, Volle M (eds) Generalized convexity, generalized monotonicity. Kluwer Academic Publishers, Dordrecht-Boston-London 257-275
5. Chadli O, Chbani Z, Riahi H (2000) Equilibrium problems with generalized monotone bifunctions and applications to variational inequalities. J. Optim. Theory Appl. 105:299–323
6. Bianchi M, Pini R (2005) Coercivity conditions for equilibrium problems. J. Optim. Theory Appl. 124:79–92
7. Chadli O, Chbani Z, Riahi H (1999) Recession methods for equilibrium problems and applications to variational and hemivariational inequalities. Discrete and Continuous Dynamical Systems 5:185–195
8. Chadli O, Chbani Z, Riahi H (2001) Equilibrium problems and noncoercive variational inequalities. Optimization 50:17-27
9. Flores-Bazán F (2000) Existence theorems for generalized noncoercive equilibrium problems:The quasiconvex case. SIAM Journal on Optimization 11:675-690
10. Chadli O, Schaible S, Yao J C (2004) Regularized equilibrium problems with application to noncoercive hemivariational inequalities. J. Optim. Theory Appl. 121:571–596
11. Bianchi M, Kassay G, Pini R (2005) Existence of equilibria via Ekeland's principle. J. Math. Anal. Appl. 305:502–512
12. Brézis H, Haraux A (1976) Image d'une somme d'opérateurs monotones et applications, Israel Journal of Mathematics 23:165-186
13. Hadjisavvas N (2005) Generalized convexity, generalized monotonicity and nonsmooth analysis. In: Hadjisavvas N, Komlósi S, Schaible S (eds) Handbook of generalized convexity and generalized mononicity. Springer, Berlin Heidelberg New York 463-499
14. Chang K C (1981) Variational methods for non-differentiable functionals and their applications to partial differential equations, J. Math. Anal. Appl. 80:102–129
15. Shi S-Z, Chang K-C (1987) A local minimax theorem without compactness. In: Bor-Luh Lin, Stephen Simons (eds) Nonlinear and Convex Analysis. Marcel Dekker, 211-233
16. Clarke FH (1983) Nonsmooth analysis and optimization. Willey, New York.

Points of Efficiency in Vector Optimization with Increasing-along-rays Property and Minty Variational Inequalities

Giovanni P. Crespi[1], Ivan Ginchev[2], and Matteo Rocca[3]

[1] Université de la Vallé d'Aoste, Faculty of Economics, 11100 Aosta, Italy
g.crespi@univda.it
[2] Technical University of Varna, Department of Mathematics, 9010 Varna, Bulgaria ginchev@ms3.tu-varna.acad.bg
[3] University of Insubria, Department of Economics, 21100 Varese, Italy
mrocca@eco.uninsubria.it

Summary. Minty variational inequalities are studied as a tool for vector optimization. Instead of focusing on vector inequalities, we propose an approach through scalarization which allows to construct a proper variational inequality type problem to study any concept of efficiency in vector optimization.
This general scheme gives an easy and consistent extension of scalar results, providing also a notion of increasing along rays vector function. This class of generalized convex functions seems to be intimately related to the existence of solutions to a Minty variational inequality in the scalar case, we now extend this fact to vector case.
Finally, to prove a reversal of the main theorem, generalized quasiconvexity is considered and the notion of $*$-quasiconvexity plays a crucial role to extend scalar evidences. This class of functions, indeed, guarantees a Minty-type variational inequality is a necessary and sufficient optimality condition for several kind of efficient solution.

Key words: Minty variational inequalities, vector optimization, generalized quasiconvexity, scalarization.

1 Introduction

Variational inequalities (for short, VI) provide suitable mathematical models for a range of practical problems, see e.g. [3] or [29]. Vector VI were introduced first in [20] and thereafter studied intensively. For a survey and some recent results we refer to [2, 8, 17, 21, 47, 30]. Stampacchia [41] and Minty [37] type VI are the main versions which are studied. One of the most challenging field of research studies relations among solutions of a "differential" VI and those of a primitive optimization problem. This task has been the core of investigation for vector extension of the variational inequalities since the beginning [5, 6, 9, 31, 32, 46, 47]. Nowadays, when the problem focuses on Minty VI (both scalar and vector) and optimization, the guideline for the research is the well known "Minty Variational Principle" (see e.g. [21]). The principle provides a test of goodness for any formulation of a vector Minty VI. Namely an inequality is a differentiable Minty VI if and only if its solution is an optimal solution (of a certain type) to the primitive multiobjective optimization problem. This must hold without any assumption on the objective function, but differentiability. In [10, 12] and [16], two different approaches have been proposed. First in [10] and [16] a vector inequality has been related to efficiency and proper efficiency (for differentiable vector optimization problems), filling a gap left by the formulation given by Giannessi. Then in [12] the same gap has been filled by means of generalized scalar Minty VI, that is without using vector inequality but a scalar VI involving set-valued maps. Meanwhile, also the Minty variational principle in the scalar case has been deeply investigated. By means of lower Dini derivatives we enlarged the class of optimization problems to be studied through the inequalities and we pointed out a more detailed form of the principle itself (see e.g. [13, 11]), which involves also the notion of Increasing-Along-Rays (IAR) functions as generalized convex. The next step has been to go back to vector optimization to try to prove a vector version of this principle. However, as several notions of vector-optima can be given, also different results on the Minty VI can be proved. Some attempts can be found in [14] and [15]. Basically one could think to develop a theorem for each kind of efficient solution known in the literature based on a suitable inequality for the vector-valued formulation. Here we wish to prove that results known and new ones can be basically related to a common scheme, based on scalarization. To see this, in Section 2 we define vector solution through general scalarization, as a mean to construct the general approach to the problem. Once this result is achieved, next sections are devoted to show how very special results can be obtained within this scheme and one shall recognize how classical vector concepts lay behind those in Section 2. Clearly when the problem is scalar, the classical results come as a special case. To clarify the idea, we stress that notions as ideal, efficient, weak efficient and the wide variety of proper efficient solutions are well established in the literature. Several of them are also studied by means of a proper scalarization technique. On the contrary, in the field of (Minty) variational inequalities, we think the theory is not yet established and what are the notions of solution is still to be clarified. For some of the results we also discuss relation (if any) with concept of solution to a vector Minty VI introduced by other authors. Finally the concept of vector IAR function is far to be a classic in optimization and generalized convexity community. Only few results can be guessed by analogy with classical monotonicity for vector valued functions. Here we propose a more rigorous approach to the definition of this property. To complete the exploration we also extend the properties of quasiconvex functions as related to Minty VI. We show that some widely accepted concepts of vector quasiconvex functions can be fitted within the scalarization scheme we develop in Section 2.

2 A General Scheme

In the sequel X denotes a real linear space and K is a convex subset of X. Further Y is a real topological vector space and $C \subset Y$ is a closed convex cone.

In [11] we consider the scalar case $Y = \boldsymbol{R}$ and investigate the scalar (generalized) Minty VI of differential type

$$f'(x, x^0 - x) \leq 0, \quad x \in K\,, \tag{1}$$

were $f'(x, x^0 - x)$ is the Dini directional derivative of f at x in direction $x^0 - x$. For x in K and $u \in X$ we define the Dini derivative

$$f'(x, u) = \liminf_{t \to 0^+} \frac{1}{t}\ (f(x + tu) - f(x)) \tag{2}$$

as an element of the extended real line $\overline{\boldsymbol{R}} = \boldsymbol{R} \cup \{-\infty\} \cup \{+\infty\}$.

The following result is established in [11].

Theorem 1. *Let K be a set in a real linear space and let the function $f : X \to \mathbb{R}$ be radially lsc on the rays starting at $x^0 \in ker\,K$. Then x^0 is a solution of the Minty VI (1) if and only if f increases along rays starting at x^0. In consequence, each such solution x^0 is a global minimizer of f.*

Recall that $f : K \to \boldsymbol{R}$ is said radially lower semicontinuous on the rays starting at x^0 if for all $u \in X$ the composition function $t \to f(x^0 + tu)$ is lsc on the set $\{t \geq 0 \mid x^0 + tu \in K\}$. We write then $f \in RLSC(K, x^0)$. In a similar way we can introduce other "radial notions". We write also $f \in IAR(K, x^0)$ if f increases along rays starting at x^0, the latter means that for all $u \in X$ the function $t \to f(x + tu)$ is increasing on the set $\{t \geq 0 \mid x^0 + tu \in K\}$. We call this property the IAR property. The kernel $ker\,K$ of K is defined as the set of all $x^0 \in K$, for which $x \in K$ implies that $[x^0, x] \subset K$, where $[x^0, x] = \{(1-t)x^0 + tx \mid 0 \leq t \leq 1\}$ is the segment determined by x^0 and x. Obviously, for a convex set $ker\,K = K$. Sets with nonempty kernel are star-shaped and play an important role in the abstract convexity [39].

In [14] we generalize some results of [11] to a vector VI of the form

$$f'(x, x^0 - x) \cap (-C) \neq \emptyset\,, \quad x \in K\,, \tag{3}$$

where the Dini derivative $f'(x, u)$ is defined as

$$f'(x, u) = Limsup_{t \to 0^+} \frac{1}{t}(f(x + tu) - f(x)) \tag{4}$$

and the *Limsup* is intended in the sense of Painlevé-Kuratowski.

Let us underline some of the difficulties, which arise in the previous formulation. First, due to the use of infinite elements, the vector Dini derivative (4) is not exactly a generalization of the scalar Dini derivative (2). This motivated us to study in [14] infinite elements in the topological vector space Y and subsequently to introduce slightly different notions of the vector Dini derivative and the vector VI. Second, IAR property for vector-valued functions can be introduced in various ways. In [14] we propose two such generalizations, called respectively IAR^- and IAR^+ properties, based on alternative understandings of vector inequalities. Third, since a global minimizer (point of efficiency) for a vector function can be defined in different ways, Theorem 1 can be extended to copy with any of these.

We recall here some of the most standard points of efficiency. The point $x^0 \in K$ is said to be an ideal (or absolute) efficient point for $f : K \to Y$ if $f(K) \subset f(x^0) + C$. We call the ideal efficient points a-minimizers. The point $x^0 \in K$ is said to be an efficient point (e-minimizer) for $f : K \to Y$ if $f(K) \cap (f(x^0) - C \setminus \{0\}) = \emptyset$. The point $x^0 \in K$ is said to be a weakly efficient point (w-minimizer) for $f : K \to Y$ if $f(K) \cap (f(x^0) - \mathit{int}\, C) = \emptyset$.

In [14] we studied the vector VI (3) and the extension of Theorem 1 requires the notion of a-minimizers. Despite they might be unlike to happen, if the vector optimization problem

$$\min_C f(x)\,, \quad x \in K\,, \tag{5}$$

possesses a-minimizers, we certainly wish to distinguish them, since they represent a rather nice property.

Since problem (5) possesses rather e-minimizers (or w-minimizers) than a-minimizers, the natural question is whether f can define a type of VI and a type of IAR property, so that the equivalence of the properties x^0 *is a solution of the VI* and f *is increasing-along-rays starting at* x^0 remains true, and x^0 is either e-minimizers or w-minimizers. A step further appears in [15], where the notion of w-minimizer is characterized by a Minty VI. However [14] and [15] reveal a common approach in the proofs, which motivates our interest to develop a general scheme which may define the suitable Minty VI for any point of efficiency.
Since in both [14] and [15] proofs are based on scalarization, we need first to set a general notation for this technique.

Let the function $f : K \to Y$ be given and Ξ be a set of functions $\xi : Y \to \boldsymbol{R}$. For $x^0 \in \mathit{ker}\, K$ put $\Phi(\Xi, x^0)$ is the set of all functions $\phi : K \to \boldsymbol{R}$ such that $\phi(x) = \xi(f(x) - f(x^0))$ for some $\xi \in \Xi$. Instead of a single VI we consider the system of scalar VI

$$\phi'(x, x^0 - x) \le 0\,, \quad x \in K\,, \quad \text{for all} \quad \phi \in \Phi(\Xi, x^0)\,. \tag{6}$$

A solution of (6) is any point x^0, which solves all the scalar VI of the system.

Now we say that f is increasing-along-rays with respect to Ξ (Ξ-IAR) at x^0 along the rays starting at $x^0 \in \mathit{ker}\, K$, and write $f \in \Xi\text{-}IAR(K, x^0)$, if $\phi \in IAR(K, x^0)$ for all $\phi \in \Phi(\Xi, x^0)$. We say that $x^0 \in K$ is a Ξ-minimizer of f on K if x^0 is a minimizer on K of each of the scalar functions $\phi \in \Phi(\Xi, x^0)$. We say that the function f is radially Ξ-lsc at the rays starting at x^0, and write $f \in \Xi\text{-}RLSC(K, x^0)$, if all the functions $\phi \in \Phi(\Xi, x^0)$ satisfy $\phi \in RLSC(K, x^0)$.

Under this formulation we extend every concept involved in Theorem 1 to the vector case by using a family of scalar VI, which depends upon Ξ. The main idea is similar to the approach we proposed in [12], that is, to avoid troubles with vector inequalities, we look for scalar counterparts. The price we pay is that the scalar problem may not be easy to solve.

Remark 1. The variational inequality (1), where $Y = \boldsymbol{R}$, can be treated within this scheme if we put $\Xi = \{\xi\}$ to be the set consisting of the identical function $\xi : \boldsymbol{R} \to \boldsymbol{R}$, $\xi(y) = y$. Now the system (6) consists of just one VI, in which $\phi(x) = f(x) - f(x^0)$, $x \in K$. Obviously $\phi'(x, u) = f'(x, u)$ for all $x \in K$ and $u \in X$.

Let us underline that there is certain advantage in defining (6) through compositions $\xi(f(x) - f(x^0))$, in which the inner function $f(x)$ is translated by $f(x^0)$, as we can deal with some applications.

The main advantage to deal with families of scalarized inequalities is that we can easily apply scalar results as Theorem 1. Although trivial to prove, the following theorem is the general scheme we are looking for.

Theorem 2. *Let K be a set in a real linear space and Ξ be a set of functions $\xi : Y \to \boldsymbol{R}$ on the topological vector space Y. Let the function $f : K \to Y$ satisfy $f \in \Xi\text{-}RLSC(K, x^0)$ at the point $x^0 \in \mathit{ker}\, K$. Then x^0 is a solution of the system of VI (6) if and only if $f \in \Xi\text{-}IAR(K, x^0)$. In consequence, any solution $x^0 \in \mathit{ker}\, K$ of (6) is a Ξ-minimizer of f.*

Proof. The system (6) consists in fact of independent VI

$$\phi'(x,\, x^0 - x) \leq 0\,, \quad x \in K\,, \tag{7}$$

where $\phi \in \Phi(\Xi, x^0)$. The assumption that f is radially Ξ-lsc along the rays starting at x^0 means that the function ϕ in each of scalar VI (7) is radially lsc along the rays starting at x^0. According to Theorem 1 the point $x^0 \in \mathit{ker}\, K$ is a solution of (7) if and only if $\phi \in IAR(K, x^0)$. In consequence x^0 is a global minimizer of ϕ. Since this is true for each of the inequalities of the system (6), we get immediately the theses.

The generality of Ξ find a reason in the following sections, where it is clear that, given a choice of which points of efficiency we are willing to study, we can find a class Ξ to construct a Minty VI and a concept of IAR function.

Remark 2. Despite when dealing with VI in the vector case an ordering cone should be given in advance (see e.g. [20, 16], C does not appear explicitly neither in the system of VI (6) nor in the statement of the theorem. Therefore, the result of Theorem 2 depends on the set Ξ, but not on C. Actually, since the VI is related to a vector optimization problem, the cone C is given in advance because of the nature of the problem itself. The adequate system of VI claims then for a reasonable choice of Ξ depending in some way on C. In such a case the result in Theorem 2 depends implicitly on C through Ξ.

Remark 3. When $\Xi = \{\xi^0\}$ is a singleton, then Theorem 2 easily reduces to Theorem 1, where f should be substituted by $\phi : K \to \boldsymbol{R}$, $\phi(x) = \xi^0(f(x) - f(x^0))$, and the VI (1) by a single scalar VI of the form (7). Obviously, now f radially Ξ-lsc means ϕ radially lsc, $f \in \Xi\text{-}IAR(K, x^0)$ means $\phi \in IAR(K, x^0)$, x^0 a Ξ-minimizer of f means x^0 a minimizer of ϕ.

Another problem faced in [11] is the reversal of Theorem 1. namely we cannot ensure that when a minimizer exists, it is the solution of Minty VI 1. A counter example can be stated even in a differentiable case.

Example 1. Consider the scalar VI (1) with $f : \boldsymbol{R}^2 \to \boldsymbol{R}$ given by

$$f(x_1, x_2) = \begin{cases} x_1^2 x_2^2 \,, & x_1 \geq 0 \text{ or } x_2 \geq 0\,, \\ 0\,, & x_1 < 0 \text{ and } x_2 < 0\,. \end{cases}$$

The function f is C^1. The scalar VI (1) has a solution $x^0 = 0$, hence the set of solutions is non-empty (we have $f'(x, x^0 - x) = -4x_1^2 x_2^2$ if $x_1 \geq 0$ or $x_2 \geq 0$, and $f'(x, x^0 - x) = 0$ if $x_1 < 0$ and $x_2 < 0$). The set of the global minimizers of f is $\{x \in \boldsymbol{R}^2 \mid x_1 \leq 0,\, x_2 \leq 0\}$. At the same time a point x with $x_1 < 0$ and $x_2 < 0$ is not a solution of the VI, since for such a point the IAR property does not hold. (There is a ray starting at x, which intersects the coordinate axes in the different points x^1 and x^2. The function f takes value 0 at both x^1 and x^2, and f is strictly positive at the relative interior of the segment determined by x^1 and x^2.)

Reversal of Theorem 1 hold only with further assumptions on f, namely quasiconvexity. We say that $f : K \to \boldsymbol{R}$ is radially quasiconvex along rays starting at $x^0 \in ker\, K$ if the restriction of f to any such ray is quasiconvex. If this property is satisfied we write $f \in RQC(K, x^0)$.

The following assertion is a straightforward consequence of the definitions of quasiconvexity.

Theorem 3. *The function $f : K \to \boldsymbol{R}$ defined on a set K in a real linear space is quasiconvex if and only if K is convex and $f \in RQC(K, x^0)$ for all points $x^0 \in K$.*

The following theorem together with Theorem 1 establishes the equivalence among solutions of the scalar VI (1) and minimizers of f.

Theorem 4. *Let K be a set in a real linear space and let the function $f : X \to \mathbb{R}$ have the property $f \in RQC(K, x^0)$ at $x^0 \in \ker K$. If x^0 is a minimizer of f, then x^0 is a solution of the scalar VI (1). In particular, if f is quasiconvex, then any minimizer of f is a solution of VI (1).*

Proof. The properties that x^0 is a minimizer of f and $f \in RQC(K, x^0)$ imply that $f \in IAR(K, x^0)$. In consequence, according to Theorem 1 the point x^0 is a solution of the VI. Let us remark that in Theorem 1 for this implication the assumption for radial lower semicontinuity is superfluous.

To generalize Theorem 4 from the scalar VI (1) to the system of VI (6) we introduce Ξ-quasiconvexity.

Let Ξ be a set of functions $\xi : Y \to \boldsymbol{R}$. For $x^0 \in \ker K$ define the functions $\phi \in \Phi(\Xi, x^0)$ as in (6). For any such x^0 we say that f is radially Ξ-quasiconvex along the rays starting at x^0, and write $f \in \Xi\text{-}RQC(K, x^0)$, if $\phi \in RQC(K, x^0)$ for all $\phi \in \Phi(\Xi, x^0)$. We say that f is Ξ-quasiconvex if K is convex and $f \in \Xi\text{-}RQC(K, x^0)$ for all $x^0 \in K$. Clearly, for scalar f, the choice of Ξ as in Remark 1 guaranties $\Xi\text{-}RQC(K, x^0)$ reduces to $RQC(K, x^0)$.

The following theorem generalizes Theorem 4. The proof follows straightforward from Theorem 4 and is omitted.

Theorem 5. *Let K be a set in a real linear space and Ξ be a set of functions $\xi : Y \to \boldsymbol{R}$ on the topological vector space Y. Let the function $f : K \to Y$ have the property $f \in \Xi\text{-}RQC(K, x^0)$ at the point $x^0 \in \ker K$. If x^0 is a Ξ-minimizer of f, then x^0 is a solution of the system of VI (6). In particular, if f is Ξ-quasiconvex, then any Ξ-minimizer of f is a solution of (6).*

3 a-Minimizers

We assume that the topological vector space Y is locally convex and denote by Y' its dual space. The positive polar cone of C is defined by

$$C' = \{\xi \in Y' \mid \langle \xi,\, y \rangle \geq 0 \quad \text{for all} \quad y \in C\}\,.$$

Due to the Separation Theorem for topological vector spaces, see Theorem 9.1 in [40], for the closed and convex cone C we have $C = \{y \in Y \mid \langle \xi, y \rangle \geq 0 \text{ for all } \xi \in C'\}$.

Consider the system of VI (6) with $\Xi = C'$. Now $\Phi(\Xi, x^0)$ is the set of functions $\phi : K \to \boldsymbol{R}$ such that $\phi(x) = \langle \xi,\, f(x) - f(x^0) \rangle$, $x \in K$, for some $\xi \in C'$.

Under this setting, Theorem 2 allows to characterize a-minimizers by means of Minty VI. We first recall in [14] it was named IAR$^+$ the property of a function $f : K \to Y$ to increase along rays starting at $x^0 \in K$, when for every $0 \leq t_1 < t_2$

$$f(x^0 + t_2 u) \in f(x^0 + t_1 u) + C.$$

This fact has been denoted as $f \in IAR^+(K, x^0)$.

Theorem 6. *Let K be a set in a real linear space and C be a closed convex cone in the locally convex space Y. Let the function $f : K \to Y$ be such that all the functions $x \in K \to \langle \xi, f(x) \rangle \in \boldsymbol{R}$, $\xi \in C'$, are radially lsc along the rays starting at the point $x^0 \in \ker K$. Then x^0 is a solution of the system of VI (6) with $\Xi = C'$ if and only if $f \in IAR^+(K, x^0)$. In consequence, any solution $x^0 \in \ker K$ of (6) is an a-minimizer of f.*

Proof. Now $f \in \Xi\text{-}IAR(K, x^0)$ means that for arbitrary $u \in X$ and $0 \leq t_1 < t_2$ in the set $\{t \geq 0 \mid x^0 + tu \in K\}$, it holds

$$\langle \xi, f(x^0 + t_1 u) - f(x^0 + t_2 u) \rangle \leq 0$$

for all $\xi \in C'$. The latter gives that $f(x^0 + t_1 u) - f(x^0 + t_2 u) \in -C$, or equivalently $f(x^0 + t_2 u) \in f(x^0 + t_1 u) + C$, i.e. $f \in IAR^+(K, x^0)$. Similarly we get that x^0 is a Ξ-minimizer of f if $\langle \xi, f(x^0) - f(x) \rangle \leq 0$ for all $x \in K$ and $\xi \in C'$, which is equivalent to $f(x) \in f(x^0) + C$ for all $x \in K$, i.e. x^0 is an a-minimizer of f on K.

The next example shows an application of this result.

Example 2. Let $X = \boldsymbol{R}$, $Y = \boldsymbol{R}^2$, $C = \boldsymbol{R}^2_+$ and $K = \boldsymbol{R}_+$. The function $f(x) = \left(x,\, 2\sqrt{x}\right)$ defines, for $x^0 = 0$ the VI system

$$\left(\xi_1 + \frac{\xi_2}{\sqrt{x}}\right)(-x) = -\left(\xi_1 x + \xi_2 \sqrt{x}\right) \leq 0 \quad \forall x \in K$$

Since this inequality holds true for any choice of $\xi \in C'$, we conclude $x^0 = 0$ is an a-minimizer, as can be easily checked directly.

Remark 4. The Minty VI we are using is not a vector inequality. However it can be related to some vector inequality similar to VI 3. More problems arise than just those involved by the order. Indeed to copy with the infinite values allowed by scalar VI 2, it is necessary to explore the concept of infinite elements for vectors. In [14] this effort has been done and the relations with scalarized inequalities are studied. Leaving details to [14], here we stress that Theorem 6 provides a more straightforward result to handle similar cases which are presented in [14].

The approach we just have used bases on linear scalarization. One may complain that in real world applications the feasible values of ξ may be given in advance, following some constraints. In Remark 2 we have stressed it is not necessary to know C in advance to have optimality. We can start from Ξ.

Let Ξ be an arbitrary set in the dual space Y'. Now $\Phi(\Xi, x^0)$ is the set of functions $\phi : K \to \boldsymbol{R}$ such that $\phi(x) = \langle \xi,\, f(x) - f(x^0) \rangle$, $x \in K$, with some $\xi \in \Xi$. Define the cone $C_\Xi = \{y \in Y \mid \langle \xi,\, y \rangle \geq 0 \text{ for all } \xi \in \Xi\}$. Its positive polar cone is $C'_\Xi = cl\, cone\, conv\, \Xi$. We note that, despite $\Xi \subset C'_\Xi$, the solutions of system of VI (6) coincide with those of the system of VI obtained from (6) by replacing Ξ with C'_Ξ. However the new system allows to recover the case already described, with the cone C_Ξ replacing C

Corollary 1. *Let K be a set in a real linear space, Y be a locally convex space, and $\varXi$ be an arbitrary set in the dual space Y'. Let the function $f : K \to Y$ be such that all the functions $x \in K \to \langle \xi, f(x) \rangle \in \boldsymbol{R}$, $\xi \in \varXi$, are radially lsc along the rays starting at the point $x^0 \in ker\, K$. Then x^0 is a solution of the system of VI (6) with the chosen $\varXi$ if and only if with respect to the cone $C_\varXi$ it holds $f \in IAR^+(K, x^0)$. In consequence, any solution $x^0 \in ker\, K$ of (6) is an a-minimizer of f with respect to the cone $C_\varXi$.*

We shall now exploit some possibilities, for given closed convex cone C and linear scalarization.
If $\varXi \subset C'$, then $C \subset C_\varXi$ and $C'_\varXi \subset C'$. The system of VI is a subset of that defined at the beginning of the section.
Often in optimization happens to deal with the set $\varXi = \{\xi \in C' \mid \langle \xi, y^0 \rangle = 0\}$, where $y^0 \in C$. Then $C_\varXi$ is the contingent cone (see e.g. [1]) of C at y^0, at least when Y is a normed space.
It is worth to mention the case $\varXi = \{\xi^0\}$, $\xi^0 \in C$, a singleton. Then (6) reduces to the single equation (7) with $\phi(x) = \langle \xi^0, f(x) - f(x^0) \rangle$ to which Theorem 1 directly can be applied. If ϕ is radially lsc along the rays starting at x^0, then $x^0 \in ker\, K$ is a solution of (7) if and only if $\phi \in IAR(K, x^0)$. Consequently, x^0 is a minimizer of ϕ, and therefore $\langle \xi^0, f(x^0) \rangle = \min\{\langle \xi^0, f(x) \rangle \mid x \in K\}$. In the literature sometimes the points x^0 satisfying this condition are called linearized (through $\xi^0 \in C'$) efficient points.

We can also study the reversal of Theorem 6, through the scheme of Theorem 5. Now the core question is which kind of vector quasi-convexity lay behind $\varXi$-RQC. We see that now $f \in \varXi\text{-}RQC(K, x^0)$ at $x^0 \in ker\, K$ if all the functions

$$x \in K \to \langle \xi, f(x) - f(x^0) \rangle\,, \quad \xi \in C'\,, \tag{8}$$

are radially quasiconvex along the rays starting at x^0. The function f is $\varXi$-quasiconvex if K is convex and the functions (8) are quasiconvex for each $x^0 \in K$. Clearly, the latter is equivalent to require that functions

$$x \in K \to \langle \xi,\, f(x) \rangle\,, \quad \xi \in C'\,, \tag{9}$$

are (radially) quasiconvex. This generalized quasiconvexity was introduced in [28] as $*$-quasiconvexity. Further studies on the subject and detailed references can be found in [18, 42].

It becomes straightforward that the following version of Theorem 5.

Theorem 7. *Let K be a set in a real linear space and C be a closed convex cone in the locally convex space Y. Let the function $f : K \to Y$ be radially $*$-quasiconvex along the rays starting at $x^0 \in ker\, K$. If x^0 is an a-minimizer of f, then x^0 is a solution of the system of VI (6). In particular, if f is $*$-quasiconvex, then any a-minimizer of f is a solution of (6).*

We omit the easy proof, to focus on the study of the weaker assumption of C-quasiconvexity. Recall that a function $f : K \to Y$ is said to be C-quasiconvex on the convex set $K \subset X$ if for each $y \in Y$, the set $\{x \in K |\, f(x) \in$

$y - C\}$ is convex. Similarly, we call f radially C-quasiconvex along rays starting at $x^0 \in \ker K$, if the restriction of f on each such ray is C-quasiconvex. It is well known (see e.g. [28]), that the class of (radially) C-quasiconvex functions is broader than that of (radially) $*$-quasiconvex functions.

The following proposition, which is in fact Theorem 3.1 in [4], shows that, eventually diminishing the set Ξ, we still get equivalence of Ξ-quasiconvexity and C-quasiconvexity. We make use the following standard notions. The pair (Y, C) is directed if, for arbitrary $y^1, y^2 \in Y$, there exists $y \in Y$, such that $y - y^1 \in C$ and $y - y^2 \in C$. If Y is a Banach space, and the closed convex cone C has a nonempty interior, then the pair (Y, C) is directed. There are however important examples [4], of directed pairs in which $\operatorname{int} C = \emptyset$. Given a set $P \subset Y$, a point $x \in P$ is said to be an extreme point of P, when there exist no couple of points $x^1 \neq x^2 \in P$, such that x can be expressed as a convex combination of x^1 and x^2 with positive coefficients. Recall also that a vector $\xi \in C'$ is said to be an extreme direction of C' when $\xi \in C' \backslash \{0\}$ and for all $\xi^1, \xi^2 \in C'$ such that $\xi = \xi^1 + \xi^2$, there exist positive reals λ_1, λ_2 for which $\xi^1 = \lambda_1 \xi$, $\xi^2 = \lambda_2 \xi$. We denote by $\operatorname{ext} P$ the set of extreme points of P and by $\operatorname{extd} C'$ the set of extreme directions for C'.

Proposition 1 ([4]). *Let Y be a Banach space, and C be a closed convex cone in Y, such that the pair (Y, C) is directed. Suppose that C' is the weak-$*$ closed convex hull of* $\operatorname{extd} C'$. *Then, f is C-quasiconvex if and only if f is Ξ-quasiconvex with $\Xi = \operatorname{extd} C'$.*

Remark 5. A "radial variant" of this statement can be formulated straightforward.

Now, as an application of Theorem 5 and Proposition 1, we get the following result.

Corollary 2. *Let K be a set in a real linear space and Y be a Banach space. Let C be a closed convex cone in Y, such that (Y, C) is directed and C' has a weak-$*$ compact convex base Γ (in particular these assumptions hold when C is a closed convex cone with nonempty interior). Let the function $f : K \to Y$ be radially C-quasiconvex along the rays starting at $x^0 \in \ker K$. If x^0 is an a-minimizer of f, then x^0 is a solution of the system of VI (6). Moreover, if f is C-quasiconvex, then each a-minimizer of f is a solution of (6).*

Proof. Together with (6), with $\Xi = C'$, we consider the system of VI

$$\phi'(x, x^0 - x) \le 0\,, \quad x \in K\,, \quad \text{for all} \quad \phi \in \Phi(\Gamma \cap \operatorname{extd} C', x^0)\,. \tag{10}$$

The two systems are equivalent, in the sense that x^0 is a solution of (6) if and only if x^0 is a solution of (10). Indeed, since $\Gamma \cap \operatorname{extd} C' \subset C'$, we see that if x^0 is a solution of (10), then x^0 is a solution of (6). The reverse inclusion in true, since according to Krein-Milman Theorem, $C' = \operatorname{cl} \operatorname{cone} \operatorname{co} (\Gamma \cap \operatorname{extd} C')$, hence each inequality in the system (6) is a consequence of the inequalities in (10). Suppose that f is radially C-quasiconvex along the rays starting at $x^0 \in \ker K$. This assumption according to Proposition 1 is equivalent to the condition $f \in \Xi - RQC(K, x^0)$, with $\Xi = \Gamma \cap \operatorname{extd} C'$ (replacing $\operatorname{extd} C'$ with $\Gamma \cap \operatorname{extd} C'$ does give any change). Therefore, the condition that x^0 is an a-minimizer of f is equivalent to the condition that x^0 is a Ξ-minimizer (with $\Xi = \Gamma \cap \operatorname{extd} C'$). The Ξ-quasiconvexity of f and Theorem 5 give that x^0 is a solution of the system of VI (10). Finally, the equivalence of (10) and (6) gives that x^0 is a solution of the system of VI (6), with $\Xi = C'$.

4 *w*-Minimizers

Let Y be a normed space and C be a closed convex cone in Y. The dual space is also a normed space endowed with the norm $\|\xi\| = \sup\left(\langle \xi, y\rangle / \|y\|\right)$, $\xi \in Y'$. We choose now $\Xi = \{\xi^0\}$ to be a singleton with $\xi^0 : Y \to \boldsymbol{R}$ given by

$$\xi^0(y) = \sup\{\langle \xi, y\rangle \mid \xi \in C',\ \|\xi\| = 1\}\,. \tag{11}$$

In fact $\xi^0(y) = D(y, -C)$ is the so called oriented distance from the point y to the cone $-C$. The oriented distance $D(y, A)$ from a point $y \in Y$ to a set $A \subset Y$ is defined by $D(y, A) = d(y, A) - d(y, Y \setminus A)$. Here $d(y, A) = \inf\{\|y - a\| \mid a \in A\}$. The function D, introduced in [25, 26], has found various applications to optimization. It is shown in [23] that for a convex set A it holds $D(y, A) = \sup_{\|\xi\|=1}\left(\langle \xi, y\rangle - \sup_{a\in A}\langle \xi, a\rangle\right)$, which when C is a convex cone gives (11).

According to Remark 2 the system (6) is the single VI (7) with $\phi : K \to \boldsymbol{R}$ given by

$$\phi(x) = \xi^0(f(x) - f(x^0)) = D(f(x) - f(x^0), -C)\,. \tag{12}$$

Obviously, we have $\phi \in RLSC(K, x^0)$ provided that each of the functions

$$x \in K \to \langle \xi, f(x) - f(x^0)\rangle\,, \quad \xi \in C'\,, \tag{13}$$

has the same property. The property $\phi \in IAR(K, x^0)$ means geometrically that the oriented distance $D(f(x) - f(x^0), -C)$ is increasing along the rays starting at x^0. Further, x^0 is a minimizer of ϕ if and only if x^0 is a w-minimizer of f (see e.g. [22, 48]. Theorem 2 gives now the following result, related to VI considered in [12] and [15].

Theorem 8. *Let K be a set in a real linear space X, Y be a normed space, and $\Xi = \{\xi^0\}$ be a singleton with $\xi^0 : Y \to \boldsymbol{R}$ given by (11). Let the function $f : K \to Y$ be such that the functions (13) are radially lsc along the rays starting at $x^0 \in \ker K$. Then x^0 is a solution of VI (7) with ϕ given by (12) if and only if $\phi \in IAR(K, x^0)$, which means that the oriented distance $D(f(x) - f(x^0), -C)$ is increasing along the rays starting at x^0. In consequence, any solution $x^0 \in \ker K$ of (7) is a w-minimizer of f.*

Example 3. Let $X = \boldsymbol{R}$, $Y = \boldsymbol{R}^2$ and $K = \boldsymbol{R}_+$. The function $f(x) = \left(1, x^3\right)$ is differentiable. fix $x^0 = 0$, then

$$\phi(x) = \xi^0 \left(\left(1, x^3\right) - \left(1, 0\right)\right) = D\left(\left(0, x^3\right), -C\right)$$

which is $IAR(K, 0)$. Therefore, from Theorem 8, we gain $x^0 = 0$ is solution of VI system 1 and w-minimizer for f over K. This can be easily seen by direct calculations.

In [15] we refer to the $\Xi - IAR(K, x^0)$ property with $\Xi = \left\{\xi^0\right\}$, where ξ^0 is given by (11), as $VIAR(K, x^0)$ property. It is clear from (11) that ξ^0 depends on the norm chosen on Y, as shown by the next example, which we quote from [15] for the sake of completeness.

Example 4. i) Consider the function $f : \boldsymbol{R} \to \boldsymbol{R}^2$, defined as $f(x) = (x, g(x))$, where $g(x) = 2x$ if $x \in [0, 1]$ and $g(x) = -\frac{1}{4}x + \frac{9}{4}$ if $x \in (1, +\infty)$ and let $C = \boldsymbol{R}^2_+$, $K = \boldsymbol{R}_+$ and $x_0 = 0$. Then it is easy to show that function $f \in VIAR(K, x_0)$ if $\boldsymbol{R}^2$ is endowed with the Euclidean norm l^2, but $f \notin VIAR(K, x_0)$ if $\boldsymbol{R}^2$ is endowed with the norm l^∞.

ii) Consider the function $f : \boldsymbol{R}^2 \to \boldsymbol{R}^3$, defined as $f\left(x_1, x_2\right) = \left(x_1,\ g\left(x_1\right),\ 0\right)$, where $g(x)$ is defined in the previous point i). Let $C \subset \boldsymbol{R}^3$ be the polyhedral cone generated by the linearly independent vectors $\xi^1 = (-1/30,\ 1,\ 0)$, $\xi^2 = (1,\ 1/30,\ 0)$ and $\xi^3 = (0,\ 0,\ 1)$ (hence it is easily seen that $C' = C$). Let $K = \boldsymbol{R}^2_+$ and $x^0 = 0$. Then $f \in VIAR\left(K, x^0\right)$ if $\boldsymbol{R}^3$ is endowed with the Euclidean norm l^2, while $f \notin VIAR\left(K, x^0\right)$ if $\boldsymbol{R}^3$ is endowed by the l^∞_C norm.

We now move to explore the reversal of Theorem 8. Again it turns out that we can make a fruitful use $*$-quasiconvexity.

Corollary 3. *Let K be a set in a real linear space, Y be a normed space, C be a closed convex cone in the normed space Y, and $\Xi = \{\xi^0\}$ with ξ^0 given by (11). Let the function $f : K \to Y$ be radially $*$-quasiconvex along the rays starting at $x^0 \in ker\, K$. If x^0 is a w-minimizer of f, then x^0 is a solution of the system of VI (6). In particular, if f is $*$-quasiconvex, then any w-minimizer of f is a solution of (6).*

Proof. The proof is an immediate consequence of Theorem 5. It is enough to observe that if f is $*$-quasiconvex, then $f \in \Xi - RQC(K, x^0)$ with $\Xi = \{\xi^0\}$. This comes immediately, since $\phi(x) = \xi^0(f(x) - f(x^0)) = D(f(x) - f(x^0), -C)$ is the supremum of radially quasiconvex functions and hence is radially quasiconvex.

Since the class of (radially) C-quasiconvex functions is broader than that of (radially) $*$-quasiconvex functions, it arises naturally the question whether a result similar to Corollary 7 holds under radial C-quasiconvexity assumptions. We are going to show that in this case a result analogous to Corollary 7 holds when Y is a Banach space and its dual Y^* is endowed by a suitable norm, equivalent to the original one. From now on we assume that C is a closed convex cone in the normed space Y with both $\operatorname{int} C \neq 0$ and $\operatorname{int} C' \neq 0$. Fix $c \in \operatorname{int} C$. The set $G = \{\xi \in C' | \langle \xi, c\rangle = 1\}$ is a weak-$*$ compact convex base for C' [27]. Let $\tilde{B} = \{G \cup (-G)\}$. Since $\tilde{B}$ is a balanced, convex, absorbing and bounded set, with $0 \in \operatorname{int} \tilde{B}$ (here we apply $\operatorname{int} C' \neq \emptyset$), the Minkowsky functional $\gamma_{\tilde{B}}(y) = \{\lambda \in \boldsymbol{R} | \, \lambda > 0 \, , y \in \lambda\tilde{B}\}$ is a norm on Y^*, see e.g. [43, 27]. We denote this norm by $\|\cdot\|_1$. Since $\operatorname{int} \tilde{B} \neq \emptyset$ and $\tilde{B}$ is bounded, it is easily seen that the norm $\|\cdot\|_1$ is equivalent to the original norm $\|\cdot\|$ in Y^*.

Theorem 9. *Let K be a set in a linear space and Y be a normed space. Let the function $f : K \to Y$ be radially C-quasiconvex along the rays starting at $x^0 \in \ker K$ and assume Y^* is endowed with the norm $\|\cdot\|_1$. If x^0 is a w-minimizer of f, then $\phi(x) = D(f(x) - f(x^0), -C)$ is radially quasiconvex along the rays starting at x^0 (i.e. $f \in \Xi - RQC(K, x^0)$) and hence x^0 is a solution of the VI (7), with ϕ given by (12). In particular, if f is C-quasiconvex, then any w-minimizer of f is a solution of such VI.*

Proof. To prove the theorem it is enough to show that if f is radially C-quasiconvex and Y^* is endowed by the norm $\|\cdot\|_1$, then $\phi(x) = D(f(x) - f(x^0), -C)$ is radially C-quasiconvex along the rays starting at x^0. Indeed, we recall that

$$\phi(x) = \sup\{\langle \xi, f(x) - f(x^0)\rangle | \, \xi \in C', \|\xi\|_1 = 1\}\,, \tag{14}$$

and we observe that $\{\xi \in C' | \, \|\xi\|_1 = 1\} = G$. Hence the supremum in (14) is attained, since G is weak-$*$ compact, and, in particular, this happens at extreme points of G, which in turns are extreme vectors of C'. We have

$$\phi(x) = \max\{\langle \xi, f(x) - f(x^0)\rangle \, | \, \xi \in \operatorname{ext} G\}\,. \tag{15}$$

Due to Proposition 1, we get that ϕ^0 is the maximum of radially quasiconvex functions and hence is radially quasiconvex.

Next example shows that the previous theorem does not hold if Y^* is not endowed with the norm $\|\cdot\|_1$.

Example 5. Let $K \subset \boldsymbol{R} = [1/2, +\infty)$ and consider the function $f : K \to \boldsymbol{R}^2$, $f = (f_1, f_2)$, defined as $f_1(x) = -\frac{1}{2}x$, if $x \in [1/2, 1]$, $f_1(x) = -\frac{1}{2}x^3$, if $x \in (1, +\infty)$ and $f_2(x) = x$. Further, let $C = \boldsymbol{R}^2_+$. Clearly, here $Y = Y^* = \boldsymbol{R}^2$ and if we fix $(1, 1) \in \operatorname{int} C$, we obtain that the norm $\|\cdot\|_1$ coincides with the l^1 norm on $\boldsymbol{R}^2$, so that $\phi^0(x) = \max\{f_1(x), f_2(x)\}$ (see e.g. [15, 23]). Clearly, x^0 is a w minimizer and by Theorem 9, if $\boldsymbol{R}^2$ is endowed with the l^1 norm, x^0 is a solution of the VI (7), with ϕ given by (12), which it is also easily seen directly.
Assume now that $\boldsymbol{R}^2$ is endowed with a different norm, constructed as follows. Consider the set $A = \{(x_1, x_2) \in \boldsymbol{R}^2 | x_2 = -x_1 + 3, x_1 \in [1, 2]\}$ and let

$\tilde{A} = \mathrm{conv}\,(A \cup (-A))$. The Minkowsky functional of the set $\tilde{A}$ defines a norm on $\boldsymbol{R}^2$ and direct calculations show that when this norm is considered, $\phi^0 \notin IAR(K, x^0)$, so that x^0 does not solve the VI (7), with ϕ given by (12).

Let us underline, that in the previous theorem we proved in fact, that when Y^* is endowed with the norm $\|\cdot\|_1$, if x^0 is a w-minimizer of f, and f is C-quasiconvex along the rays starting at x^0, then $f \in \Xi\text{-}RQC(K, x^0)$. Easy examples show however, that the converse is not true.

We close this section with a remark on the choice of $\Xi = \{\xi^0\}$. Indeed, as a tool for w-minimizer, it is more common the Gerstewitz's function (see e.g. [7, 19, 24, 33, 34, 49], defined as

$$\tilde{\xi}\,(y) = \inf\,\{t \in \boldsymbol{R} \,|\, y \in tc - C\}\,,$$

where $c \in \mathrm{int}C$ is fixed. However the scalarizing function $\tilde{\xi}$ is just a special case of ξ^0.

Proposition 2. *If Y^* is endowed with the norm $\|\cdot\|_1$, then $\xi^0\,(y) = \tilde{\xi}(y)$, for all $y \in Y$.*

Proof. We have

$$\begin{aligned} \tilde{\xi}\,(y) &= \inf\,\{t \in \boldsymbol{R} \,|\, y - tc \in -C\} = \\ &= \inf\,\{t \in \boldsymbol{R} \,|\, \langle \xi, y - tc \rangle \le 0,\ \forall \xi \in C'\} = \\ &= \inf\,\{t \in \boldsymbol{R} \,|\, \langle \xi, y \rangle - t \le 0,\ \forall \xi \in C' \cap S\} = \\ &= \inf\,\{t \in \boldsymbol{R} \,|\, t \ge \langle \xi, y - tc \rangle,\ \forall \xi \in C'\} = \\ &= \sup\,\{\langle \xi, y \rangle, \xi \in C' \cap S\}\,. \end{aligned}$$

5 *p*-Minimizers

In this section for the sake of simplicity we assume Y is a finite dimensional normed space and C is a closed convex pointed cone with nonempty interior. However, the results can be easily extended to the infinite dimensional case. We recall that, besides the notions of efficiency presented in the Section 2, in problem (5) it is widely studied also that of proper efficiency. A point $x^0 \in K$ is a proper efficient point (p-minimizer) for $f : K \to Y$ when there exists a closed convex and pointed cone $\tilde{C}$ with $C \setminus \{0\} \subset \mathrm{int}\,\tilde{C}$ such that x^0 is a w-minimizer with respect to $\tilde{C}$. Let $\bar{a} = \sup\{a > 0 \,|\, \exists y \in C' \cap S\,, D(y, C') \le -a\}$ and for a given $a \in (0, \bar{a})$ consider the set

$$A(a) = \{\xi \in C' \,|\, D(\xi, C') \le -a\|\xi\|\}\,.$$

We easily have that $A(a)$ is a nonempty closed pointed cone with $A(a) \subseteq \mathrm{int}\,C'$ (observe $A(a)$ is not necessarily convex). We choose $\Xi = \{\xi^1\}$ to be a singleton with

$$\xi^1(y) = \max\{\langle \xi, y\rangle | \xi \in A(a) \cap S\}.$$

Taking into account that $\xi^1(y) = \max\{\langle \xi, y\rangle | \xi \in \operatorname{conv}(A(a) \cap S)\} = \max\{\langle \xi, y\rangle | \xi \in \operatorname{conv} A(a) \cap S\}$ we obtain

$$\xi^1(y) = D(y, -\tilde{C}(a)), \tag{16}$$

where $\tilde{C}(a) = [\operatorname{cone}\operatorname{conv}(A(a) \cap S)]'$ and $C \backslash \{0\} \subseteq \operatorname{int} \tilde{C}(a)$. Hence the point x^0 is a p-minimizer of f if and only if there exists a positive number a such that x^0 is a minimizer of

$$\phi(x) = \xi^1(f(x) - f(x^0)) \tag{17}$$

Therefore, next result follows as a rephrasing of Theorem 8 with respect to the cone $\tilde{C}(a)$.

Theorem 10. *Let K be a set in a real linear space X, Y be a finite dimensional normed space, and $\Xi = \{\xi^1\}$ be a singleton with $\xi^1 : Y \to \boldsymbol{R}$ given by (16) for some positive number a. Let the function $f : K \to Y$ be such that the functions (13) are radially lsc along the rays starting at $x^0 \in \ker K$. Then if x^0 is a solution of VI (7) with ϕ given by (17), the oriented distance $D(f(x) - f(x^0), -\tilde{C}(a))$ is increasing along the rays starting at x^0. In consequence, any solution $x^0 \in \ker K$ of (7) is a p-minimizer of f.*

The following result gives a reversal of Theorem 10 under $*$-quasiconvexity assumptions. We omit the easy proof.

Theorem 11. *Let K be a set in a real linear space, Y be a finite dimensional normed space, and $\Xi = \{\xi^1\}$ with ξ^1 given by (16). Let the function $f : K \to Y$ be radially $*$-quasiconvex along the rays starting at $x^0 \in \ker K$. If x^0 is a p-minimizer of f, then there exists a positive number a such that x^0 is a solution of the system of VI (6).*

6 Conclusions

In this paper we present a general scheme to define Minty variational inequalities to study vector optimization. The guideline is the Minty variational principle for the scalar case, as deepened in [13]. Other researches have introduced Minty vector variational inequalities. Here we follow the suggestion in [12] to associate to a primitive vector optimization problem suitable scalar inequalities. The key of such an approach is scalarization. Despite the Minty inequality which comes from our Theorem 2 may look too abstract, we show it can be related even to vector Minty inequalities. However we have the advantage that some gap which was in [21] is filled by this formulation.

We trust this result may be a reference for new researches in the field of differentiable variational inequalities.

References

1. Aubin JP Frankowska H (1990) Set-valued analysis. Birkhäuser, Boston
2. Aussel D Hadjisavvas N (2004) On quasimonotone variational inequalities J. Optim. Theory Appl. 121 no. 2:445–450
3. Baiocchi C Capelo A (1984) Variational and Quasivariational Inequalities, Applications to Free-Boundary Problems. John Wiley & Sons, New York
4. Benoist J Borwein JM Popovici N (2002) A characterization of quasiconvex vector-valued functions Proc. Amer. Math. Soc. 131 no. 4:1109-1113.
5. Chen GY Cheng GM (1987) Vector variational inequality and vector optimization. In: Lecture notes in Economics and Mathematical System 285. Springer–Verlaag, Berlin:408–416
6. Chen GY Craven BD (1990) A vector variational inequality and optimization over an efficient set Z. Oper. Res. 34:1–12
7. Chen GY Goh CJ Yang XQ (1999) Vector network equilibrium problems and nonlinear scalarization methods Math. Methods Oper. Res. 49 no. 2:239– 253
8. Chen GY Huang XX and Yang XQ (2005) Vector optimization, Set-valued and variational analysis. In: Lecture Notes in Economics and Mathematical Systems, 541, Springer-Verlag, Berlin
9. Crespi GP (2002) Proper efficiency and vector variational inequalities J. Inform. Optim. Sci. 23 no. 1:49–62
10. Crespi GP Ginchev I Rocca M (2004) Minty vector variational inequality, efficiency and proper efficiency Vietnam J. Math. 32 no. 1:95–107
11. Crespi GP Ginchev I Rocca M (2004) Minty variational inequalities, increase along rays property and optimization J. Optim. Theory Appl. 123 no. 3:479–496
12. Crespi GP Ginchev I Rocca M (2005) Variational inequalities in vector optimization. In: Giannessi F, Maugeri A (eds) Variational Analysis and Applications, Proc. Erice, June 20– July 1 2003. Springer, New York, 2005
13. Crespi GP Ginchev I Rocca M (2005) Existence of solutions and star-shapedness in Minty variational inequalities J. Global Optim. 32 no. 4:485–494
14. Crespi GP Ginchev I Rocca M (2006) A note on Minty type vector variational inequalities RAIRO Oper. Res.:in press
15. Crespi GP Ginchev I Rocca M (2006) Increasing-along-rays property for vector functions J. Nonlinear Convex Anal. 7 no. 6:in press
16. Crespi GP Guerraggio A Rocca M (2005) Minty variational inequality and optimization: scalar and vector case. In: Generalized convexity, generalized monotonicity and applications. Nonconvex Optim. Appl. 77 Springer, New York:193–211
17. Daniilidis A Hadjisavvas N (1999) Coercivity conditions and variational inequalities Math. Program. 86 no. 2 Ser. A:433–438
18. Flores-Bazán F Hadjisavvas N Vera C (to appear) An optimal alternative theorem and applications to matheamtical programming J. Global Optim.
19. Gerstewitz C (1986) Nichtkonvexe Trennungssätze und deren Anwendung in der Theorie der Vektoroptimierung. In:17. Jahrestagung “Mathematische Optimierung” (Sellin, 1984) (German) Seminarberichte 80 Humboldt Univ. Berlin:19–31
20. Giannessi F (1980) Theorems of alternative, quadratic programs and complementarity problems. In: Cottle RW, Giannessi F, Lions JL (eds.) Variational Inequalities and Complementarity Problems John Wiley & Sons, New York:151–186

21. Giannessi F (1998) On Minty variational principle. In: Giannessi F, Komlósi S, Rapcsák T (eds.) New Trends in Mathematical Programming, Kluwer, Dordrecht:93–99
22. Ginchev I Guerraggio A Rocca M (2005) Isolated minimizers and proper efficiency for $C^{0,1}$ constrained vector optimization problems J. Math. Anal. Appl. 309:353–368
23. Ginchev I Hoffmann A (2002) Approximation of set-valued functions by single-valued one Discuss. Math. Differ. Incl. Control Optim. 22:33–66
24. Goh CJ Yang XQ (2002) Duality in optimization and variational inequalities. Optimization Theory and Applications, 2, Taylor & Francis Ltd., London
25. Hiriart-Urruty JB (1979) New concepts in nondifferentiable programming, Analyse non convexe Bull. Soc. Math. France 60:57–85
26. Hiriart-Urruty JB (1979) Tangent cones, generalized gradients and mathematical programming in Banach spaces Math. Oper. Res. 4:79–97
27. Jameson G (1970) Ordered linear spaces. Lecture Notes in Mathematics, Vol. 141, Springer, Berlin
28. Jeyakumar V Oettly W Natividad V (1993) A solvability theorem for a class of quasiconvex mappings with applications to optimization J. Math. Anal. Appl. 179:537–546
29. Kinderlehrer D Stampacchia G (1980) An Introduction to Variational Inequalities and Their Applications. Academic Press, New York
30. Komlósi S (1999) On the Stampacchia and Minty variational inequalities. In: Giorgi G, Rossi F (eds) Generalized Convexity and Optimization for Economic and Financial Decisions, Proc. Verona, Italy, May 28–29 1998. Pitagora Editrice, Bologna:231–260.
31. Lee GM Kim DS Lee BS Cho SJ (1996) On vector variational inequality Bull. Korean Math. Soc. 33 no. 4:553–564.
32. Lee GM Kim DS Lee BS Yen ND (2000) Vector Variational Inequality as a tool for studying vector optimization Problems. In: Vector variational inequalities and vector equilibria. Nonconvex Optim. Appl. 38 Kluwer Acad. Publ. Dordrecht:277–305
33. Li SJ Yang XQ Chen GY (2003) Nonconvex vector optimization of set-valued mappings J. Math. Anal. Appl. 283 no.2:337– 350
34. Li SJ Chen GY Teo KL Yang XQ (2003) Generalized minimax inequalities for set-valued mappings J. Math. Anal. Appl. 281 no.2:707– 723
35. Luc DT (1989) Theory of Vector Optimization. Springer Verlag, Berlin
36. Mastroeni G (1999) Some remarks on the role of generalized convexity in the theory of variational inequalities. In: Giorgi G, Rossi F (eds) Generalized Convexity and Optimization for Economic and Financial Decisions, Proc. Verona, Italy, May 28–29 1998. Pitagora Editrice, Bologna:271–281
37. Minty GJ (1967) On the generalization of a direct method of the calculus of variations Bull. Amer. Math. Soc. 73:314–321
38. Mordukhovich B (1994) Stability theory for parametric generalized equations and variational inequalities via nonsmooth analysis Trans. Amer. Math. Soc. 343 no. 2:609–657
39. Rubinov AM (2000) Abstract Convexity and Global Optimization. Kluwer Dordrecht
40. Schaefer HH (1966) Topological Vector Spaces. The MacMillan Company, New York, London

41. Stampacchia G (1960) Formes bilinéaires coercitives sur les ensembles convexes C. R. Acad. Sci. Paris (Groupe 1) 258:4413–4416
42. Tanaka T (1994) Generalized quasiconvexity, cones saddle points and minimax theorem for vector-valued functions J. Optim. Theory Appl. 81:355-377
43. Taylor AE Lay DC (1980) Introduction to functional analysis. Second edition. John Wiley & Sons, New York-Chichester-Brisbane
44. Thach PT Kojima M (1996) A generalized convexity and variational inequality for quasi-convex minimization SIAM J. Optim. 6 no. 1:212–226
45. Yang XQ (1993) Generalized convex functions and vector variational inequalities J. Optim. Theory Appl. 79 no. 3:563–580
46. Yang XQ (1997) Vector Variational Inequality and Vector Pseudolinear Optimization J. Optim. Theory Appl. 95 no.3:729–734
47. Yang XQ Goh CJ (1997) On vector variational inequalities: applications to vector equilibria J. Optim. Theory Appl. 95 no.2:431–443.
48. Zaffaroni A (2003) Degrees of efficiency and degrees of minimality SIAM J. Control Optim. 42:1071–1086.
49. Zălinescu C (1987) On two notions of proper efficiency. In: Optimization in mathematical physics (Oberwolfach, 1985) Methoden Verfahren Math. Phys. 34 Lang Frankfurt am Main:77–86

Higher Order Properly Efficient Points in Vector Optimization

Ivan Ginchev[1], Angelo Guerraggio[2], and Matteo Rocca[3]

[1] Technical University of Varna, Department of Mathematics, Studentska Str. 9010, Varna, Bulgaria, iginchev@yahoo.com
[2] University of Insubria, Department of Economics, via Monte Generoso, 71, 21100, Varese, Italy, aguerraggio@eco.uninsubria.it
[3] University of Insubria, Department of Economics, via Monte Generoso, 71, 21100, Varese, Italy, mrocca@eco.uninsubria.it

Summary. We consider the constrained vector optimization problem $\min_C f(x)$, $g(x) \in -K$, where $f : \mathbb{R}^n \to \mathbb{R}^m$ and $g : \mathbb{R}^n \to \mathbb{R}^p$ are given functions and $C \subseteq \mathbb{R}^m$ and $K \subseteq \mathbb{R}^p$ are closed convex cones. Two type of solutions are important for our considerations, namely i-minimizers (isolated minimizers) of order k and p-minimizers (properly efficient points) of order k (see e.g. [11]). Every i-minimizer of order $k \geq 1$ is a p-minimizer of order k. For $k = 1$, conditions under which the reversal of this statement holds have been given in [11]. In this paper we investigate the possible reversal of the implication i-minimizer $\Longrightarrow$ p-minimizer in the case $k = 2$. To carry on this study, we develop second-order optimality conditions for p-minimizers, expressed by means of Dini derivatives. Together with the optimality conditions obtained in [13] and [12] in the case of i-minimizers, they play a crucial role in the investigation. Further, to get a satisfactory answer to the posed reversal problem, we deal with sense I and sense II solution concepts, as defined in [11] and [5].

Key words: Vector optimization, locally Lipschitz data, properly efficient points, isolated minimizers, optimality conditions.

1 Introduction

In this paper we consider the vector optimization problem

$$\min_C f(x)\,, \quad g(x) \in -K, \tag{1}$$

where $f : \mathbb{R}^n \to \mathbb{R}^m$, $g : \mathbb{R}^n \to \mathbb{R}^p$. Here n, m and p are positive integers and $C \subseteq \mathbb{R}^m$ and $K \subseteq \mathbb{R}^p$ are closed convex cones.

Usually the solutions of problem (1) are called points of efficiency. We prefer, as in scalar optimization, to call them minimizers. In particular we call w-minimizers the weakly efficient points of problem (1).

We focus on two kinds of solutions of problem (1), introduced in [8] and [11], namely those of isolated minimizer (i-minimizer) and properly efficient point (p-minimizer) of a given order $k \geq 1$. In particular, the concept of i-minimizer extends to the vector case a notion known in scalar optimization [3], [25], [28], while p-minimizers of order k generalize the classical notion of properly efficient point [19]. Further generalizations of these concepts to set-valued optimization are given in [6]. As it will be clear from the definitions, p-minimizers give informations on the behavior of the image of the objective function f, while i-minimizers involve an interplay between the domain and the image of f and hence give a more complete picture for the problem data near the solution. These features are strengthened by the study of the stability properties of p-minimizers and i-minimizers of a given order (see [5], [9]). Indeed, while p-minimizers of a given order are stable under perturbations of the ordering cone, i-minimizers show stability with respect both to the ordering cone and the objective function f. Further, the i-minimizers are the most appropriate notion of a solution when one has to deal with optimality conditions. Indeed, for such solutions, the necessary conditions appear (changing the weak inequalities to strict ones) to be also sufficient [10], [13], [12].
At this point it is not surprising that, when f is of class $C^{0,1}$, every i-minimizer of a given order is also a p-minimizer of the same order [11]. Recall that a function is said to be of class $C^{k,1}$ when it is k times Fréchet differentiable, with locally Lipschitz k-th derivative. The $C^{0,1}$ functions are the locally Lipschitz functions.
Further, from the considerations above, arises naturally the question under what conditions the implication i-minimizer $\Longrightarrow$ p-minimizer admits a reversal, that is under what conditions a p-minimizer is also an i-minimizer.
A satisfactory answer to the problem of comparison of p-minimizers and i-minimizers of first order is given in [11]. A key role in this comparison is played by the first order optimality conditions given in [10] for $C^{0,1}$ functions.

In this paper, we concentrate on the comparison between the second order notions. Similarly to the first order case some optimality conditions are crucial in this investigation. For this reason we establish second order optimality conditions for p-minimizers of order 2, which, together with those given in [13] and [12] for the case of i-minimizers, constitute the main tool for the comparison. In order to achieve a unified treatment of necessary optimality conditions for p-minimizers, we prove also some first order optimality conditions.
Finally, we show that as in [11], our investigation leads us to relate to the original constrained problem an unconstrained one and to speak about sense I and sense II optimality concepts (i.e. respectively those of the original constrained problem and those of the associated unconstrained one). Relations between sense I and sense II concepts are obtained.

The outline of the paper is the following. Section 2 is devoted to notions of optimality for problem (1) and their scalarization. In Section 3 we prove necessary optimality conditions for p-minimizers of first and second order, respectively in the case of $C^{0,1}$ and $C^{1,1}$ data. The given first order necessary conditions extend the classical Kuhn-Tucker optimality conditions for properly efficient points (see e.g. [24]). In section 4 we show that the first order necessary conditions for a p-minimizer become sufficient under convexity assumptions on the functions f and g. In Section 5, as a consequence of the obtained optimality conditions we discuss the reversal of the implication i-minimizer $\Longrightarrow$ p-minimizer, in the case $k = 2$. In section 6 we show that a satisfactory solution to the reversal problem leads to consider sense I and sense II concepts.

2 Vector Optimality Concepts and Scalar Characterizations

We denote by $\|\cdot\|$ and $\langle\cdot,\cdot\rangle$ the Euclidean norm and the scalar product in the considered finite-dimensional spaces. The open unit ball is denoted by B, while S stands for the unit sphere. From the context it should be clear to which spaces these notations are applied. The results of the paper can be immediately extended to finite dimensional real Banach spaces.

There are different concepts of solution for problem (1). In the following definitions we assume that the considered point x^0 is feasible, i.e. $g(x^0) \in -K$ (equivalently $x^0 \in g^{-1}(-K)$). The definitions below are given in a local sense. We omit this specification in the text.

The feasible point x^0 is said to be weakly efficient (efficient), if there is a neighborhood U of x^0, such that if $x \in U \cap g^{-1}(-K)$ then $f(x) - f(x^0) \notin -\operatorname{int} C$ (respectively $f(x) - f(x^0) \notin -(C\backslash\{0\})$).

In this paper the weakly efficient and the efficient points of problem (1) are called respectively w-minimizers and e-minimizers.

We say that the feasible point x^0 is a strong e-minimizer if there exists a neighborhood U of x^0, such that $f(x) - f(x^0) \notin -C$, for $x \in U\backslash\{x^0\} \cap g^{-1}(-K)$.

The unconstrained problem

$$\min\nolimits_C f(x)\,, \quad x \in \mathbb{R}^n, \tag{2}$$

is a particular case of problem (1) and the defined notions of optimality concern also this problem.

For the cone $M \subseteq \mathbb{R}^k$ its positive polar cone M' is defined by $M' = \{\zeta \in \mathbb{R}^k | \, \langle\zeta,\, \phi\rangle \geq 0 \text{ for all } \phi \in M\}$. The cone M' is closed and convex and it is well known that $M'' := (M')' = \operatorname{cl}\operatorname{conv}\operatorname{cone} M$, see e. g. [23, Chapter III, § 15]. Here $\operatorname{cone} A$ stands for the cone generated by the set A. In particular, for a closed convex cone M we have $M = M'' = \{\phi \in \mathbb{R}^k | \, \langle\zeta,\, \phi\rangle \geq 0 \text{ for all } \zeta \in M'\}$.

If $\phi \in \operatorname{cl}\operatorname{cone} M$, then $\langle\zeta,\, \phi\rangle \geq 0$ for all $\zeta \in M'$. We set $M'[\phi] = \{\zeta \in M' | \, \langle\zeta,\, \phi\rangle = 0\}$. Then $M'[\phi]$ is a closed convex cone and $M'[\phi] \subseteq M'$. Consequently its positive polar cone $M[\phi] := (M'[\phi])'$ is a closed convex cone,

$M \subseteq M[\phi]$ and its positive polar cone satisfies $(M[\phi])' = M'[\phi]$. In fact it can be shown that $M[\phi]$ is the contingent cone [2] of M at ϕ. In this paper we apply the notation $M[\phi]$ for $M = K$ and $\phi = -g(x^0)$. In other words we will denote by $K[-g(x^0)]$ the contingent cone of K at $-g(x^0)$.

The solutions of problem (1) can be characterized in terms of a suitable scalarization. Given a set $A \subseteq \mathbb{R}^k$, then the distance from $y \in \mathbb{R}^k$ to A is given by $d(y, A) = \inf\{\|a - y\| \mid a \in A\}$. This definition can be applied also to the set $A = \emptyset$ with the agreement $d(y, \emptyset) = \inf \emptyset = +\infty$. The oriented distance from y to A is defined by $D(y, A) = d(y, A) - d(y, \mathbb{R}^k \setminus A)$. This definition gives $D(y, A) = +\infty$ when $A = \emptyset$ and $D(y, A) = -\infty$ when $A = \mathbb{R}^k$.

The function D is introduced in Hiriart-Urruty [17], [18] and is used later in Amahroq, Taa [1], Ciligot-Travain [4], Miglierina, Molho [21], Miglierina, Molho, Rocca [22]. Zaffaroni [29] gives different notions of efficiency and uses the function D for their scalarization and comparison. Ginchev, Hoffmann [15] use the oriented distance to study approximation of set-valued functions by single-valued ones and in the case of a convex cone C prove the representation $D(y, -C) = \sup_{\|\xi\|=1,\, \xi \in C'} \langle \xi,\, y \rangle$. Turn attention that this formula works also in the case of the improper cones $C = \{0\}$ (then $D(y, -C) = \sup_{\|\xi\|=1} \langle \xi, y \rangle = \|y\|$) and $C = \mathbb{R}^m$ (then $D(y, -C) = \sup_{\xi \in \emptyset} \langle \xi, y \rangle = -\infty$)).

Proposition 1 ([10]). *The feasible point $x^0 \in \mathbb{R}^n$ is a w-minimizer for problem (1), if and only if there exists a neighborhood U of x^0, such that $D(f(x) - f(x^0), -C) \geq 0$, $\forall x \in U \cap g^{-1}(-K)$.*

Proposition 2 ([10]). *The feasible point x^0 is a strong e-minimizer of problem (1) if and only if there exists a neighborhood U of x^0, such that $D(f(x) - f(x^0), -C) > 0$, $\forall x \in U \cap g^{-1}(-K)$, $x \neq x^0$.*

The concept of an isolated minimizer for scalar problems has been popularized by Auslender [3]. It is natural to introduce a similar concept of optimality for the vector problem (1) (see e.g. [13], [11], [12]).

Definition 1. *We say that the feasible point x^0 is an isolated minimizer (i-minimizer) of order k for the vector problem (1) if and only if there exists a neighborhood U of x^0 and a positive number A, such that $D(f(x) - f(x^0), -C) \geq A\|x - x^0\|^k$, $\forall x \in U \cap g^{-1}(-K)$.*

When $m = 1$ the notion of the isolated minimizer of order k from Definition 1 coincides with the known in the scalar case notion of an isolated minimizer of order k (see [3]). This remark justifies the importance of the oriented distance D for vector optimization problems.

Applying the oriented distance function we can generalize also the concept of proper efficiency. We recall that when C is a closed convex pointed cone, the feasible point x^0 is said to be properly efficient (in the sense of Henig) for problem (1), when there exists a closed convex pointed cone $\tilde{C}$, with $C \setminus \{0\} \subset \operatorname{int} \tilde{C}$, such that x^0 is w-minimizer with respect to $\tilde{C}$ [16].

For given $k \geq 1$ and $a > 0$, we define the set

$$C^k(a) = \{y \in \mathbb{R}^m \mid D(y, C) \leq a \|y\|^k\}.$$

It is easily seen that when $k = 1$ the set $C^1(a)$ is a closed cone (not necessarily convex, see e.g. [6]).

Definition 2. *We say that the feasible point x^0 is a properly efficient point (p-minimizer) of order $k \geq 1$ for problem (1) if there exist a neighborhood U of x^0 and a constant $a > 0$ such that if $x \in U \cap g^{-1}(-K)$ then $f(x) - f(x^0) \notin -\operatorname{int} C^k(a)$.*

In [14] it has been proved that when C is a closed convex pointed cone, then p-minimizers of first order are just properly efficient points in the sense of Henig. In the same paper, the Geoffrion characterization [7] of properly efficient points has been generalized to p-minimizers of higher order.

Although the notions of i-minimizer and p-minimizer are defined through the norms, it can be shown [12] that in fact they are norm independent, as a consequence of the equivalence of any two norms in a finite dimensional space over the reals. As a consequence of this, when $C = \mathbb{R}^m_+$, we can replace the function $D(y, -C)$ with the function $\phi(y) = \max\{y_1, \cdots, y_m\}$.

When f is a $C^{0,1}$ function, the following relation holds between i-minimizers of order k and p-minimizers of order k.

Theorem 1 ([11]). *Let f be of class $C^{0,1}$. If a point x^0 is an i-minimizer of order $k \geq 1$ for problem (1) then x^0 is a p-minimizer of order k.*

3 First and Second Order Optimality Conditions for *p*-Minimizers

In [10], [13], [12], we obtained first and second order optimality conditions for w-minimizers and i-minimizers of order 1 and 2, given in terms of suitable Dini directional derivatives for $C^{0,1}$ or $C^{1,1}$ functions. In this section we give similar necessary optimality conditions for p-minimizers of first and second order.

Given a $C^{0,1}$ function $\Phi : \mathbb{R}^n \to \mathbb{R}^k$ we define the Dini directional derivative (we use to say just Dini derivative) $\Phi'_u(x^0)$ of Φ at x^0 in direction $u \in \mathbb{R}^n$, as the set of the cluster points of $(1/t)(\Phi(x^0 + tu) - \Phi(x^0))$ as $t \to 0^+$, that is as the Kuratowski limit

$$\Phi'_u(x^0) = \operatorname*{Limsup}_{t \to 0^+} \frac{1}{t} \left(\Phi(x^0 + tu) - \Phi(x^0)\right).$$

It can be shown (see e.g. [10]) that if Φ is of class $C^{0,1}$, then $\Phi'_u(x^0)$ is a nonempty compact subset of $\mathbb{R}^k$ for all $u \in \mathbb{R}^n$.

In connection with problem (1) we deal with the Dini directional derivative of the function $\Phi : \mathbb{R}^n \to \mathbb{R}^{m+p}$, $\Phi(x) = (f(x),\, g(x))$ and then we use to write $\Phi'_u(x^0) = (f,\, g)'_u(x^0)$. If at least one of the derivatives $f'_u(x^0)$ and $g'_u(x^0)$ is a singleton, then $(f, g)'_u(x^0) = f'_u(x^0) \times g'_u(x^0)$. Let us turn attention that always $(f,\, g)'_u(x^0) \subseteq f'_u(x^0) \times g'_u(x^0)$, but in general these two sets do not coincide.

Given a $C^{1,1}$ function $\Phi : \mathbb{R}^n \to \mathbb{R}^k$, we define the second-order Dini directional derivative $\Phi''_u(x^0)$ of Φ at x^0 in direction $u \in \mathbb{R}^n$ as the set of the cluster points of $(2/t^2)(\Phi(x^0 + tu) - \Phi(x^0) - \Phi'(x^0)u)$ as $t \to 0^+$, that is as the upper limit

$$\Phi''_u(x^0) = \operatorname*{Limsup}_{t \to 0^+} \frac{2}{t^2} \left(\Phi(x^0 + tu) - \Phi(x^0) - \Phi'(x^0)u\right) .$$

If Φ is twice Fréchet differentiable at x^0 then the Dini derivative is a singleton and can be expressed in terms of the Hessian $\Phi''_u(x^0) = \Phi''(x^0)(u, u)$. In connection with problem (1) we deal with the Dini derivative of the function $\Phi : \mathbb{R}^n \to \mathbb{R}^{m+p}$, $\Phi(x) = (f(x),\, g(x))$. Then we use the notation $\Phi''_u(x^0) = (f,\, g)''_u(x^0)$. Let us turn attention that always $(f,\, g)''_u(x^0) \subset f''_u(x^0) \times g''_u(x^0)$, but in general these two sets do not coincide. In the sequel we need the next result [10], [12].

Proposition 3. *i) Let* $\Phi : \mathbb{R}^n \to \mathbb{R}^k$ *be Lipschitz with constant* L *in* $x^0 + r\,\mathrm{cl}\,B$, *where* $x^0 \in \mathbb{R}^n$ *and* $r > 0$. *Then for* $u,\, v \in \mathbb{R}^n$ *and* $0 < t < r/\max(\|u\|,\, \|v\|)$ *it holds*

$$\left\| \frac{1}{t} \left(\Phi(x^0 + tv) - \Phi(x^0)\right) - \frac{1}{t} \left(\Phi(x^0 + tu) - \Phi(x^0)\right) \right\| \le L\, \|v - u\| \,, \tag{3}$$

In particular for $v = 0$ *and* $0 < t < r/\|u\|$ *we get*

$$\left\| \frac{1}{t} \left(\Phi(x^0 + tu) - \Phi(x^0)\right) \right\| \le L\, \|u\| \,. \tag{4}$$

ii) Let $\Phi : \mathbb{R}^n \to \mathbb{R}^k$ *be a* $C^{1,1}$ *function and* Φ' *be Lipschitz with constant* L *on the ball* $\{x \mid \|x - x^0\| \le r\}$, *where* $x^0 \in \mathbb{R}^n$ *and* $r > 0$. *Then, for* $u, v \in \mathbb{R}^m$ *and* $0 < t < r/\max(\|u\|, \|v\|)$ *we have*

$$\left\| \frac{2}{t^2} \left(\Phi(x^0 + tv) - \Phi(x^0) - t\Phi'(x^0)v\right) - \frac{2}{t^2} \left(\Phi(x^0 + tu) - \Phi(x^0) - t\Phi'(x^0)u\right) \right\|$$
$$\le L\,(\|u\| + \|v\|)\, \|v - u\| \,.$$

In particular, for $v = 0$ *we get*

$$\left\| \frac{2}{t^2} \left(\Phi(x^0 + tu) - \Phi(x^0) - t\Phi'(x^0)u\right) \right\| \le L\, \|u\|^2 \,.$$

In the formulation of Theorem 2 below we use the following constraint qualification, which is a generalization for $C^{0,1}$ constraints of the well known Kuhn-Tucker constraint qualification (compare with Mangasarian [20, page 102]).

$$\mathcal{Q}_{0,1}(x^0): \begin{array}{c} \text{If } g(x^0) \in -K \text{ and } \frac{1}{t_k}\big(g(x^0+t_k u^0) - g(x^0)\big) \to z^0 \in -K[-g(x^0)] \\ \text{then } \exists\, u^k \to u^0 : \exists\, k_0 \in \mathbb{N} : \forall k > k_0 : g(x^0 + t_k u^k) \in -K\,. \end{array}$$

Theorem 2. *Assume that in the constrained problem (1) f and g are $C^{0,1}$ functions. If x^0 is a p-minimizer of first order and the constraint qualification $\mathcal{Q}_{0,1}(x^0)$ holds, then for each $u \in \mathbb{R}^n\backslash\{0\}$ and for every $(y^0, z^0) \in (f,g)'_u(x^0)$, we have*

$$\mathcal{NP}'_{0,1}: \qquad (y^0, z^0) \notin -\big(C\backslash\{0\} \times K[-g(x^0)]\big)\,.$$

Proof. Let $(y^0, z^0) \in (f,g)'_u(x^0)$, that is

$$(y^0, z^0) = \lim \frac{1}{t_\nu}\big(f(x^0+t_\nu u) - f(x^0),\, g(x^0+t_\nu u) - g(x^0)\big)$$

for some sequence $t_\nu \to 0+$. If $z^0 \notin -K[-g(x^0)]$, then clearly $(y^0, z^0) \notin -\big(C\backslash\{0\} \times K[-g(x^0)]\big)$. Assume now that $z^0 \in -K[-g(x^0)]$. Since the constraint qualification $\mathcal{Q}_{0,1}(x^0)$ holds, there exist a sequence $u^\nu \to u$, such that $g(x^0 + t_\nu u^\nu) \in -K$. Since f is of class $C^{0,1}$, using Proposition 3, we can assume

$$\frac{1}{t_\nu}(f(x^0+t_\nu u^\nu) - f(x^0)) \to y^0 \in f'_u(x^0)\,,$$

and since x^0 is p-minimizer, for ν large enough we have

$$D(f(x^0+t_\nu u^\nu) - f(x^0), -C) \ge a\|f(x^0+t_\nu u^\nu) - f(x^0)\|\,,$$

for some positive number a. Dividing by t_ν and passing to the limit we get $D(y^0, -C) \ge a\|y^0\|$, whence $y^0 \notin (-C\backslash\{0\})$.

The conclusion of Theorem 2 holds also obviously for $u = 0$, then $y^0 = 0 \notin -(C \setminus \{0\})$ and $z^0 = 0$. This trivial case is however not interesting for the next Theorem 3 stating necessary optimality conditions in dual form.

Theorem 3. *Let in problem (1) f and g be $C^{0,1}$ functions and assume that the constraint qualification $\mathcal{Q}_{0,1}(x^0)$ holds. If x^0 is a p-minimizer of first order for problem (1), then for every $u \in \mathbb{R}^n \setminus \{0\}$ and for every couple $(y^0, z^0) \in (f,g)'_u(x^0)$, such that $y^0 \neq 0$ whenever $z^0 \in -K[-g(x^0)]$, one can find $(\xi^0, \eta^0) \in C' \times K'[-g(x^0)]$, $(\xi^0, \eta^0) \neq (0,0)$, such that*

$$\mathcal{ND}'_{0,1}: \qquad \langle \xi^0, y^0\rangle + \langle \eta^0, z^0\rangle > 0.$$

Proof. From Theorem 2 we know that $(y^0, z^0) \notin -\big(C\backslash\{0\} \times K[-g(x^0)]\big)$. If $z^0 \notin -K[-g(x^0)]$, then one can find an element $\eta^0 \in K'[-g(x^0)]$, $\eta^0 \neq 0$, such that $\langle \eta^0, z^0\rangle > 0$. Choosing $\xi^0 = 0 \in C'$ we get the desired inequality.

Assume now $z^0 \in -K[-g(x^0)]$. Then we must have $y^0 \notin -C\backslash\{0\}$. If $y^0 \neq 0$, there exists a vector $\xi^0 \in C'\backslash\{0\}$, such that $\langle \xi^0, y^0 \rangle > 0$ and for $\eta^0 = 0 \in K'[-g(x^0)]$ the conclusion is obtained.

Remark 1. When C is a closed convex pointed cone with nonempty interior, then in Theorem 2, we can assume $\xi^0 \in \operatorname{int} C'$.

Constraint qualification $\mathcal{Q}_{0,1}(x^0)$ is essential in order that Theorems 2 and 3 hold, as shown by the next example.

Example 1. Let $f : \mathbb{R}^2 \to \mathbb{R}^2$, $g : \mathbb{R}^2 \to \mathbb{R}^2$ be defined as $f(x_1, x_2) = (x_1, -x_1^2)$, if $x_1 \leq 0$, $f(x_1, x_2) = (x_1^2, -x_1^2)$, if $x_1 > 0$ and $g(x_1, x_2) = (-x_1^3 + x_2, -x_2)$. Let $C = K = \mathbb{R}^2_+$. The point $x^0 = (0, 0)$ is a p-minimizer of first order for problem (1), but constraint qualification $\mathcal{Q}_{0,1}(x^0)$ is not satisfied and the necessary optimality conditions of Theorem 2 and 3 do not hold at x^0.

Indeed, consider the vector $u = (-1, 0)$. We have $g'_u(x^0) = 0 \in -K[-g(x^0)]$, but it is easily seen that constraint qualification $\mathcal{Q}_{0,1}(x^0)$ does not hold. The optimality conditions of Theorems 2 and 3 are not satisfied since we have $f'_u(x^0) = (-1, 0)$.

Let us underline that in $\mathcal{N}\mathcal{D}'_{0,1}$ the multipliers $(\xi^0,\ \eta^0)$ depend on the direction u and differently from some known results the inequality is strict. The strict inequality applied to p-minimizers gives similarly looking necessary conditions and sufficient conditions, compare Theorems 3 and 5. The eventual independence of the multipliers on the direction when f and g are differentiable is discussed in [10]. Then the optimality conditions of Theorems 2 and 3 show similarity with the classical Kuhn-Tucker conditions for properly efficient points (see e.g. [24]).

Now we establish second order necessary optimality conditions when f and g are $C^{1,1}$ functions. We give directly the dual formulation of the proposed conditions. In the next theorem we use the following second order constraint qualification.

$\mathcal{Q}_{1,1}(x^0)$: The following is satisfied, in the sense that provided x^0, u^0, t_k and z^0 are such that 1^0, 2^0, 3^0 and the first line of 4^0 are satisfied, then the last line of 4^0 must hold

1^0: $$g(x^0) \in -K\,,$$

2^0: $$g'(x^0)u^0 \in -(K[-g(x^0)] \setminus \operatorname{int} K[-g(x^0)])\,,$$

3^0: $$\frac{2}{t_k^2}\ \left(g(x^0 + t_k u^0) - g(x^0) - t_k g'(x^0) u^0\right) \to z^0\,,$$

4^0: $$\begin{array}{c}\left(\forall\, \eta \in K'[-g(x^0)], (\forall z \in \operatorname{img}'(x^0) : \langle \eta, z \rangle = 0) : \langle \eta, z^0 \rangle \leq 0\right) \Rightarrow \\ \forall w \in \mathbb{R}^n\ \exists w^k \to w : \exists k_0 \in \mathbb{N} : \forall k > k_0 : g(x^0 + t_k u^0 + \frac{t_k^2}{2} w^k) \in -K.\end{array}$$

Roughly speaking the geometrical meaning of the constraint qualification $\mathcal{Q}_{0,1}(x^0)$ is the following. If the differential quotient of g determines a tangent direction z^0 in the the contingent cone $-K[-g(x^0)]$, then the same tangent direction can be determined by the differential quotient constituted by feasible points. Though the definition of the constraint qualification $\mathcal{Q}_{1,1}(x^0)$ is more complicated, the geometrical meaning is similar. Namely, if the second-order tangent direction z^0 belongs to the second-order contingent cone, then this direction can be determined by a second-order quotient of feasible points.

For $x^0 \in \mathbb{R}^n$ we put

$$\Delta(x^0) = \left\{(\xi,\, \eta) \in C' \times K' \mid (\xi, \eta) \neq 0,\ \langle \eta,\, g(x^0)\rangle = 0,\, \xi\, f'(x^0) + \eta\, g'(x^0) = 0\right\}\, .$$

Theorem 4. *Let in problem (1) f and g be $C^{1,1}$ functions and assume C and K are closed convex cones with nonempty interior. If x^0 is a p-minimizer of second order for problem (1) and the constraint qualification $\mathcal{Q}_{1,1}(x^0)$ holds, then for every $u \in \mathbb{R}^n \backslash \{0\}$ one of the following conditions is satisfied*

$\mathcal{NP}'_{1,1}$: $\quad (f'(x^0)u,\, g'(x^0)u) \notin -(C\backslash\{0\} \times K[-g(x^0)])\,,$

$\mathcal{ND}''_{1,1}$: $\quad (f'(x^0)u, g'(x^0)u) \in -\big((C\backslash\{0\} \times K[g(x^0)])\backslash(\text{int}C \times \text{int}K[-g(x^0)])\big)$ *and* $\forall\, (y^0, z^0) \in (f,g)''_u(x^0) : \exists\, (\xi^0, \eta^0) \in \Delta(x^0) :$ $\langle \xi^0,\, y^0\rangle + \langle \eta^0,\, z^0\rangle > 0.$

Proof. Let x^0 be a p-minimizer of order two for problem (1), which means that $g(x^0) \in -K$ and there exists $r > 0$ and $A > 0$ such that $g(x) \in -K$ and $\|x - x^0\| \le r$ implies

$$D(f(x) - f(x^0), -C) \ge A\, \|f(x) - f(x^0)\|^2\, . \tag{5}$$

Therefore it satisfies the condition

$\mathcal{N}'_{1,1}$: $\quad (f'(x^0)u,\, g'(x^0)u) \notin -(\text{int}\, C \times \text{int}\, K[-g(x^0)])\,,$

(see [10], [12]). Now it becomes obvious, that for each $u = u^0$ one and only one of the first-order conditions in $\mathcal{NP}'_{1,1}$ and the first part of $\mathcal{ND}''_{1,1}$ is satisfied. Suppose that $\mathcal{NP}'_{1,1}$ is not satisfied. Then the first part of condition $\mathcal{ND}''_{1,1}$ holds, that is

$$(f'(x^0)u^0,\, g'(x^0)u^0) \in -\big((C\backslash\{0\} \times K[-g(x^0)])\setminus(\text{int}\, C \times \text{int}\, K[-g(x^0)])\big)\, .$$

We prove, that also the second part of condition $\mathcal{ND}''_{1,1}$ holds. Let $t_\nu \to 0^+$ be a sequence such that

$$\frac{2}{t_\nu^2}\, \big(f(x^0 + t_\nu u^0) - f(x^0) - t_\nu f'(x^0)u^0\big) \to y^0$$

and

$$\frac{2}{t_\nu^2}\,\left(g(x^0+t_\nu u^0)-g(x^0)-t_\nu g'(x^0)u^0\right)\to z^0\,.$$

One of the following two cases may arise:

1^0. *There exists* $\eta^0\in K[-g(x^0)]'$ *such that* $\langle\eta^0,\,z\rangle=0$ *for all* $z\in\operatorname{im}g'(x^0)$ *and* $\langle\eta^0,\,z^0\rangle>0$.

We put now $\xi^0=0$. Then we have obviously $(\xi^0,\eta^0)\in C'\times K'[-g(x^0)]$ and $\langle\xi^0,\,f'(x^0)\rangle+\langle\eta^0,\,g'(x^0)\rangle=0$. Thus $(\xi^0,\eta^0)\in\Delta(x^0)$ and $\langle\xi^0,\,y^0\rangle+\langle\eta^0,\,z^0\rangle>0$. Therefore condition $\mathcal{ND}''_{1,1}$ is satisfied.

2^0. *For all* $\eta\in K'[-g(x^0)]$, *such that* $\langle\eta,\,z\rangle=0$ *for all* $z\in\operatorname{im}g'(x^0)$, *it holds* $\langle\eta,\,z^0\rangle\le 0$.

This condition coincides with condition 4^0 in the constraint qualification $\mathcal{Q}_{1,1}(x^0)$. Now we see that all points 1^0–4^0 in the constraint qualification $\mathcal{Q}_{1,1}(x^0)$ are satisfied. Therefore, for every $w\in\mathbb{R}^n$, there exists $w^\nu\to w$ and a positive integer ν_0 such that for all $\nu>\nu_0$ it holds $g(x^0+t_\nu u^0+\frac{t_\nu^2}{2}w^\nu)\in -K$. Passing to a subsequence, we may assume that this inclusion holds for all ν. We have

$$D\left(f\left(x^0+t_\nu u^0+\frac{t_\nu^2}{2}w^\nu\right)-f(x^0),-C\right)$$
$$\ge A\left\|f\left(x^0+t_\nu u^0+\frac{t_\nu^2}{2}w^\nu\right)-f(x^0)\right\|^2$$

and hence, since $f'(x^0)u^0\in -C$

$$D\left(f\left(x^0+t_\nu u^0+\frac{t_\nu^2}{2}w^\nu\right)-f(x^0)-t_\nu f'(x^0)u^0,-C\right)$$
$$\ge D\left(f\left(x^0+t_\nu u^0+\frac{t_\nu^2}{2}w^\nu\right)-f(x^0),-C\right)$$
$$\ge A\left\|f\left(x^0+t_\nu u^0+\frac{t_\nu^2}{2}w^\nu\right)-f(x^0)\right\|^2\,.$$

We get further

$$\frac{2}{t_\nu^2}\left(f\left(x^0+t_\nu u^0+\frac{t_\nu^2}{2}w^\nu\right)-f(x^0)-t_\nu f'(x^0)u^0\right)$$
$$=\frac{2}{t_\nu^2}\left(f\left(x^0+t_\nu(u^0+\frac{t_\nu}{2}w^\nu)\right)-f(x^0)-t_\nu f'(x^0)(u^0+\frac{t_\nu}{2}w^\nu)\right.$$
$$\left.+\frac{t_\nu^2}{2}f'(x^0)w^\nu\right)$$

and since f is of class $C^{1,1}$, the second term in this equality converges to $y^0+f'(x^0)w$, while

$$\frac{1}{t_\nu}\left(f\left(x^0 + t_\nu u^0 + \frac{t_\nu^2}{2}w^\nu\right) - f(x^0)\right) \to f'(x^0)u$$

(this is an easy consequence of Proposition 3). From

$$\frac{2}{t_\nu^2}D\left(f\left(x^0 + t_\nu u^0 + \frac{t_\nu^2}{2}w^\nu\right) - f(x^0) - t_\nu f'(x^0)u^0, -C\right)$$
$$\geq \frac{2A}{t_\nu^2}\left\|f\left(x^0 + t_\nu u^0 + \frac{t_\nu^2}{2}w^\nu\right) - f(x^0)\right\|^2$$

we obtain

$$D(y^0 + f'(x^0)w, -C) \geq 2A\|f'(x^0)u\|^2 > 0$$

since $f'(x^0)u \neq 0$. Because w is arbitrary, we obtain

$$\inf\{D(y^0{+}f'(x^0)w, -C) \,|\, w \in \mathbb{R}^n\} := D(y^0{+}\mathrm{im}\, f'(x^0), -C) \geq 2A\|f'(x^0)u\|^2\,,$$

which implies $0 \notin \mathrm{cl}\,(y^0 + \mathrm{im}\, f'(x^0) + C)$. Hence, according to Theorem 11.4 in [23], the convex sets $-C$ and $y^0 + \mathrm{im}\, f'(x^0)$ are are strongly separated, i.e. there exists a vector $\xi \in \mathbb{R}^m$, such that

$$\inf\{\langle \xi, y\rangle \,|\, y \in y^0 + \mathrm{im}\, f'(x^0)\} > \sup\{\langle \xi, y\rangle \,|\, y \in -C\}\,.$$

Let $\beta = \sup\{\langle \xi, y\rangle \,|\, y \in -C\}$. Since $-C$ is a cone, we have

$$\langle \xi, \lambda y\rangle \leq \beta\,,\ \forall y \in -C\,,\ \forall \lambda > 0.$$

This implies $\beta \geq 0$ and $\langle \xi, y\rangle \leq 0$ for every $y \in -C$. Hence $\beta = 0$ and $\xi \in C'$. Further, from $\inf\{\langle \xi, y\rangle | y \in y^0 + \mathrm{im}\, f'(x^0)\} > 0$, we get easily $\xi^0 f'(x^0) = 0$. Indeed, otherwise we would find a vector $w \in \mathbb{R}^n$ such that $\langle \xi^0, f'(x^0)w\rangle < 0$ and hence we would have

$$\langle \xi^0, y^0 + \lambda f'(x^0)w\rangle > 0\,,\ \forall \lambda > 0\,,$$

but this is impossible, since the left side tends to $-\infty$ as $\lambda \to +\infty$, while the right side does not depend on λ. This completes the proof.

Theorems 2, 3 and 4 are valid and simplify in an obvious way when we consider the unconstrained problem (2). Let us underline that in this case we do not need the constraint qualifications $\mathcal{Q}_{0,1}(x^0)$ and $\mathcal{Q}_{1,1}(x^0)$.

4 Optimality Under Convexity Conditions

In this section we revert the necessary conditions of Theorem 2, under convexity assumptions on the functions f and g. We recall that given a convex cone $D \subseteq \mathbb{R}^k$, a function $\Phi : \mathbb{R}^n \to \mathbb{R}^k$ is said to be D-convex when $\Phi((1-t)x^1 + tx^2) \in (1-t)\Phi(x^1) + t\Phi(x^2) - D$, for every $t \in [0,1]$ and $x^1, x^2 \in \mathbb{R}^n$. We recall also that a set $A \subseteq \mathbb{R}^k$ is D-convex when $A + D$ is convex.

We need the following lemma.

Lemma 1. *Let C be a closed convex pointed cone with nonempty interior and let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a C-convex function. Then*

i) f is of class $C^{0,1}$,

ii) for every $u \in \mathbb{R}^n$ and every $x^0 \in \mathbb{R}^n$, $f'_u(x^0)$ is a singleton, i.e.

$$f'_u(x^0) = \lim_{t \to 0^+} \frac{1}{t} \left(f(x^0 + tu) - f(x^0) \right) .$$

Proof. i) The proof is analogous to that of Lemma 2.1 in [27] and hence is omitted.

ii) It is well known that f is C-convex if and only if for every $\lambda \in C'$, the scalar function $f_\lambda(x) = \langle \lambda, f(x) \rangle$ is convex and hence directionally differentiable. That is the limit $\lim_{t \to 0^+} \frac{1}{t} \left(f_\lambda(x^0 + tu) - f_\lambda(x^0) \right) = (f_\lambda)'_u(x)$ exists and is finite for all $x \in \mathbb{R}^n$ and $u \in \mathbb{R}^n$. It follows that for every $y \in f'_u(x)$ and $\lambda \in C'$ it holds $(f_\lambda)'_u(x) = \langle \lambda, y \rangle$. Since C is a pointed cone, then C' has nonempty interior and hence it is possible to find m linearly independent vector $\lambda^1, \dots, \lambda^m \in C'$. Thus the system

$$\begin{cases} \langle \lambda^1, y \rangle = (f_{\lambda^1})'_u(x) \\ \qquad \cdots \\ \langle \lambda^m, y \rangle = (f_{\lambda^m})'_u(x) \end{cases} \tag{6}$$

admits a unique solution. This proves that there exists a unique $y \in f'_u(x^0)$, whence

$$f'_u(x^0) = \lim_{t \to 0^+} \frac{1}{t} \left(f(x^0 + tu) - f(x^0) \right) .$$

Remark 2. Results similar to that of Lemma 1 ii) can be found in [26].

Theorem 5. *Let in problem (1) C and K be closed convex pointed cones with nonempty interior. Assume f is C-convex and g is K-convex. If the first order necessary optimality conditions of Theorem 2 hold at the feasible point x^0, then x^0 is a p-minimizer of first order.*

Proof. From the previous lemma we know that both $f'_u(x^0)$ and $g'_u(x^0)$ are singletons. Since g is K-convex, the sets $A = \{x : g(x) \in -K\}$ and $T(x^0, A) = \operatorname{cl} \operatorname{co}(A - x^0)$ are convex. We observe that if $u \in T(x^0, A)$, then $g'_u(x^0) \in -K[-g(x^0)]$. Indeed, if $u \in \operatorname{co}(A - x^0)$, we have $u = \alpha(x - x^0)$ for some $\alpha > 0$ and $x \in A$. If $t_\nu \to 0^+$ we have, since g is K-convex

$$g(x^0 + t_\nu u) - g(x^0) = g(x^0 + t_\nu \alpha(x - x^0)) - g(x^0) \in t_\nu \alpha g(x) - t_\nu \alpha g(x^0) - K .$$

Let $\xi \in K'[-g(x^0)]$ be arbitrarily chosen. Hence from the previous inclusion we obtain $\langle \xi, g(x^0 + t_\nu u) - g(x^0) \rangle \le 0$, that is $g(x^0 + t_\nu u) - g(x^0) \in -K[-g(x^0)]$. Dividing by t_ν, passing to the limit and taking into account that $K[-g(x^0)]$ is a closed cone, we get $g'_u(x^0) \in -K[-g(x^0)]$. Let now $u \in T(x^0, A)$, i.e. $u = \lim u^\nu$, with $u^\nu \in \operatorname{co}(A - x^0)$. Since $g'_u(x^0)$ is single valued, Proposition 3 gives $\|g'_u(x^0) - g'_{u^\nu}(x^0)\| \le L\|u - u^\nu\|$, whence $g'_{u^\nu}(x^0) \to g'_u(x^0)$, as $\nu \to +\infty$. From $g'_{u^\nu}(x^0) \in -K[-g(x^0)]$, we get $g'_u(x^0) \in -K[-g(x^0)]$. Consider the set

$$P = \{y \in \mathbb{R}^m : y = f'_u(x^0)\,, u \in T(x^0, A)\}\,.$$

Clearly P is a cone. We show that it is closed and C-convex. The set $P_1 = \{y \in \mathbb{R}^m : y = f'_u(x^0), u \in T(x^0, A) \cap S\}$ is closed. Indeed, let $y^\nu \in P_1$, $y^\nu \to y$. We have $y^\nu = f'_{u^\nu}(x^0)$, $u^\nu \in T(x^0, A) \cap S$ and since $T(x^0, A) \cap S$ is compact, we can assume $u^\nu \to u \in T(x^0, A) \cap S$. Using Proposition 3, we have $y^\nu \to f'_u(x^0)$, whence $y \in P_1$. From Proposition 3 it follows that P_1 is also bounded and hence compact. The closedness of the cone P now follows since it is generated by the compact base P_1 and easily we obtain that $P + C$ is closed too. Now, trivial calculations show that $f'_u(x^0)$ is C-convex as a function of the direction $u \in T(x^0, A)$ and this gives that $P + C$ is convex.
The first order condition of Theorem 2 implies $(P + C) \cap (-C) = \{0\}$. Hence (see e.g. [24]), there exists a vector $\xi \in \operatorname{int} C'$, such that $\langle \xi, w\rangle \ge 0$, $\forall w \in P + C$. From the C-convexity of f we get $\langle \xi, f(x) - f(x^0)\rangle \ge \langle \xi, f'(x^0, x - x^0)\rangle \ge 0\,.$ The existence of such vector ξ proves that x^0 is a p-minimizer (see e.g. [19], [24]).

Remark 3. The previous theorem holds under the weaker requirement that g is $K[-g(x^0)]$-convex (see e.g. [5]).

5 *p*-Minimizers and *i*-Minimizers of Second Order

In [11] we investigated the problem under which conditions Theorem 1 admits a reversal in the case $k = 1$, that is under which conditions a p-minimizer of first order is also an i-minimizer of first order. In particular, we obtained the following result, in whose proof play a crucial role the first order optimality conditions in terms of Dini derivatives.

Theorem 6. *Let f and g be $C^{0,1}$ functions and let x^0 be a p-minimizer of first order for the constrained problem (1), which has the property*

$$(y^0, z^0) \in (f, g)'_u(x^0) \text{ and } z^0 \in -K(x^0) \text{ implies } y^0 \ne 0\,. \tag{7}$$

Then x^0 is an i-minimizer of first order for (1). If in particular we consider the unconstrained problem (2), then every p-minimizer of first order is an i-minimizer of first order under the condition $0 \notin f'_u(x^0)$.

In this section and in the next one, we apply the second order optimality conditions of Section 3 in order to solve a similar problem for p-minimizers and i-minimizers of second order. We leave as an open problem the reversal of Theorem 1 for arbitrary real $k \ge 1$.

We begin observing that a reversal of Theorem 1 is not possible in general, as shown by the following example.

Example 2. Consider the unconstrained problem (2) and let $f : \mathbb{R} \to \mathbb{R}^2$, $f(x) = (f_1(x), f_2(x))$, with $f_1(x) = -x^2 \sin \frac{1}{x} - x^2$ and $f_2(x) = [f_1(x)]^2$, if $x \neq 0$, and $f_1(0) = f_2(0) = 0$. The point $x_0 = 0$ is a p-minimizer of any order $k \geq 2$, but there exists no positive number k, such that x_0 is an i-minimizer of order k.

Together with the optimality conditions of Section 3, we need the following result.

Theorem 7 ([12]). *Consider problem (1) with f and g being $C^{1,1}$ functions and C and K closed convex cones with nonempty interior. Let x^0 be a feasible point for problem (1). Suppose that for each $u \in \mathbb{R}^n \setminus \{0\}$ one of the following two conditions is satisfied:*

$$\mathcal{S}'_{1,1}: \qquad (f'(x^0)u,\, g'(x^0)u) \notin -(C \times K[-g(x^0)])\,,$$

$$\mathcal{S}''_{1,1}: \quad \begin{array}{c} (f'(x^0)u,\, g'(x^0)u) \in -\big((C \times K[-g(x^0)]) \setminus (\operatorname{int} C \times \operatorname{int} K[-g(x^0)])\big) \\ \textit{and}\ \forall\, (y^0, z^0) \in (f,g)''_u(x^0) \exists\, (\xi^0, \eta^0) \in \Delta(x^0) : \langle \xi^0,\, y^0 \rangle + \langle \eta^0,\, z^0 \rangle > 0\,. \end{array}$$

Then x^0 is an i-minimizer of order two for problem (1). Conversely, if x^0 is an i-minimizer of second-order for problem (1) and the constraint qualification $\mathcal{Q}_{1,1}(x^0)$ holds, then one of the conditions $\mathcal{S}'_{1,1}$ and $\mathcal{S}''_{1,1}$ is satisfied.

Theorem 7 is valid and simplifies in an obvious way when instead of problem (1), we consider the unconstrained problem (2). Let us underline that in this case the reversal of the sufficient conditions does not require the use of the constraint qualifications.

Theorem 8. *Consider problem (2) with f being a $C^{1,1}$ function and C a closed convex cone with nonempty interior. Let x^0 be a feasible point for problem (2). Suppose that for each $u \in \mathbb{R}^n \setminus \{0\}$ one of the following two conditions is satisfied*

$$1^0: \qquad f'(x^0)u \notin -C\,,$$

$$2^0: \quad \begin{array}{c} f'(x^0)u \in -C \setminus \operatorname{int} C \ \textit{and} \\ \forall\, y^0 \in f''_u(x^0) : \exists\, \xi^0 \in C' \setminus \{0\} : \xi^0 f'(x^0) = 0 \ \ \textit{and}\ \langle \xi^0,\, y^0 \rangle > 0\,. \end{array}$$

Then x^0 is an i-minimizer of order two for problem (2).
Conversely, if x^0 is an i-minimizer of second-order for problem (2) then one of the conditions above is satisfied.

Theorem 9. *Let f and g be $C^{1,1}$ functions, C and K be closed convex cones with nonempty interior and let x^0 be a p-minimizer of second order for the constrained problem (1), which has the property $\forall u \in \mathbb{R}^n \setminus \{0\}$*

$$\begin{array}{c} g'(x^0)u \in -K[-g(x^0)] \ \ \textit{and} \ \ f'(x^0)u = 0 \\ \textit{implies}\ \forall (y^0, z^0) \in (f,g)''_u(x^0)\ \exists (\xi^0, \eta^0) \in \Delta(x^0), \\ \textit{with} \ \ \langle \xi^0, y^0 \rangle + \langle \eta^0, z^0 \rangle > 0. \end{array} \tag{8}$$

If constraint qualification $\mathcal{Q}_{1,1}(x^0)$ holds, then x^0 is an i-minimizer of second order for (1).

Proof. Since x^0 is a p-minimizer of second order and the constraint qualification $\mathcal{Q}_{1,1}(x^0)$ holds, then the necessary optimality conditions of Theorem 4 are satisfied. Under the made assumptions these conditions coincide with the sufficient conditions of Theorem 7, whence x^0 is an i-minimizer of second order.

As an immediate consequence of the previous result we get the following

Corollary 1. *Let f and g be $C^{1,1}$ functions, C and K be closed convex cones with nonempty interior and let x^0 be a p-minimizer of second order for the constrained problem (1), which has the property*

$$g'(x^0)u \in -K[-g(x^0)] \quad \textit{implies} \quad f'(x^0)u \neq 0 \quad \forall u \in \mathbb{R}^n \backslash \{0\} \tag{9}$$

and assume that the constraint qualification $\mathcal{Q}_{1,1}(x^0)$ holds. Then x^0 is an i-minimizer of second order for (1).

Remark 4. If we formulate the previous theorem with regard to the unconstrained problem (2), then we get that a p-minimizer of second order is an i-minimizer under the condition

$$\begin{gathered} f'(x^0)u = 0 \quad \textit{implies} \quad \forall y^0 \in f''_u(x^0)\ \exists \xi^0 \in C'\backslash\{0\}, \\ \textit{with} \quad \xi^0 f'(x^0) = 0 \quad \textit{and} \quad \langle \xi^0, y^0 \rangle > 0\,. \end{gathered} \tag{10}$$

6 Two Approaches Towards Proper Efficiency

In the unconstrained case, as observed in [11], condition $0 \notin f'_u(x^0)$ is both necessary and sufficient in order that a p-minimizer of first order is also an i-minimizer of first order and analogously condition (10) is also necessary in order that a p-minimizer of second order is an i-minimizer of second order (this is shown by Theorem 8). Easy examples show, instead, that when dealing with the constrained problem (1), both conditions (7) and (8) are not necessary in order the previous implication holds. This problem has been investigated in [11] for p-minimizers of first order. In this section we deal with a similar problem for p-minimizers of second order.
In the constrained case one observes that the sufficient optimality conditions of Theorem 7 involve the condition

$$\begin{gathered} (f'(x^0)u, g'(x^0)u) = 0 \\ \textit{implies}\ \forall (y^0, z^0) \in (f,g)''_u(x^0)\ \exists (\xi^0, \eta^0) \in \Delta(x^0), \\ \textit{with} \quad \langle \xi^0, y^0 \rangle + \langle \eta^0, z^0 \rangle > 0,\ \forall u \in \mathbb{R}^n \backslash \{0\}. \end{gathered} \tag{11}$$

(which is weaker than (8)). Therefore one is led to wonder whether this condition is necessary and sufficient in order that a p-minimizer of second order x^0 is also an i-minimizer of second order. The next examples show however that this is not the case.

Example 3. Consider the constrained problem (1), with $f:\mathbb{R}\to\mathbb{R}$, $f(x)=-x^2$, $C=\mathbb{R}_+$, $g:\mathbb{R}\to\mathbb{R}$, $g(x)=x^4$, $K=\mathbb{R}_+$. The point $x_0=0$ is the only feasible point and hence it is both a p-minimizer and an i-minimizer of second order. Anyway, for every $u\in\mathbb{R}^n$ we have $(f'(x^0)u,g'(x^0)u)=0$ and $(f,g)''_u(x^0)=(-2u,0)$, and hence condition (11) is not satisfied.

Example 4. Consider problem (1), with $f:\mathbb{R}^2\to\mathbb{R}^2$, $f(x)=(x_1^4+x_2^2,-x_1^4-x_2^2)$, $C=\mathbb{R}^2_+$, $g:\mathbb{R}^2\to\mathbb{R}^2$, $g(x)=(-x_1^4,-x_1^4)$, $K=\mathbb{R}^2_+$ and let $x^0=(0,0)$. For $u=(0,1)$, we have $(f'(x^0)u,g'(x^0)u)=(0,0)$ and $(f,g)''_u(x^0)=((2,-2),(0,0))$, whence condition (11) is satisfied. Further, constraint qualification $\mathcal{Q}_{1,1}(x^0)$ holds and x^0 is a p-minimizer of second order, but not an i-minimizer of second order.

In virtue of Example 4, to obtain a result similar to Theorem 9 under condition (11), we need a new approach toward the concepts of i-minimizer and p-minimizer. In [11], while investigating the links between i-minimizers and p-minimizers of first order, we have related to the constrained problem (1) and the feasible point x^0, the unconstrained problem

$$\min{}_{C\times K[-g(x^0)]}\,(f(x),g(x))\,,\quad x\in\mathbb{R}^n\,. \tag{12}$$

Definition 3. *We say that x^0 is a p-minimizer of order k in sense II for the constrained problem (1) if x^0 is a p-minimizer of order k for the unconstrained problem (12).*

Similarly, we say that x^0 is an isolated minimizer of order k in sense II for the constrained problem (1) if x^0 is an isolated minimizer of order k for the unconstrained problem (12).

We will preserve the names for the concepts used so far, but sometimes we will refer to them as sense I concepts, saying e. g. p-minimizer in sense I, instead of just p-minimizer.

As an immediate application of Corollary 1 we get the following result.

Theorem 10. *Let f and g be $C^{1,1}$ functions, C and K be closed convex cones with nonempty interior and let x^0 be a p-minimizer of second order in sense II for the constrained problem (1), which has property (11). Then x^0 is an i-minimizer of second order in sense II for (1).*

Next, under the hypotheses of Theorem 10, we show that x^0 is an i-minimizer in sense I. We state also relations between sense I and sense II, i-minimizers and p-minimizers.

Theorem 11. *Let f and g be $C^{1,1}$ functions, C and K be closed convex cones with nonempty interior and let x^0 be a p-minimizer of second order in sense II for the constrained problem (1), which has property (11). Then x^0 is an i-minimizer of second order in sense I for (1) and hence x^0 is also a p-minimizer of second order in sense I.*

Proof. According to Theorem 10, x^0 is an i-minimizer of second order. The reversal of the sufficient conditions of Theorem 8 applied to problem (12) gives a condition which coincides with the sufficient condition $\mathcal{S}''_{1,1}$ of Theorem 7, whence x^0 is an i-minimizer of second order in sense I for the constrained problem (1). Theorem 1 gives now that x^0 is also a p-minimizer of second order in sense I for (1).

Thus, within the set of points satisfying (11) the set of the p-minimizers of second order in sense II is a subset of the p-minimizers of second order in sense I. The reversal does not hold. In fact, the following reasoning shows, that in Example 4 the point x^0 is a p-minimizer of second order in sense I (indeed is a p-minimizer of first order), but it is not a p-minimizer of second order in sense II. Now, for the corresponding problem (12) we have

$$(f,g) : \mathbb{R}^2 \to \mathbb{R}^4 \,, \quad (f(x), g(x)) = (x_1^4 + x_2^2, -x_1^4 - x_2^2, -x_1^4, -x_1^4)$$

and $C \times K[-g(x_0)] = \mathbb{R}^2_+ \times \mathbb{R}^2_+ = \mathbb{R}^4_+$. Each point $x \in \mathbb{R}^2$ is feasible and we have $\max\{x_1^4+x_2^2, -x_1^4-x_2^2, -x_1^4, -x_1^4\} = x_1^4+x_2^2$, whence x^0 is an i-minimizer of order 4 in sense II, but it is not an i-minimizer of second order in sense II. Therefore, according to Theorem 10, in spite that x^0 is a p-minimizer of second order in sense I, it is not a p-minimizer of second order in sense II (the assumption that x^0 is a p-minimizer of second order in sense II would imply that x^0 is an i-minimizer of second order in sense II).

Let us now make some comparison between Theorems 9 and 10. In spite that condition (11) is more general than condition (8), Theorem 10 does not imply Theorem 9. Indeed, the assumption in Theorem 10 is that x^0 is a p-minimizer in sense II, which does not imply the more general assumption in Theorem 9 that x^0 is a p-minimizer in sense I.

Next we compare the i-minimizers in sense I and II.

Theorem 12. *Let f and g be $C^{1,1}$ functions, C and K be closed convex cones with nonempty interior. If x^0 is an i-minimizer of second order in sense II for the constrained problem (1), then x^0 is an i-minimizer of second order in sense I for (1). If the constraint qualification $\mathcal{Q}_{1,1}(x^0)$ holds, then also the converse is true.*

Proof. Let x^0 be an i-minimizer of second order in sense II. The reversal of the sufficient conditions of Theorem 8 gives the sufficient condition $\mathcal{S}''_{1,1}$ of Theorem 7, whence x^0 is an i-minimizer of second order in sense I.

Conversely, let x^0 be an i-minimizer of second order in sense I. Under the constraint qualification $\mathcal{Q}_{1,1}(x^0)$, we can apply the reversal of the sufficient conditions of Theorem 7, getting condition $\mathcal{S}''_{1,1}$, which is identical with the sufficient conditions of Theorem 8 applied to the unconstrained problem (12), whence x^0 is an i-minimizer of second order in sense II.

We conclude the paper with the following comments. The comparison of the p-minimizers and the i-minimizers has led us to "duplicate" the notions of optimality, introducing sense II concepts.

Indeed, we have to underline that the points we call sense II minimizers are not minimizers of the considered constrained problem but of a related unconstrained vector problem and one prefers probably to deal with this simpler unconstrained problem instead of the constrained one. The name is justified since we find a connection between the properties of a point to be sense I or sense II minimizer. This connection has been obtained in [11] for p-minimizers and i-minimizers of first order and has been investigated in this section in the second-order case. A crucial role in this study plays the observation that the optimality conditions for the related unconstrained problem (12) coincide with the optimality conditions of the constrained problem (1). A motivation to introduce sense II concepts give also the stability properties obeyed by the p-minimizers and the isolated minimizers. In [5] and [9] we have shown that sense I concepts are stable under perturbations of the objective data, while sense II concepts are stable under perturbations of both the objective and constraint data. From this point of view sense II concepts are advantageous, since it is preferable to deal with a problem, which is stable with respect to all data, than with one which is stable with only part of the data.

References

1. Amahroq T, Taa A (1997) On Lagrange-Kuhn-Tucker multipliers for multiobjective optimization problems. Optimization 41:159–172
2. Aubin J-P, Frankowska H (1990) Set-valued analysis. Birkhäuser, Boston Basel Berlin
3. Auslender A (1984) Stability in mathematical programming with nondifferentiable data. SIAM J. Control Optim. 22:239–254
4. Ciligot-Travain M (1994) On Lagrange-Kuhn-Tucker multipliers for Pareto optimization problems. Numer. Funct. Anal. Optim. 15:689–693
5. Crespi GP, Ginchev I, Rocca M (2005) Two approaches toward vector optimization and identity of the solutions. Journal of Industrial and Management Optimization 1:549–563
6. Crespi GP, Ginchev I, Rocca M (2005) First-order optimality conditions in set-valued optimization. Math. Methods Oper. Res. 63:87–106
7. Geoffrion AM (1968) Proper efficiency and the theory of vector maximization. J. Math. Anal. Appl. 22:618–630
8. Ginchev I (2002) Higher order optimality conditions in nonsmooth vector optimization. J. Stat. Manag. Syst. 5:321–339
9. Ginchev I, Guerraggio A, Rocca M (2004) Isolated minimizers, proper efficiency and stability for $C^{0,1}$ constrained vector optimization problems. Preprint 2004/9, Universitá dell'Insubria, Facoltá di Economia, Varese (http:// eco.uninsubria.it/ dipeco/ Quaderni/ files/ QF2004_9.pdf)
10. Ginchev I, Guerraggio A, Rocca M (2005) First-order conditions for $C^{0,1}$ constrained vector optimization. In: Giannessi F, Maugeri A (eds.), Variational analysis and applications. Springer, Berlin New York

11. Ginchev I, Guerraggio A, Rocca M (2005) Isolated minimizers and proper efficiency for $C^{0,1}$ constrained vector optimization problems. J. Math. Anal. Appl. 309:353–368
12. Ginchev I, Guerraggio A, Rocca M (2005) Second-order conditions in $C^{1,1}$ constrained vector optimization. Math. Program. ser. B 104:389–405
13. Ginchev I, Guerraggio A, Rocca M (2006) From scalar to vector optimization. Appl. Math. 51:5–36
14. Ginchev I, Guerraggio A, Rocca M, Geoffrion type characterization of higher order properly efficient points in vector optimization. J. Math. Anal. Appl., submitted
15. Ginchev I, Hoffmann A (2002) Approximation of set-valued functions by single-valued one. Discuss. Math. Differ. Incl. Control Optim. 22:33–66
16. Henig MI (1982) Proper efficiency with respect to cones. J. Optim. Theory Appl. 36:387–407
17. Hiriart-Urruty J-B (1979) New concepts in nondifferentiable programming. Analyse non convexe, Bull. Soc. Math. France 60:57–85
18. Hiriart-Urruty J-B (1979) Tangent cones, generalized gradients and mathematical programming in Banach spaces. Math. Oper. Res. 4:79–97
19. Luc DT (1989) Theory of vector optimization. Springer, Berlin
20. Mangasarian OL (1994) Nonlinear programming. SIAM, Philadelphia
21. Miglierina E, Molho E (2002) Scalarization and its stability in vector optimization. J. Optim. Theory Appl. 114:657–670
22. Miglierina E, Molho E, Rocca M (2005) Well-posedness and scalarization in vector optimization. J. Optim. Theory Appl. 126:391–409
23. Rockafellar RT (1970) Convex analysis. Princeton University Press, Princeton
24. Sawaragi Y, Nakayama H, Tanino T (1985) Theory of multiobjective optimization. Academic Press, New York
25. Studniarski M (1986) Necessary and sufficient conditions for isolated local minima of nonsmooth functions. SIAM J. Control Optim. 24:1044–1049
26. Tan NX, Tinh PN (2000) On conjugate maps and directional derivatives of convex vector functions. Acta Math. Vietnam. 25:315–345
27. Tanino T (1988) Stability and sensitivity analysis in convex vector optimization. SIAM J. Control Optim. 26:521-536
28. Ward DE (1994) Characterizations of strict local mininima and necessary conditions for weak sharp minima. J. Optim. Theory Appl. 80:551-571
29. Zaffaroni A (2003) Degrees of efficiency and degrees of minimality. SIAM J. Optim. 42:1071–1086

Higher-order Pseudoconvex Functions

Ivan Ginchev[1] and Vsevolod I. Ivanov[2]

[1] Technical University of Varna, Department of Mathematics
9010 Varna, Bulgaria, iginchev@yahoo.com
[2] Technical University of Varna, Department of Mathematics
9010 Varna, Bulgaria, vsevolodivanov@yahoo.com

Summary. In terms of n-th order Dini directional derivative with n positive integer we define n-pseudoconvex functions being a generalization of the usual pseudoconvex functions. Again with the n-th order Dini derivative we define n-stationary points, and prove that a point x^0 is a global minimizer of a n-pseudoconvex function f if and only if x^0 is a n-stationary point of f. Our main result is the following. A radially continuous function f defined on a radially open convex set in a real linear space is n-pseudoconvex if and only if f is quasiconvex function and any n-stationary point is a global minimizer. This statement generalizes the results of Crouzeix, Ferland, Math. Program. 23 (1982), 193–205, and Komlósi, Math. Program. 26 (1983), 232–237. We study also other aspects of the n-pseudoconvex functions, for instance their relations to variational inequalities.

Key words: Pseudoconvex functions, n-pseudoconvex functions, stationary points, n-stationary points, quasiconvex functions.

1 Introduction

Within nonsmooth setting in terms of n-th order Dini directional derivative with n positive integer we define n-th order pseudoconvex (for short, n-pseudoconvex) functions generalizing in such a way the usual pseudoconvex functions. In fact, the class of 1-pseudoconvex functions coincides with the class of the nonsmooth pseudoconvex functions, and the class of n-pseudoconvex functions is strictly contained in the class of $(n+1)$-pseudoconvex functions. Again with the n-th order Dini derivative we define the notion of n-th order stationary (for short, n-stationary) point, and prove that a point x^0 is a global minimizer of a n-pseudoconvex function f if and only if x^0 is a n-stationary point of f. For a radially continuous function f defined on a radially open convex set in a real linear space we prove, that f is n-pseudoconvex if and only if f is quasiconvex function and any n-stationary point is a global minimizer. For smooth functions with open domain the characterization that a function is pseudoconvex if and only if it is quasiconvex and each stationary point is a global minimizer is obtained in Crouzeix, Ferland [5]. This statement has been extended to nonsmooth functions in Komlósi [12], where pseudoconvexity and stationary points of a nonsmooth function are defined through the first-order Dini derivative. Hence, we propose a generalization of the above results. We study also other aspects of n-pseudoconvex functions, for instance their relations to variational inequalities and the particular class of the twice continuously differentiable 2-pseudoconvex functions.

2 Dini Derivatives and n-Pseudoconvex Functions

In this paper $\mathbf{E}$ denotes a real linear space and $f : X \to \mathbb{R}$ a finite-valued real function defined on the set $X \subset \mathbf{E}$. Here $\mathbb{R}$ is the set of the reals and $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$. Let $\overline{f} : \mathbf{E} \to \mathbb{R} \cup \{+\infty\}$ be the extension of f such that $\overline{f}(x) = +\infty$ for $x \in \mathbf{E} \setminus X$.

The lower Dini directional derivative $f_-^{(1)}(x,u)$ of f at $x \in X$ in direction $u \in \mathbf{E}$ is defined as an element of $\overline{\mathbb{R}}$ by

$$f_-^{(1)}(x,u) = \liminf_{t \to +0} \frac{1}{t} (\overline{f}(x+tu) - f(x)) .$$

The difference in the right-hand side has sense in $\overline{\mathbb{R}}$, since eventually only $\overline{f}(x+tu)$ is infinite. Applying $\overline{f}$ instead of f we get some convenience in handling the case $x+tu \notin X$. Still, we may use $f(x+tu)$ instead of $\overline{f}(x+tu)$ in the case when 0 is an accumulating point of the set $\{t > 0 \mid x+tu \in X\}$. If this condition has place, we say that u is a feasible direction of X at x. As a consequence of the definition, if u is not a feasible direction of X at x we have $f_-^{(1)}(x,u) = +\infty$.

For a given positive integer $n > 1$ we accept that the n-th order lower Dini directional derivative $f_-^{(n)}(x,u)$ exists as an element of $\overline{\mathbb{R}}$ only if for $i = 1, \dots, n-1$ all the lower Dini directional derivatives $f_-^{(i)}(x,u)$ exist as elements of $\mathbb{R}$, and then we put

$$f_-^{(n)}(x,u) = \liminf_{t \to +0} \frac{n!}{t^n} \left(\overline{f}(x+tu) - \sum_{i=0}^{n-1} \frac{t^i}{i!} f_-^{(i)}(x,u) \right) .$$

In the sum we accept $f_-^{(0)}(x,u) = f(x)$. From this definition we see, that if the higher-order derivative $f_-^{(n)}(x,u)$ exists, then u is a feasible direction of X at x. Indeed, if u is not a feasible direction of X at x, then as it was said $f_-^{(1)}(x,u) = +\infty \notin \mathbb{R}$, whence from the definition the higher-order derivative $f_-^{(n)}(x,u)$ does not exist. For short we will say Dini derivatives instead of lower Dini directional derivatives. The above definition follows Ginchev [7], where in terms of the Dini derivatives higher-order optimality conditions for nonsmooth problems are derived.

The function $f : X \to \mathbb{R}$ is said to be pseudoconvex if x^0, $x^1 \in X$ and $f(x^0) < f(x^1)$ implies $f_-^{(1)}(x^1, x^0 - x^1) < 0$. We accept this definition after Diewert [6] as a convenient modification for nonsmooth functions of the classical definition [14] of a pseudoconvex function. Let us specially underline, that in opposite to the commonly accepted restriction, we do not assume in advance that the domain X of f is convex. Still, the given here definition of a pseudoconvex function gives some implicit restriction on X. Namely, if $x^1 \in X$ and there exists $x^0 \in X$ such that $f(x^0) < f(x^1)$, then $x^0 - x^1$ is a feasible direction of X at x^1. Indeed, now $f_-^{(1)}(x^1, x^0 - x^1) < 0$ and not $+\infty$.

We generalize the notion of a pseudoconvex functions as follows.

Definition 1. *For a positive integer n we call the function $f : X \to \mathbb{R}$ pseudoconvex of order n (for short, n-pseudoconvex) if for any x^0, $x^1 \in X$ such that $f(x^0) < f(x^1)$ there exists a positive integer $m \leq n$ such that $f_-^{(i)}(x^1, x^0 - x^1) = 0$ for all positive integers $i < m$ and $f_-^{(m)}(x^1, x^0 - x^1) < 0$.*

In this definition the derivative $f_-^{(m)}(x^1, x^0 - x^1)$ exists. When $m > 1$ this follows from the existence with values 0 of all the derivatives of lower order. The existence of n-th order Dini derivatives of f is not required.

We do not assume in advance that the domain X of f is convex. Still, as in the case of a pseudoconvex function we have the implicit restriction on X, that if $x^1 \in X$ and there exists $x^0 \in X$ such that $f(x^0) < f(x^1)$, then $x^0 - x^1$ is a feasible direction of X at x^1. Now $f_-^{(1)}(x^1, x^0 - x^1) \leq 0$ and not $+\infty$, as it would be if u were not feasible. With regard to this remark one can define the following class of sets, call them convex-like, as the natural sets to serve for domains of n-pseudoconvex functions. We call the set X in $\mathbf{E}$ convex-like, if for all x^0, $x^1 \in X$ the direction $x^0 - x^1$ is feasible of X at x^1. Obviously each convex set is convex-like. A convex-like set is also each radially open set, see the definition of a radially open set in Section 4. Some of the following results as Theorems 1 and 8 do not use convexity assumptions for X. However, the convexity of X plays a role in the most of the forthcoming considerations. Whenever this is the case, we say explicitly that X is convex. We think, that these results admit generalizations to convex-like sets. For simplicity, we do not discuss these possibilities.

In the above definition when $n = 1$ the only possible choice of the positive integer $m \leq n$ is $m = 1$. Since then the set of the positive integers $i < m$ is empty, the equalities claiming the vanishing of the i-th derivatives do not occur. Therefore the definition of 1-pseudoconvex function reduces to the definition of a pseudoconvex function, hence the classes of 1-pseudoconvex functions and pseudoconvex functions are identical.

It is obvious from the definition, that for a positive integer n each n-pseudoconvex function is $(n+1)$-pseudoconvex. The following example shows that this inclusion is strict.

Example 1. The function $f_n : \mathbb{R} \to \mathbb{R}$, n positive integer, defined by

$$f_n(x) = \begin{cases} x^n\,, \ x \geq 0\,, \\ (-1)^{n-1} x^n\,, \ x < 0\,. \end{cases}$$

is n-pseudoconvex, and when $n \geq 2$ it is not $(n-1)$-pseudoconvex.

For n odd the function f_n in Example 1 is the power function $f_n(x) = x^n$, $x \in \mathbb{R}$. For n even the function f_n is of class C^{n-1} but not of class C^n.

3 Stationary Points and n-Pseudoconvex Functions

We introduce the notion of a n-stationary point as follows.

Definition 2. *For a positive integer n we call $x \in X$ a stationary point of order n (for short, n-stationary point) of the function $f : X \to \mathbb{R}$, if for each direction $u \in \mathbf{E}$ and arbitrary positive integer $m \leq n$ the equalities*

$$f_-^{(i)}(x,u) = 0 \textit{ for all positive integers } i < m \textit{ imply } f_-^{(m)}(x,u) \geq 0.$$

Let us turn attention, that in particular for $m = 1$ it holds $f_-^{(1)}(x,u) \geq 0$ for all $u \in \mathbf{E}$ which is a consequence of the nonexistence of positive integers $i < 1$.

The notion of 1-stationary point coincides with the usual notion of a stationary point applied in minimization of nonsmooth functions, e. g. in [9] and [12]. Stationary points of order two, usually in a smooth aspect, are used in the literature. The definition of n-stationary points for nonsmooth functions given here seems to be a new one.

It is obvious, that for $n \geq 2$ each n-stationary point is $(n-1)$-stationary point. Example 1 shows that the converse is not true, there the point $x = 0$ is $(n-1)$-stationary but not n-stationary.

The following theorem characterizes the global minimizers of n-pseudoconvex functions in terms of n-stationary points.

Theorem 1. *Let $f : X \to \mathbb{R}$ be a n-pseudoconvex function with n positive integer. Then $x^0 \in X$ is a global minimizer of f if and only if x^0 is a n-stationary point of f.*

Proof. Suppose that $x^0 \in X$ is a global minimizer of f. Take arbitrary direction u and let $m \leq n$ be a positive integer. Assume that $f_-^{(i)}(x^0,u) = 0$ for all positive integers $i < m$. Then

$$f_-^{(m)}(x^0,u) = \liminf_{t \to +0} \frac{m!}{t^m} \left(\overline{f}(x^0 + tu) - f(x^0) \right) \geq 0\,.$$

Therefore x^0 is a n-stationary point. Moreover, replacing in the inequality $m \leq n$ the number n with arbitrary positive integer n_0 and repeating the above reasonings, we see that x^0 is a n_0-stationary point for arbitrary positive integer n_0.

Conversely, assume that x^0 is a n-stationary point of f. Suppose in the contrary, that x^0 is not a global minimizer of x^0, hence there exists $x^1 \in X$ such that $f(x^1) < f(x^0)$. Since f is n-pseudoconvex, therefore there exists a positive integer $m \le n$ such that $f_-^{(i)}(x^0, x^1 - x^0) = 0$ for all positive integers $i < m$ and $f_-^{(m)}(x^0, x^1 - x^0) < 0$. Therefore, the condition in the definition of a n-stationary point fails with $x = x^0$ and $u = x^1 - x^0$, a contradiction. □

4 Characterization of n-Pseudoconvex Functions

We deal with functions $f : X \to \mathbb{R}$, where X is a subset of the real linear space $\mathbf{E}$. No topological structure on $\mathbf{E}$ is assumed. At the same time topological notions can be involved, when we consider restrictions on straight lines in $\mathbf{E}$. For such notions we apply the adjective radial.

For any x^0, $x^1 \in \mathbf{E}$ we put $X(x^0, x^1) = \{t \in \mathbb{R} \mid (1-t)x^0 + tx^1 \in X\}$. We say, that the set X is radially open (radially closed) if for any x^0, $x^1 \in \mathbf{E}$ the set $X(x^0, x^1)$ is open (closed) in $\mathbb{R}$.

Further we use the abbreviations lsc and usc respectively for lower semi-continuous and upper semi-continuous. We say that the function $f : X \to \mathbb{R}$ is radially lsc (radially usc) if the function

$$\varphi : X(x^0, x^1) \to \mathbb{R}\,, \quad \varphi(t) = f((1-t)x^0 + tx^1)$$

is lsc (usc) for all x^0, $x^1 \in X$. If f is both radially lsc and radially usc, we say that f is radially continuous. Let us underline, that the definition given requires that φ is lsc (usc) with respect to the relative topology on $X(x^0, x^1) \subset \mathbb{R}$. This remark is to avoid confusions of the following type. Let $\mathbf{E} = \mathbb{R}$ and $X \subset \mathbb{R}$ be an open proper set in $\mathbb{R}$ (X proper in $\mathbb{R}$ means $X \ne \emptyset$ and $X \ne \mathbb{R}$). Consider the function $f : X \to \mathbb{R}$, $f(x) = 0$. Then f is (radially) lsc. At the same time the function $\bar{f}$ being the indicator function of the open set X is not (radially) lsc as a function from $\mathbb{R}$ into $\overline{\mathbb{R}}$.

A mean value theorem for lsc functions appears in Diewert [6]. In the following precise formulation it is proved in Crespi, Ginchev, Rocca [3] and is used there to study nonsmooth variational inequalities.

Theorem 2 (Mean Value Theorem). *Let $\varphi : [t_0,\, t_1] \to \mathbb{R} \cup \{+\infty\}$ be a lsc function of a real variable, such that $\varphi(t_0) \ne +\infty$. Then there exists a real ξ, $t_0 < \xi \le t_1$, such that $\varphi_-^{(1)}(\xi, t_0 - t_1) \ge \varphi(t_0) - \varphi(t_1)$. Moreover, this inequality holds at the point ξ, $t_0 < \xi \le t_1$, which supplies the minimum of the function $\psi : [t_0,\, t_1] \to \mathbb{R} \cup \{+\infty\}$ defined by*

$$\psi(t) = \varphi(t) - \frac{t_1 - t}{t_1 - t_0}\,\varphi(t_0) - \frac{t - t_0}{t_1 - t_0}\,\varphi(t_1)\,. \tag{1}$$

The following theorem applies the mean value theorem to establish monotonicity properties of the lsc n-pseudoconvex functions.

Theorem 3. *Let n be a positive integer and let $\varphi : I \to \mathbb{R}$ be a lsc n-pseudoconvex function defined on the interval $I \subset \mathbb{R}$. Denote by $\hat{I}$ the set of the global minimizers of φ on I. Then either $\hat{I} = \emptyset$, in which case φ is strictly monotone on I, or $\hat{I}$ is a nonempty interval, in which case φ is a constant on $\hat{I}$, strictly increasing on $I_+ = \{t \in I \mid t \geq \hat{t}$ for all $\hat{t} \in \hat{I}\}$ and strictly decreasing on $I_- = \{t \in I \mid t \leq \hat{t}$ for all $\hat{t} \in \hat{I}\}$.*

Proof. *We prove, that if $t_- < t_+$ with $t_-, t_+ \in I$, and $\varphi(t_-) < \varphi(t_+)$, then φ is strictly increasing on the interval $I \cap [t_+, +\infty)$.* Assume in the contrary, that this is not true. Then there exists $\bar{t} \in I \cap (t_+, +\infty)$ such that φ is not strictly increasing on the interval $[t_+, \bar{t}]$. Denote by t_0 the supremum of the set M of all $t \in [t_+, \bar{t}]$ such that φ is strictly increasing on the interval $[t_+, t]$. *We will show some properties of* t_0. Since $t_+ \in M$, obviously $M \neq \emptyset$ and $t_0 \geq t_+$. Also $t_0 \in I$, since $t_0 \in [t_+, \bar{t}] \subset I$. From the definition of t_0, φ *is strictly increasing on the interval* $[t_+, t_0)$. We will show more, that φ *is strictly increasing on the closed interval* $[t_+, t_0]$. To show this, it is enough to prove that $\varphi(t) \leq \varphi(t_0)$ *for all* $t \in [t_+, t_0)$. Indeed, if this is true, and $t_+ \leq t < t_0$, then for $t' \in (t, t_0)$ we would have $\varphi(t) < \varphi(t') \leq \varphi(t_0)$, which shows that φ is strictly increasing on the interval $[t_+, t_0]$. *Now we show that* $\varphi(t) \leq \varphi(t_0)$ *for all* $t \in [t_+, t_0]$. Assume that this is not true. Then $t_+ < t_0$ (if $t_+ = t_0$ then the interval $[t_+, t_0)$ is empty, hence we would have the truthfulness of the claimed property), and $\varphi(\bar{t}_+) > \varphi(t_0)$ for some $t_+ \leq \bar{t}_+ < t_0$. Since φ is n-pseudoconvex, there exists $m \leq n$, such that $\varphi_-^{(i)}(\bar{t}_+, t_0 - \bar{t}_+) = 0$ for all positive integer $i < m$ and $\varphi_-^{(m)}(\bar{t}_+, t_0 - \bar{t}_+) < 0$. On the other hand, since φ is increasing on $[\bar{t}_+, t_0)$, we have

$$\varphi_-^{(m)}(\bar{t}_+, t_0 - \bar{t}_+) = \liminf_{\tau \to 0^+} \frac{m!}{\tau^m} \left(\varphi(\bar{t}_+ + \tau(t_0 - t_+)) - \varphi(\bar{t}_+)\right) \geq 0\,,$$

a contradiction. *Thus, φ is strictly increasing on the interval* $[t_+, t_0]$, whence with respect to the assumptions for $\bar{t}$, we have $t_0 < \bar{t}$. Hence, *the interval* $I \cap [t_0, +\infty)$ *contains the non-degenerate interval* $[t_0, \bar{t}]$.

Resuming, from the assumption that φ is not strictly increasing on the interval $I \cap [t_+, +\infty)$, it follows that there exists $t_0 \geq t_+$ with the properties: $t_0 \in I$ and $I \cap [t_0, +\infty)$ contains a non-degenerate interval $[t_0, \bar{t}]$; the function φ is strictly increasing on the interval $[t_+, t_0]$; on any larger interval $[t_+, t]$, $t > t_0$, the function φ is not strictly increasing.

We will show that the above properties lead to a contradiction. Since φ is lsc, we can choose $t_1 > t_0$, such that $\varphi(t_0) \geq \varphi(t_1)$ and still $\varphi(t_-) < \varphi(t)$ for any $t \in [t_0, t_1]$. Moreover, we can choose t_1 with these properties in a way that the following alternative holds: either $\varphi(t_0) > \varphi(t_1)$ or $\varphi(t_0) = \varphi(t_1) = \min\{\varphi(t) \mid t_0 \leq t \leq t_1\}$. Take now ξ, $t_0 < \xi \leq t_1$, such that ξ supplies the minimum of the function $\psi : [t_0, t_1] \to \mathbb{R}$ in (1), whence in particular

$$\varphi(\xi) \le \frac{t_1 - \xi}{t_1 - t_0}\, \varphi(t_0) + \frac{\xi - t_0}{t_1 - t_0}\, \varphi(t_1) \le \varphi(t_0)\,.$$

The Mean Value Theorem gives $\varphi_-^{(1)}(\xi, t_0 - t_1) \ge \varphi(t_0) - \varphi(t_1)$. On the other hand $\varphi(t_-) < \varphi(\xi)$, whence from the n-pseudoconvexity of φ, there exists $m \le n$, such that $\varphi_-^{(i)}(\xi,\, t_- - \xi) = 0$ for all positive integers $i < m$ and $\varphi_-^{(m)}(\xi,\, t_- - \xi) < 0$. By the choice of t_1 the following cases may occur:

1^0. $\varphi(t_0) > \varphi(t_1)$.

This leads immediately to a contradiction, since the positive homogeneity of the Dini derivative gives

$$0 \ge \varphi_-^{(1)}(\xi,\, t_- - \xi) = \frac{\xi - t_-}{t_1 - t_0}\, \varphi_-^{(1)}(\xi,\, t_0 - t_1) \ge \frac{\xi - t_-}{t_1 - t_0}\, (\, \varphi(t_0) - \varphi(t_1)\,) > 0\,.$$

2^0. $\varphi(t_0) = \varphi(t_1) = \min\{\varphi(t) \mid t_0 \le t \le t_1\}$.

Now $\psi(t) = \varphi(t) - \varphi(t_0)$, and the point ξ, $t_0 < \xi \le t_1$, supplies the minimum of φ on the interval $[t_0,\, t_1]$. We obtain easily the contradiction

$$0 > \varphi_-^{(m)}(\xi,\, t_- - \xi) = \liminf_{s \to +0} \frac{m!}{s^m}\, (\varphi(\xi + s(t_- - \xi)) - \varphi(\xi)) \ge 0\,.$$

Similarly, if $t_- < t_+$ with $t_-,\, t_+ \in I$, and $\varphi(t_-) > \varphi(t_+)$, then φ is strictly decreasing on the interval $I \cap (-\infty,\, t_-]$. This case is reduced to the already proved by changing the variable t to $-t$.

The set $\hat{I}$ is an interval, possibly empty. We can show this by proving, that for any $t_- < t_+$ with $t_-,\, t_+ \in \hat{I}$ it holds $[t_-,\, t_+] \subset \hat{I}$. If this is not the case, then there exists a point t_0, $t_- < t_0 < t_+$, such that $\varphi(t_-) < \varphi(t_0)$. Since φ is strictly increasing on $I \cap [t_0,\, +\infty)$, we see that $\varphi(t_-) < \varphi(t_0) < \varphi(t_+)$. On the other hand, since both t_- and t_+ are global minimizers, of φ, we have $\varphi(t_-) = \varphi(t_+)$, a contradiction.

If $\hat{I} \ne \emptyset$, then φ is strictly increasing on I_+ and strictly decreasing on I_-. We prove the first assertion, the proof of the second is similar. The case when $I_+ = \emptyset$ or I_+ is a singleton is obvious. Suppose now that I_+ is neither empty nor a singleton. Let $t_- \in \hat{I}$ and $t_+ \in I_+$ be such that $t_+ > \inf I_+$. We have $t_+ \notin \hat{I}$, otherwise the whole interval $[t_-,\, t_+]$ would be contained in $\hat{I}$ and the strict inequality $t_+ > \inf I_+$ would not be satisfied. Therefore $\varphi(t_-) < \varphi(t_+)$, whence φ is strictly increasing on the interval $I \cap [t_+,\, +\infty)$. Since the strict monotonicity holds for any $t_+ > \inf I_+$ and φ is lsc, we see that φ is strictly increasing on I_+.

If $\hat{I} = \emptyset$, then φ is strictly increasing or strictly decreasing. Choose in I a sequence $\{t_n\}_{n=0}^{\infty}$ such that $\varphi(t_0) > \varphi(t_1) > \dots > \varphi(t_n) > \dots$, and $\lim_n \varphi(t_n) = \inf\{\varphi(t) \mid t \in I\}$.

Suppose that $t_0 > t_1$ (the case $t_0 < t_1$ is considered similarly). We prove that then *$\{t_n\}$ is a strictly decreasing sequence and the function φ is strictly* *increasing on* I (when $t_0 < t_1$ the sequence $\{t_n\}$ is strictly increasing and the function φ is strictly decreasing on I). To see that $\{t_n\}$ is strictly decreasing, it is enough to show that $t_2 < t_1$ (similarly $t_2 < t_1$ implies $t_3 < t_2$ etc.). Indeed, it holds $t_2 \notin [t_0, +\infty)$, since for $t \in I \cap [t_0, +\infty)$, we have $\varphi(t_0) \le \varphi(t)$ (the inequality $\varphi(t_1) < \varphi(t_0)$ implies that φ is strictly increasing on $I \cap [t_0, +\infty)$), while by choice $\varphi(t_2) < \varphi(t_0)$. Also $t_2 \notin [t_1, t_0]$. Otherwise, if $t_1 < t_2 < t_0$, from $\varphi(t_1) > \varphi(t_2) < \varphi(t_0)$, it would follow that φ is strictly decreasing on $I \cap (-\infty, t_1]$ and it is strictly increasing on $I \cap [t_0, +\infty)$. Therefore, since $\hat{I}$ is the set of global minimizers of φ, we have $\hat{I} = \{\xi \in [t_1, t_0] \mid \varphi(\xi) = \min_{t_1 \le t \le t_0} \varphi(t)\}$. The contradiction is that once $\hat{I} = \emptyset$ by assumption, and otherwise the set in the right hand side is not empty, since any lsc functions achieves its minimum on a compact interval.

Now $\varphi(t_{n+1}) < \varphi(t_n)$ implies that φ is strictly increasing on the interval $I \cap [t_n, +\infty)$, hence also on the set $J = \bigcup_{n=1}^{\infty} (I \cap [t_n, +\infty))$. It remains to show that $J = I$. If this is not the case, then there exists a point $t_- \in I \setminus J$, whence $t_- < t_n$ for $n = 0, 1, \dots$. We have $\varphi(t_-) \le \varphi(t_n)$ for all n. Otherwise for some n we would have $\varphi(t_-) > \varphi(t_{n+1}) < \varphi(t_n)$, which as we have observed leads to a contradiction with $\hat{I} = \emptyset$. Now $\varphi(t_-) \le \varphi(t_n)$ for all n implies that $\varphi(t_-) \le \inf\{\varphi(t) \mid t \in I\}$, whence $t_- \in \hat{I}$, which contradicts again to $\hat{I} = \emptyset$. □

The assumption that φ is lsc is essential for the validity of Theorem 3. This is seen in the following example borrowed from Crespi, Ginchev, Rocca [3].

Example 2. Define the function $\varphi : \mathbb{R} \to \mathbb{R}$ by $\varphi(x) = 0$ for $x = 0$ or x irrational, and $\varphi(x) = -q$ for $x \ne 0$ rational with $x = p/q$, $q > 0$ and p and q relatively prime. The function φ is not lsc. The first-order Dini derivative is $\varphi_-^{(1)}(x, u) = -\infty$ for each $x \in \mathbb{R}$ and $u \in \mathbb{R} \setminus \{0\}$. Therefore φ is 1-pseudoconvex. For the set $\hat{I}$ of the global minimizers of φ we have $\hat{I} = \emptyset$. At the same time φ is neither increasing nor decreasing.

Recall that a function $f : X \to \mathbb{R}$ is said to be quasiconvex if its domain X is convex and for all x^0, $x^1 \in X$ and $t \in [0, 1]$ it holds

$$f((1-t)x^0 + tx^1) \le \max\left(f(x^0),\, f(x^1)\right) .$$

The function $f : X \to \mathbb{R}$ is called semistrictly quasiconvex if its domain X is convex and for all x^0, $x^1 \in X$ such that $f(x^0) \ne f(x^1)$ and $t \in (0, 1)$ the strict inequality $f((1-t)x^0 + tx^1) < \max\left(f(x^0),\, f(x^1)\right)$ is satisfied.

In general a semistrictly quasiconvex function is not quasiconvex, but each semistrictly quasiconvex radially lsc function is quasiconvex [11].

The following theorem generalizes a well-known relation between pseudoconvex and quasiconvex functions, see Diewert [6].

Theorem 4. *Let $f : X \to \mathbb{R}$ be radially lsc and n-pseudoconvex function with n positive integer on the convex set X in a real linear space. Then f is quasiconvex, and moreover, f is semistrictly quasiconvex.*

Proof. We prove that f is quasiconvex. Let $x^0,\, x^1 \in X$ be such that $f(x^0) \leq f(x^1)$. Put $x(t) = (1-t)x^0 + tx^1$. Define the function $\varphi : [0,\, 1] \to \mathbb{R}$, $\varphi(t) = f(x(t))$. We must show that for each $\bar{t} \in (0,\, 1)$ it holds $\varphi(\bar{t}) \leq \varphi(1)$. The function φ is lsc and n-pseudoconvex on $I = [0,\, 1]$. Therefore φ satisfies the hypotheses of Theorem 3. Since I is a compact interval, the set $\hat{I} \subset I$ of the global minimizers of φ is a nonempty compact interval $[t_-,\, t_+]$. The function φ strictly decreases on the interval $I_- = [0,\, t_-]$ and strictly increases on the interval $I_+ = [t_+,\, 1]$. Now our claim follows observing that for $\bar{t}$ we have the possibilities:

1^0. $\bar{t} \in \hat{I}$, then (from $\hat{I}$ the set of global minimizers) $\varphi(\bar{t}) \leq \varphi(1)$;
2^0. $\bar{t} \in I_- \backslash \{0\}$, then (from φ strictly decreasing on I_-) $\varphi(\bar{t}) < \varphi(0) \leq \varphi(1)$;
3^0. $\bar{t} \in I_+ \setminus \{1\}$, then (from φ strictly increasing on I_+) $\varphi(\bar{t}) < \varphi(1)$.

To show the semistrict quasiconvexity of f, we must show in the above notations that $\varphi(0) < \varphi(1)$ implies $\varphi(\bar{t}) < \varphi(1)$. This can be shown by repeating up to obvious small changes the same reasonings. For instance, in the case 1^0 we have $\varphi(\bar{t}) \leq \varphi(0) < \varphi(1)$. □

The following example shows that Theorem 4 cannot be reverted.

Example 3. The function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} \exp(-\frac{1}{x}),\; x > 0, \\ 0,\; x = 0, \\ -\exp(\frac{1}{x}),\; x < 0, \end{cases}$$

is semistrictly quasiconvex, but for any positive integer n it is not n-pseudoconvex.

Our main result is the following characterization of the radially continuous n-pseudoconvex functions with radially open convex domains.

Theorem 5. *Let $f : X \to \mathbb{R}$ be a radially continuous function on the radially open convex set X in a real linear space. Then f is n-pseudoconvex with n positive integer if and only if f is quasiconvex and each n-stationary point of f is a global minimizer.*

Proof. The necessity holds under the weaker assumption that f is radially lsc function on the convex set X. Let f be n-pseudoconvex. Then f is quasiconvex according to Theorem 4. Further, each n-stationary point of f is a global minimizer according to Theorem 1.

The sufficiency holds under the weaker assumption that f is radially usc on the radially open convex set X. Let f be quasiconvex and let each n-stationary point of f be a global minimizer. We prove that f is n-pseudoconvex. Take $x^0,\, x^1 \in X$ with $f(x^0) < f(x^1)$. Hence x^1 is not a global minimizer. By the hypotheses x^1 is not a n-stationary point. Therefore there exists a positive integer $m \leq n$ and a direction v such that $f_-^{(i)}(x^1, v) = 0$ for all positive integers $i < m$ and $f_-^{(m)}(x^1, v) < 0$. We may assume that m is the minimal with the above property, that is for each $k < m$ and arbitrary direction u the inequalities $f_-^{(i)}(x^1, u) = 0$ for all positive integers $i < k$ imply $f_-^{(k)}(x^1, u) \geq 0$. Now we show that $f_-^{(i)}(x^1, x^0 - x^1) = 0$ for $i < m$ and $f_-^{(m)}(x^1, x^0 - x^1) < 0$. Indeed, by the quasiconvexity of f we have

$$f_-^{(1)}(x^1, x^0 - x^1) = \liminf_{t \to +0} t^{-1}(f(x^1 + t(x^0 - x^1)) - f(x^1)) \leq 0.$$

Using the minimality of m and induction we conclude that $f_-^{(i)}(x^1, x^0 - x^1) = 0$ for all $i < m$. Since f is radially upper semicontinuous and X is radially open, there exists $\tau > 0$ such that $f(p) < f(x^1)$ where $p = x^0 - \tau v$. Put $z(t) = x^1 + t(x^0 - x^1)$ with $t \in (0,1)$. Let $w(t) = x^1 + \alpha(t)\, v$ be the point of intersection of the ray $\{x^1 + tv \mid t \geq 0\}$ and the straight line passing through p and $z(t)$. An easy calculation gives that $\alpha(t) = t\tau/(1-t)$. Since f is quasiconvex, we have

$$f(z(t)) \leq \max\,(f(p),\, f(w(t))) \quad \text{for} \quad 0 < t < 1\,.$$

Therefore

$$t^{-m}(f(z(t)) - f(x^1)) \leq \max\,(t^{-m}(f(p) - f(x^1)),\, t^{-m}(f(w(t)) - f(x^1))).$$

Since $f(p) < f(x^1)$, if t tends to 0 with positive values, then the first term of the above maximum tends to $-\infty$. Let the sequence α_n be such that

$$f_-^{(m)}(x, v) = \liminf_{n \to \infty} m!\, \alpha_n^{-m}(f(x^1 + \alpha_n v) - f(x^1))$$

and $t_n = \alpha_n/(\alpha_n + \tau)$. Therefore

$$\begin{aligned} f_-^{(m)}(x^1, x^0 - x^1) &= \liminf_{t \to +0} m!\, t^{-m}(f(z(t)) - f(x^1)) \\ &\leq \liminf_{n \to 0} m!\, t_n^{-m}(f(w(t_n)) - f(x^1)). \end{aligned} \tag{2}$$

Using the equality

$$\frac{f(w(t)) - f(x^1)}{t^m} = \frac{f(x^1 + \alpha(t)v) - f(x^1)}{\alpha^m(t)} \cdot \frac{\alpha^m(t)}{t^m},$$

where $\alpha(t) = t\tau/(1-t)$ we get that

$$\begin{aligned} &\liminf_{n \to \infty} m!\, t_n^{-m}(f(w(t_n)) - f(x_1))) \\ &= \liminf_{n \to \infty} m!\, \alpha_n^{-m}(f(x^1 + \alpha_n v) - f(x^1))\, \alpha_n^m t_n^{-m} = \tau^m f_-^{(m)}(x^1, v) < 0. \end{aligned}$$

The above inequality along with (2) yields that f is n-pseudoconvex. □

The close relation between pseudoconvex and quasiconvex functions is underlined in Crouzeix [4]. In Crouzeix, Ferland [5] it is shown that a smooth function on an open convex set is pseudoconvex if and only if it is quasiconvex and attains a minimum at each stationary point. This characterization is extended to nonsmooth functions by Komlósi [12] and Giorgi, Komlósi [9]. The Dini derivatives are the main tool in this extension and the coincidence of the stationary points and the global minimizers is shown. Later this characterization is analysed in Giorgi [8], Tanaka [15], Aussel [1] and Ivanov [10]. Theorem 5 above is a generalization of the results mentioned.

The proof of the Sufficiency in Theorem 5 is made under the assumption that f is radially usc and X is radially open and convex, while the Necessity is done for radially lsc functions and does not apply the radially openness of X. The following two examples show that the Sufficiency fails to be true for radially lsc functions, or if the radially openness of X is not assumed.

Example 4. Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$f(x_1, x_2) = \begin{cases} \arctan(x_2/x_1)\,, & x_1 > 0\,, \\ \pi/2\,, \; x_1 = 0, \; x_2 > 0\,, & \\ -\pi/2\,, \; x_1 = 0, \; x_2 \le 0\,, & \\ \pi/2 - x_1\,, & x_1 < 0\,. \end{cases}$$

Then f is lsc and quasiconvex on $\mathbb{R}^2$, that is on an open set, and each 1-stationary point is a global minimizer, but f is not 1-pseudoconvex.

It is easy to check that each level set $L(f, r) = \{(x_1, x_2) \in \mathbb{R}^2 \mid f(x_1, x_2) \le r\}$ is closed and convex, whence f is lsc and quasiconvex. Take the points $x^0 = (0, 0)$ and $x^1 = (0, 1)$. We have $f(x^0) = -\pi/2 < \pi/2 = f(x^1)$ and $f_-^{(1)}(x^1, x^0 - x^1) = 0$, hence f is not 1-pseudoconvex. The set of the global minimizers is $\{x = (0, x_2) \mid x_2 \le 0\}$ and coincides with the set of 1-stationary points. To check this we put $x = (x_1, x_2)$, $u = (u_1, u_2)$, and observe that $f_-^{(1)}(x, u) = -1$ in each of the cases: if $x_1 = 0$, $x_2 > 0$, $u_1 = x_2$, $u_2 = 0$; if $x_1 > 0$, $u_1 = 0$, $u_2 = -(x_1^2 + x_2^2)/x_1$; and if $x_1 < 0$, $u_1 = 1$, $u_2 = 0$.

Example 5. Consider the function $f : X \to \mathbb{R}$, where

$$X = \{x = (x_1, x_2) \in \mathbb{R}^2 \mid 0 \le x_1 \le 1, \; 0 \le x_2 \le 1 - x_1\}$$

and

$$f(x_1, x_2) = 1 - x_2 - \frac{1}{2}\,x_1^2 - x_1\sqrt{x_2 + \frac{1}{4}\,x_1^2}\,.$$

Then f is continuous and quasiconvex on X, which is a not radially open set, and each 1-stationary point is a global minimizer, but f is not 1-pseudoconvex.

In this example the level sets $L(f, r) = \{(x_1, x_2) \in \mathbb{R}^2 \mid f(x_1, x_2) \le r\}$ are the convex sets $L(f, r) = \emptyset$ for $r < 0$, $L(f, r) = X$ for $r > 1$ and

$$L(f,r) = \{(x_1,x_2) \in X \mid x_1\sqrt{1-r} + x_2 \geq 1-r\} \quad \text{for} \quad 0 \leq r \leq 1\,,$$

whence f is quasiconvex. Obviously f is also continuous. The function f is not 1-pseudoconvex, since for $x^1 = (0,\,0)$ and $x^0 = (1,\,0)$ we have $f(x^1) = 1 > 0 = f(x^0)$ and at the same time $f_-^{(1)}(x^1, x^0 - x^1) = 0$. The point x^1 is not 1-stationary point, since $f_-^{(1)}(x^1, (0,\,1)) = -1$. The set of the global minimizers is $\{(x_1,x_2) \in X \mid x_1 + x_2 = 1\}$, where f attends value zero, and this set coincides with the 1-stationary points. To show this it remains to observe that for $(x_1,x_2) \in X$ with $x_1 + x_2 \notin \{0,\,1\}$ we have

$$f_-^{(1)}((x_1,x_2),(0,\,1)) = -1 - \frac{x_1}{2\sqrt{x_2 + \frac{1}{4}\,x_1^2}} < 0\ .$$

Though we consider functions f defined on an arbitrary real linear space, the n-pseudoconvexity is in fact one dimensional notion in the sense, that it is determined by a property, which is supposed to hold on one-dimensional affine manifolds. That is why the most phenomena occurring for n-pseudoconvex functions we observe on one-dimensional examples. Theorem 5 is however of another nature due to the involvement of the concept of n-stationary point, which is not one-dimensional. This difference causes that the assumptions of Theorems 4 cannot be kept in Theorem 5, as it has been shown in the Examples 4 and 5. In the next Theorem 6 applying one-dimensional n-stationarity, we are still able to prove the sufficiency within the hypotheses of Theorem 4, without the additional assumptions f to be radially usc and X to be radially open.

Theorem 6. *Let $f : X \to \mathbb{R}$ be quasiconvex function defined on the convex set X in a real linear space and let f obey the following property: For arbitrary x^0, $x^1 \in X$ such that $f(x^0) < f(x^1)$ the point $t_1 = 1$ is not n-stationary point with n positive integer for the function of one variable $\varphi : X(x^0,x^1) \to \mathbb{R}$, $\varphi(t) = f(x(t))$, where $x(t) = (1-t)x^0 + tx^1$ and $X(x^0,x^1) = \{t \in \mathbb{R} \mid x(t) \in X\}$. Then f is n-pseudoconvex.*

Proof. Let x^0, $x^1 \in X$ be arbitrary such that $f(x^0) < f(x^1)$. Since X is convex, then the closed interval $[0,1]$ belongs to the domain of φ. If $t = 1$ is boundary of the domain of φ, then $\varphi_-^{(1)}(1,1) = +\infty$. By the hypothesis of the theorem there exists integer $m \leq n$ such that $f_-^{(i)}(x^1, x^0 - x^1) = 0$ for $i < m$ and $f_-^{(m)}(x^1, x^0 - x^1) < 0$. Consider the case when $t = 1$ is an interior point of the domain of φ. For all sufficiently small values $t > 0$, by quasiconvexity, $\varphi(1) \leq \max(\varphi(0), \varphi(1+t))$. Using that $\varphi(0) < \varphi(1)$ we get the inequality $\varphi(1+t) \geq \varphi(1)$. Therefore for every integer i equality $\varphi_-^{(i-1)}(1,1) = 0$ implies $\varphi_-^{(i)}(1,1) \geq 0$. Since $t = 1$ is not a n-stationary point, then there exists integer $m \leq n$ such that $f_-^{(i)}(x^1, x^0 - x^1) = 0$ for $i < m$ and $f_-^{(m)}(x^1, x^0 - x^1) < 0$. Therefore f is n-pseudoconvex. $\square$

In Example 1 the n-pseudoconvexity of f_n can be established from the definition of this notion, but also directly from Theorem 5 (or Theorem 6). Indeed, f_n is continuous and quasiconvex and the set of its n-stationary points is empty (hence we can endorse the elements of this empty set with the property that they are global minimizers). From the Sufficiency of Theorem 5 we establish that f_n is n-pseudoconvex.

5 Twice Continuously Differentiable Functions

For $n > 1$ odd the function $f_n(x) = x^n$, $x \in \mathbb{R}$, is of class C^{n}, and as it was noticed in Example 1 it is n-pseudoconvex and not $(n-1)$-pseudoconvex. For an even n the function f_n is not of class C^{n}. Therefore, one can pose the question, whether if n is even there exists a n-pseudoconvex function defined on an open set in a finite-dimensional Euclidean space, which is C^{n} and not $(n-1)$-pseudoconvex. In this section we show that the answer is negative when $n = 2$. We show in fact, that the set of the 2-pseudoconvex twice continuously differentiable functions defined on an open set in the finite-dimensional Euclidean space $\mathbb{R}^n$ coincides with the set of the pseudoconvex twice continuously differentiable functions. For such a function the Dini derivatives of order one and two coincide with the usual directional derivatives of first and second order and are expressed through the Jacobian $f_-^{(1)}(x,u) = f'(x)u$ and the Hessian $f_-^{(2)}(x,u) = f''(x)(u,u)$. With this remark, the following lemma becomes obvious.

Lemma 1. *Let $X \subset \mathbb{R}^n$ be an open set and $f \in \mathrm{C}^2(\mathrm{X})$. Then $x \in X$ is a 2-stationary point of f if and only if the following conditions hold:*

$$f'(x) = 0\,, \tag{3}$$

$$f''(x)(u,u) \geq 0 \quad \textit{for all} \quad u \in \mathbb{R}^n. \tag{4}$$

The following claim gives characterizations of 2-stationary points of a quasiconvex function.

Lemma 2. *Let $f \in \mathrm{C}^2(\mathrm{X})$ be a quasiconvex finite-valued function defined on the open set $X \subset \mathbb{R}^n$. Then $x \in X$ is a 2-stationary point of f if and only if it is a stationary point.*

Proof. It is obviously that each 2-stationary point is a stationary point. Conversely, suppose that $x \in X$ is a stationary point, which means that x satisfies assumption (3). Thanks to quasiconvexity x fulfills condition (4), since for twice continuously differentiable quasiconvex function as it is proved in Avriel [2] it holds

$$f'(x)u = 0 \quad \text{implies} \quad f''(x)(u,u) \geq 0. \tag{5}$$

This completes the proof. □

The following theorem characterizes the 2-pseudoconvex twice continuously differentiable functions.

Theorem 7. *Let $X \subset \mathbb{R}^n$ be an open convex set and $f \in \mathrm{C}^2(\mathrm{X})$. Then f is 2-pseudoconvex on X if and only if it is pseudoconvex on X.*

Proof. It is obvious that each twice continuously differentiable pseudoconvex function is 2-pseudoconvex.

Conversely, assume that f is 2-pseudoconvex. We show that it is pseudoconvex. According to Theorem 4 f is quasiconvex. It follows from Theorem 1 that the set of global minimizers of f coincides with the set of 2-stationary points. Taking into account that X is open, by Lemma 2, we conclude that the set of the stationary points coincides with the set of 2-stationary points. Therefore the set of the global minimizers coincides with the set of stationary points. Hence from the Theorem of Crouzeix, Ferland [5], which is particular case of our Theorem 5, we obtain that f is pseudoconvex. □

Let us underline, that the conclusion of Theorem 7 fails to be true in the class of C^1 functions, where the class of 2-pseudoconvex functions is larger than the class of pseudoconvex functions. For instance the function f_2 in Example 1 is C^1 and 2-pseudoconvex, but it is not pseudoconvex.

6 Pseudoconvex of Infinite Order Functions

The classical pseudoconvex functions are defined as C^1 functions, for which the directional derivative $f'(x^1)(x^0 - x^1)$ is negative each time when $f(x^0) < f(x^1)$. Remaining in the classical setting, we would have to call 2-pseudoconvex functions each C^2 function, for which each time when $f(x^0) < f(x^1)$ we would have either $f'(x^1)(x^0 - x^1) < 0$ or $f'(x^1)(x^0 - x^1) = 0$ and $f''(x^1)(x^0 - x^1, x^0 - x^1) < 0$. Then, as it is seen from the results of the previous section, we would have that the class of the "classical" 2-pseudoconvex functions is essentially smaller than the class of the pseudoconvex functions. The reason is that the latter contains C^1 functions, which are not C^2 functions. This is seen on the following example.

Example 6. The function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} x^2\,, \; x > 0\,, \\ \;0\,, \; x \le 0\,, \end{cases}$$

is C^1 and pseudoconvex, but not C^2 function.

In opposite to the "classical approach", in the framework of the "nonsmooth approach" accepted in this paper, we get strictly increasing classes of functions $\mathcal{F}_n \subset \mathcal{F}_{n+1}$, $n = 1, 2, \ldots$. Here $\mathcal{F}_n$ denotes the class of n-pseudoconvex functions as defined in Definition 1.

Wishing to extend the class of functions being in some sense similar to n-pseudoconvex functions defined in the paper, we come easily to the definition of pseudoconvex of infinite order functions.

Definition 3. *We call the function $f : X \to \mathbb{R}$ pseudoconvex of infinite order (for short, $+\infty$-pseudoconvex) if for each x^0, $x^1 \in X$ such that $f(x^0) < f(x^1)$ there exists a positive integer m such that $f_-^{(i)}(x^1, x^0 - x^1) = 0$ for all positive integers $i < m$ and $f_-^{(m)}(x^1, x^0 - x^1) < 0$.*

Obviously, each n-pseudoconvex function is $+\infty$-pseudoconvex. A natural question is, whether the class of $+\infty$-pseudoconvex functions coincides with the union of the classes of n-pseudoconvex functions with n positive integer. The next example gives a negative answer to this question.

Example 7. Consider the function $f : (0, +\infty) \to \mathbb{R}$ defined by

$$f(x) = n + 1 + f_{n+1}(x - n - 1)\,, \quad n < x \leq n + 1\,, \quad n = 0,\, 1,\, \ldots\,,$$

where f_n are the functions from Example 1. Then the function f is $+\infty$-pseudoconvex but it is not n-pseudoconvex for arbitrary positive integer n. The latter follows from Theorem 1, since the point $x_n = n+1$ is n-stationary but not a global minimizer of f.

A central place in our investigation play the relations between n-pseudoconvex functions and n-stationary points obtained in Theorems 1 and 5. To discuss the possible extension of these relations to $+\infty$-pseudoconvex functions we introduce the notion of $+\infty$-stationary point.

Definition 4. *We call $x \in X$ a stationary point of infinite order (for short, $+\infty$-stationary point) of the function $f : X \to \mathbb{R}$, if for each direction $u \in \mathbf{E}$ and arbitrary positive integer m the equalities*

$$f_-^{(i)}(x, u) = 0 \text{ for all positive integers } i < m \text{ imply } f_-^{(m)}(x, u) \geq 0.$$

Theorem 1 remains true if n is replaced by $+\infty$, which follows by repeating nearly the same proof.

Theorem 8. *Let $f : X \to \mathbb{R}$ be a $+\infty$-pseudoconvex function. Then $x^0 \in X$ is a global minimizer of f if and only if x^0 is a $+\infty$-stationary point of f.*

In Example 3 function f possesses a $+\infty$-stationary point $x^0 = 0$. Since this point is not a global minimizer, in view of Theorem 8 this function is not $+\infty$-pseudoconvex.

Theorem 5 also remains true if n is replaced by $+\infty$, which follows by repeating nearly the same proof.

Theorem 9. *Let $f : X \to \mathbb{R}$ be a radially continuous function on the radially open convex set X in a real linear space. Then f is $+\infty$-pseudoconvex if and only if f is quasiconvex and each $+\infty$-stationary point of f is a global minimizer.*

In Example 7 the function f is continuous and quasiconvex, and while it possesses n-stationary points for arbitrary positive integer n, it does not posses $+\infty$-stationary points. Therefore it satisfies the hypotheses of Theorem 8 and on this base we can conclude that f is $+\infty$-pseudoconvex. In fact, the lack of $+\infty$-stationary points implies that f is $+\infty$-pseudoconvex. The next example is of a function which possesses $+\infty$-stationary points.

Example 8. The function $f : \mathbb{R}^n \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} \exp(-\frac{1}{\|x\|}),\ x \neq 0, \\ \hfill 0,\ x = 0, \end{cases}$$

is $+\infty$-pseudoconvex, since it is continuous and quasiconvex and the unique $+\infty$-stationary point $x^0 = 0$ is a global minimizer.

7 The Related Variational Inequality

To a function $f : X \to \mathbb{R}$ we can relate the variational inequality of differential type in Dini derivatives

$$f_-^{(1)}(x^1,\, x^0 - x^1) \le 0\,, \quad x^1 \in X\,. \tag{6}$$

When the set X is convex and f is radially lsc it is shown in Crespi, Ginchev, Rocca [3] that x^0 is a solution of (6) if and only if f increases along rays starting at x^0. Consequently, if x^0 is a solution of (6) then x^0 is a global minimizer of f. In general the set of the global minimizers of f is larger than the set of the solutions of (6). However, when f is quasiconvex, these two sets coincide. Since according to Theorem 4 under the assumptions made, a n-pseudoconvex function is quasiconvex, we get immediately the following result.

Theorem 10. *Let $f : X \to \mathbb{R}$ be radially lsc and n-pseudoconvex function with n positive integer on the convex set X in a real linear space. Then $x^0 \in X$ is a global minimizer of f if and only if x^0 is a solution of the variational inequality (6).*

Theorem 10 is true with both n positive integer and $n = +\infty$. Denote by $S(f, X)$ the set of the solutions of (6). Let us underline, that if f is n-pseudoconvex, then any point $x^0 \in S(f, X)$ satisfies the following property:

$$\begin{gathered} \forall\, x^1 \in X \setminus S(f, X) : \exists\, m \le n : \\ f_-^{(i)}(x^1, x^0 - x^1) = 0 \text{ for } i < m \text{ and } f_-^{(m)}(x^1, x^0 - x^1) < 0\,. \end{gathered} \tag{7}$$

Now problem (6)–(7) is the problem to find the solutions of (6), which satisfy also (7). This problem can be considered as some refinement of the variational inequality. When $X \subset \mathbb{R}$ is convex and f is lsc, then problem (6)–(7) has a solution if and only if f is n-pseudoconvex with nonempty set of its global minimizers. When X is a convex subset of a real linear space and f is radially lsc, then problem (6)–(7) has a solution if and only if f has a nonempty set of its global minimizers and f is radially n-pseudoconvex along rays starting at x^0. Obviously, one can look for other relations between problem (6)–(7) in the case when it possesses solutions and the notion of n-pseudoconvexity.

8 Final Remarks

The pseudoconvex functions were introduced in Tuy [16] and Mangasarian [13] in an attempt to find out a larger class of functions preserving the good properties of the convex functions. For instance each local minimizer of a convex function is a global minimizer. The same property obey the pseudoconvex functions and n-pseudoconvex functions introduced here. Also programming problems with pseudoconvex data preserve some of the good properties of the convex programming problems, see e. g. [14]. Among these properties is the relation between local solutions and Kuhn-Tucker points. The Kuhn-Tucker points appear in programming in connection with necessary optimality conditions. In convex programming these conditions turn to be also sufficient, that is each Kuhn-Tucker point is a minimizer, moreover, it is a global minimizer. Similar properties obey programming problems with pseudoconvex objective functions and quasiconvex inequality constraints. The possibility of further extensions of these results motivates us to introduce the notion of n-pseudoconvex function. Naturally, then we need to introduce higher order Kuhn-Tucker points instead of the usual ones. In a future work we intend to discuss programming problems with n-pseudoconvex functions.

Acknowledgement: The authors thank the anonymous referees for the appreciation and the valuable remarks.

References

1. Aussel D (1998) Subdifferential properties of quasiconvex and pseudoconvex functions: Unified approach. J. Optim. Theory Appl. 97:29–45
2. Avriel M (1972) r-convex functions. Math. Program. 2:309–323
3. Crespi GP, Ginchev I, Rocca M (2004) Minty variational inequalities, increase along rays property and optimization. J. Optim. Theory Appl. 123:479–496
4. Crouzeix JP (1998) Characterizations of generalized convexity and generalized monotonicity, a survey. In: Crouzeix JP et al (eds) Generalized convexity, generalized monotonicity: recent results. Kluwer Academic Publishers, Dordrecht

5. Crouzeix JP, Ferland JA (1982) Criteria for quasi-convexity and pseudo-convexity: relations and comparisons. Math. Program. 23:193–205
6. Diewert WE (1981) Alternative characterizations of six kind of quasiconvexity in the nondifferentiable case with applications to nonsmooth programming. In: Schaible S, Ziemba WT (eds) Generalized Concavity in Optimization and Economics. Academic Press, New York
7. Ginchev I (2002) Higher order optimality conditions in nonsmooth optimization. Optimization 51:42–72
8. Giorgi G (1990) A note on the relationships between convexity and invexity. J. Austral. Math. Soc. Ser. B 32:97–99
9. Giorgi G, Komlósi S (1992) Dini derivatives in optimization - part 2. Riv. Mat. Sci. Econ. Soc. 15:3–24
10. Ivanov VI (2001) First-order characterizations of pseudoconvex functions. Serdica Math. J. 27:203–218
11. Karamardian S (1967) Strictly quasiconvex (concave) functions and duality in mathematical programming. J. Math. Anal. Appl. 20:344–358
12. Komlósi S (1983) Some properties of nondifferentiable pseudoconvex functions. Math. Program. 26:232–237
13. Mangasarian OL (1965) Pseudo-convex functions. SIAM J. Control 3:281–290
14. Mangasarian OL (1994) Nonlinear programming. SIAM, Philadelphia PA
15. Tanaka Y (1990) Note on generalized convex functions. J. Optim. Theory Appl. 66:345–349
16. Tuy H (1964) Sur les inégalités linéaires. Colloq. Math. 13:107–123

Sufficient Optimality Conditions and Duality in Nonsmooth Multiobjective Optimization Problems under Generalized Convexity*

Giorgio Giorgi[1], Bienvenido Jiménez[2], and Vicente Novo[3]

[1] Dipartimento di Ricerche Aziendali. Università degli Studi di Pavia. Via S. Felice, 5, 27100 Pavia, Italy (ggiorgi@eco.unipv.it)
[2] Departamento de Matemática Aplicada, E.T.S.I. Industriales, Universidad Nacional de Educación a Distancia, c/ Juan del Rosal, 12, Ciudad Universitaria, 28040 Madrid, Spain (bjimen1@encina.pntic.mec.es)
[3] Departamento de Matemática Aplicada, E.T.S.I. Industriales, Universidad Nacional de Educación a Distancia, c/ Juan del Rosal, 12, Ciudad Universitaria, 28040 Madrid, Spain (vnovo@ind.uned.es)

Summary. We consider a multiobjective optimization problem in $\mathbb{R}^n$ with a feasible set defined by inequality and equality constraints and a set constraint. All the involved functions are, at least, directionally differentiable. We provide sufficient optimality conditions for global and local Pareto minimum under several kinds of generalized convexity. Also Wolfe-type and Mond-Weir-type dual problems are considered, and weak and strong duality theorems are proved.

Key words: Multiobjective optimization problems, sufficient conditions for a Pareto minimum, Lagrange multipliers, tangent cone, quasiconvexity, Dini differentiable functions, Hadamard differentiable functions, duality theorems.

1 Introduction

In this paper, the next multiobjective optimization problem is considered:

$$\text{(MP)} \quad \text{Min } f(x) \quad \text{subject to } g(x) \leq 0, \; h(x) = 0, \; x \in Q,$$

where $f,\ g,\ h$ are functions from $\mathbb{R}^n$ to $\mathbb{R}^p$, $\mathbb{R}^m$ and $\mathbb{R}^r$, respectively, and Q is a subset of $\mathbb{R}^n$.

*This research for the second and third authors was partially supported by Ministerio de Ciencia y Tecnología (Spain), project BMF2003-02194. The authors are grateful to the referee for his comments.

There are many papers dealing with optimality criteria of Fritz John type or Kuhn-Tucker type (K-T). Singh [16] obtained a necessary K-T type optimality condition for the problem (MP), with $Q = \mathbb{R}^n$ and $f,\ g$ and h differentiable, under the Abadie constraint qualification (CQ), which establishes that the contingent cone (from now on tangent) to the feasible set equals the linearized cone:

$$T(S, x_0) = C(S), \tag{1}$$

where $C(S) = \{v \in \mathbb{R}^n :\ \nabla g_j(x_0)v = 0\ \forall j \in J_0,\ \nabla h_k(x_0)v = 0\ k = 1, \ldots, r\}$, $J_0 = \{j :\ g_j(x_0) = 0\}$.

Di and Poliquin [2] supposed that g and h are differentiable at x_0 and continuous on a neighborhood and obtained equality (1) under the linear independence CQ, that is, the linear independence of the gradients for the active constraints and equality constraints.

None of these works incorporated a set constraint. Giorgi and Guerraggio [4] considered different kinds of scalar problems with C^1 functions and incorporated a set constraint, but in the most general case they considered (with the three constraint types) they did not give K-T type conditions.

Di [1] with the same kind of functions as that of [2] proved condition (1) using a Mangasarian-Fromovitz CQ (Theorem 3.3). In the same work, this author also considered a closed convex set constraint Q and obtained the following expression for the tangent cone to the feasible set of (MP) (Theorem 4.1):

$$T(S \cap Q, x_0) = C(S) \cap T(Q, x_0), \tag{2}$$

using a CQ, that is a generalization of the Mangasarian-Fromovitz one.

In [9] (see also [6]), Jiménez and Novo generalized the result of Singh obtaining a very general necessary K-T type condition, because they incorporated a set constraint and considered directionally differentiable functions (in the sense of Hadamard for the objective functions) with convex derivative (linear derivative for the equality constraints). Such a necessary condition is satisfied under the so called extended Abadie CQ (with "$\supset$" in (2) instead of "$=$" and changing the gradients by the Dini derivatives in the definition of $C(S)$). In [5], the authors obtained different necessary optimality conditions for the problem (MP) considering a (convex or arbitrary) set constraint Q. These necessary optimality conditions generalize the ones obtained by the previously aforementioned authors.

In this work, we are going along this way and, after introducing in Section 2 the notations and some previous results, in Section 3 several sufficient optimality conditions are provided. Most of them require some kind of generalized convexity (quasiconvexity or Dini-pseudoconvexity) for the functions or linear combinations of its components. On the other hand, the requirements on the derivatives are, generally, less restrictive than those which are needed in the necessary conditions. These sufficient conditions generalize some results considered by Singh [16], Islam [8], Di [1] and Majumdar [12]. In Section 4, we suppose that Q is an open set and study Wolfe-type and Mond-Weir-type dual problems. For these problems we obtain both weak and strong duality theorems.

2 Notations and Preliminaries

Let x and y be in $\mathbb{R}^n$ and $A \subset \mathbb{R}^n$. Let us denote $x \leq y$ if $x_i \leq y_i,\ i = 1, \ldots, n$ and $x < y$ if $x_i < y_i,\ i = 1, \ldots, n$. We use $\operatorname{cl} A$, $\operatorname{co} A$ and $\operatorname{cone} A$ to denote the closure, convex hull and generated cone by A, respectively. We denote by v^T the transpose vector of a column vector v.

Given a point $x_0 \in A$ and a function $f : \mathbb{R}^n \to \mathbb{R}^p$, we consider the multiobjective optimization problem

$$\operatorname{Min}\{f(x) :\ x \in A\}.$$

A point $x_0 \in A$ is said to be a local Pareto minimum, denoted by $x_0 \in \operatorname{LMin}(f, A)$, if there exists a neighborhood U of x_0 such that

$$A_f \cap A \cap U = \emptyset, \tag{3}$$

where $A_f = \{x \in \mathbb{R}^n :\ \ f(x) \leq f(x_0),\ f(x) \neq f(x_0)\}$. The definition of (global) Pareto minimum is obvious (take $U = \mathbb{R}^n$). A point $x_0 \in A$ is said to be a weak Pareto minimum if there exists no $x \in A$ such that $f(x) < f(x_0)$. We denote by $\operatorname{Min}(f, A)$ and $\operatorname{WMin}(f, A)$ the sets of Pareto minimum points and weak Pareto minimum points, respectively.

Given the considerable difficulty to verify condition (3), different approximations to A and A_f at x_0 are usual. These (first order) approximations are the so-called "tangent cones". In Definition 1 we recall the notion of contingent cone. If the sets, we want to approximate (locally) by means of some cone, are defined through functional constraints, the related approximate cone will be also defined via some directional derivative of the functions (see Definition 2 and the definition of the cone $C(S)$).

Definition 1. *Let $A \subset \mathbb{R}^n$, $x_0 \in \operatorname{cl} A$. The tangent cone to A at x_0 is defined as follows:*

$$T(A, x_0) = \{v \in \mathbb{R}^n :\ \exists t_k > 0,\ \exists x_k \in A,\ x_k \to x_0 \text{ such that } t_k(x_k - x_0) \to v\}.$$

If $D \subset \mathbb{R}^n$, the polar cone to D is $D^* = \{v \in \mathbb{R}^n :\ v^T d \leq 0\ \forall d \in D\}$.

The normal cone to A at x_0 is the polar to the tangent cone, that is, $N(A, x_0) = T(A, x_0)^*$.

Definition 2. *Let $f : \mathbb{R}^n \to \mathbb{R}^p,\ x_0, v \in \mathbb{R}^n$.*
(a) The Dini derivative (or directional derivative) of f at x_0 in the direction v is

$$Df(x_0, v) = \lim_{t \to 0^+} \frac{f(x_0 + tv) - f(x_0)}{t}.$$

(b) The Hadamard derivative of f at x_0 in the direction v is

$$df(x_0, v) = \lim_{(t,u)\to(0^+,v)} \frac{f(x_0 + tu) - f(x_0)}{t}.$$

(c) *f is Dini differentiable or directionally differentiable (resp. Hadamard differentiable) at x_0 if its Dini derivative (resp. Hadamard derivative) exists finite in all directions.*

Definition 3. *The Dini subdifferential of $f : \mathbb{R}^n \to \mathbb{R}$ at x_0 is*

$$\partial_D f(x_0) = \{\xi \in \mathbb{R}^n : \ \xi^T v \le Df(x_0, v) \ \forall v \in \mathbb{R}^n\}.$$

If $Df(x_0, \cdot)$ is a convex function, then there exists the subdifferential (∂) of the Convex Analysis of that function at $v = 0$. This set is nonempty, compact and convex in $\mathbb{R}^n$ and it satisfies

(i) $\partial_D f(x_0) = \partial Df(x_0, \cdot)(0)$,

(ii) $Df(x_0, v) = \mathrm{Max}\{\xi^T v : \ \xi \in \partial_D f(x_0)\}$.

If $Df(x_0, v)$ is not convex, then $\partial_D f(x_0)$ can be empty.

In this work the following generalized convexity notions will be used.

Definition 4. *Let $\Gamma \subset \mathbb{R}^n$ be a convex set, $f : \Gamma \to \mathbb{R}$, and $x_0 \in \Gamma$.*

(a) *f is quasiconvex at x_0 if*

$\forall x \in \Gamma,\ f(x) \le f(x_0) \ \Rightarrow \ f(\lambda x + (1-\lambda)x_0) \le f(x_0) \ \forall \lambda \in (0, 1).$

(b) *f is Dini-pseudoconvex at x_0 if*

$\forall x \in \Gamma,\ f(x) < f(x_0) \ \Rightarrow \ Df(x_0, x - x_0) < 0.$

(c) *f is strictly Dini-pseudoconvex at x_0*

if $\forall x \in \Gamma \setminus \{x_0\},\ f(x) \le f(x_0) \Rightarrow Df(x_0, x - x_0) < 0.$

(d) *f is Dini-quasiconvex at x_0 if*

$\forall x \in \Gamma,\ f(x) \le f(x_0) \ \Rightarrow \ Df(x_0, x - x_0) \le 0.$

(e) *f is quasilinear, Dini-pseudolinear or Dini-quasilinear at x_0, if f and $-f$ are quasiconvex, Dini-pseudoconvex or Dini-quasiconvex at x_0, respectively.*

(f) *f is quasiconvex on Γ if f is quasiconvex at each point of Γ. Analogously for the other concepts.*

(g) *$f = (f_1, f_2, \ldots, f_p) : \Gamma \to \mathbb{R}^p$ is quasiconvex at x_0 if f_i is quasiconvex at x_0 for each $i = 1, 2, \ldots, p$. Similarly for the other concepts.*

The most relevant properties for our purposes, related to these notions, are collected in the following proposition.

Proposition 1.

(a) *If f is Dini differentiable at x_0 and quasiconvex at x_0, then f is Dini-quasiconvex at x_0.*

(b) *([7, Theorem 3.5]) If f is Dini-pseudoconvex at x_0 and continuous on Γ, then f is quasiconvex at x_0.*

(c) *([7, Theorem 3.2]) If f is Dini-quasiconvex at each point of Γ and continuous on Γ, then f is quasiconvex on Γ.*

In [9] the following generalized Motzkin Theorem is shown, that will be used later.

Theorem 1. *([9, Theorem 3.12]). Let $Q \subset \mathbb{R}^n$ be a convex set with $0 \in Q$, $h : \mathbb{R}^n \to \mathbb{R}^r$ be a linear function of components $h_k,\ k \in K = \{1, \dots, r\}$ given by $h_k(u) = c_k^T u$ with $c_k \in \mathbb{R}^n$, $f_1, \dots, f_p$ and $g_1, \dots, g_m$ be sublinear functions from $\mathbb{R}^n$ to $\mathbb{R}$ and $f = (f_1, \dots, f_p)$, $g = (g_1, \dots, g_m)$. Consider the following propositions:*
(a) $0 \in \sum_{i=1}^p \lambda_i \partial f_i(0) + \sum_{j=1}^m \mu_j \partial g_j(0) + \sum_{k=1}^r \nu_k c_k + N(Q, 0)$, $(\lambda, \mu) \geq 0 \Rightarrow \lambda = 0$.
(b) $\sum_{i=1}^p \lambda_i f_i(u) + \sum_{j=1}^m \mu_j g_j(u) + \sum_{k=1}^r \nu_k h_k(u) \geq 0\ \forall u \in Q$, $(\lambda, \mu) \geq 0 \Rightarrow \lambda = 0$.
(c) There exists $v \in \mathbb{R}^n$ such that $f(v) < 0,\ g(v) \leq 0,\ h(v) = 0,\ v \in Q$.
Then:
(i) (a) and (b) are equivalent.
(ii) $(c) \Rightarrow (a)$.

Consider the problem (MP) of Section 1, with Q an arbitrary nonempty set. We will use the following notations. Let

$$S = \{x \in \mathbb{R}^n :\ g(x) \leq 0,\ h(x) = 0\} \tag{4}$$

the set defined by the constraint functions.

The feasible set of (MP) is $S \cap Q$. Let $f_i,\ i \in I = \{1, \dots, p\}$, $g_j,\ j \in J = \{1, \dots, m\}$, $h_k,\ k \in K = \{1, \dots, r\}$ be the component functions of $f,\ g$ and h, respectively. Given $x_0 \in S$, the active index set at x_0 is $J_0 = \{j \in J : g_j(x_0) = 0\}$.

Assuming that the functions are Dini differentiable at x_0, another cone that will be used to approximate S at x_0 is the linearizing cone, defined as follows:

$$C(S) = \{v \in \mathbb{R}^n :\ Dg_j(x_0, v) \leq 0\ \forall j \in J_0,\ Dh_k(x_0, v) = 0\ \forall k \in K\}.$$

3 Sufficient Optimality Conditions

If the functions involved in the problem have certain kinds of generalized convexity, some of the necessary conditions studied in [6] are also sufficient. Furthermore, we state other sufficient conditions that extend previous results from differentiable problems to directionally differentiable problems. Also a sufficient condition of local minimum without convexity requirements (Theorem 6) is given.

Theorem 2. *Let $Q \subset \mathbb{R}^n$ be a convex set and $x_0 \in S \cap Q$, where S is given by (4). Let us suppose the following conditions are verified:*
(a) $f,\ g_j,\ j \in J_0,\ h$ are Dini differentiable at x_0 with convex derivatives.
(b) f is Dini-pseudoconvex at x_0 and $g_j,\ j \in J_0$ and h are Dini-quasiconvex at x_0.
(c) There exist multipliers $(\lambda, \mu, \nu) \in \mathbb{R}^p \times \mathbb{R}^{|J_0|} \times \mathbb{R}^r$ such that $(c1)$ $(\lambda, \mu, \nu) \geq 0$, $(c2)$ $\lambda \neq 0$ and

(*c*3) $0 \in \sum_{i=1}^{p} \lambda_i \partial_D f_i(x_0) + \sum_{j \in J_0} \mu_j \partial_D g_j(x_0) + \sum_{k=1}^{r} \nu_k \partial_D h_k(x_0) + N(Q, x_0)$.
Then $x_0 \in \mathrm{WMin}(f, S \cap Q)$.

Proof. Let us suppose that x_0 is not a weak minimum of f on $S \cap Q$. Then there exists $x \in S \cap Q$ such that $f(x) < f(x_0)$. Because of the Dini-pseudoconvexity of f, $Df_i(x_0, x - x_0) < 0$ $\forall i \in I$; by the Dini-quasiconvexity $Dg_j(x_0, x - x_0) \leq 0$ $\forall j \in J_0$ since $g_j(x) \leq g_j(x_0) = 0$ and also by Dini-quasiconvexity $Dh_k(x_0, x - x_0) \leq 0$ $\forall k \in K$ since $h_k(x) = 0 \leq h_k(x_0) = 0$. Consequently, the system

$$\begin{cases} Df_i(x_0, v) < 0 \ \forall i \in I \\ Dg_j(x_0, v) \leq 0 \ \forall j \in J_0 \\ Dh_k(x_0, v) \leq 0 \ \forall k \in K \\ v \in Q - x_0 \end{cases}$$

is compatible (it has at least the solution $v = x - x_0$). By the generalized Motzkin theorem (Theorem 1(*ii*)), there exist no multipliers (λ, μ, ν) satisfying conditions (*c*1), (*c*2) and (*c*3), obtaining a contradiction. □

Remark 1. This theorem is stated in a more general setting than Theorems 3.1 and 3.3 of Majumdar [12] and it corrects Theorems 3.1 and 3.3 of Majumdar [12], because Majumdar supposes that all the functions are differentiable at x_0, does not consider the set constraint Q and concludes that x_0 is a Pareto minimum (instead of weak Pareto minimum). But this is a mistake, as it is shown in the following example. For other remarks on the results of Majumdar see [10].

Example 1. Let us consider the functions $f : \mathbb{R}^2 \to \mathbb{R}^2$ given by $f(x, y) = (x, y)$ and $g : \mathbb{R}^2 \to \mathbb{R}$ given by $g(x, y) = -x$. The point $x_0 = (0, 1)$ is a weak Pareto minimum but it is not a Pareto minimum as it can easily be checked.

Remark 2. In Theorem 2, Dini-pseudoconvexity of each f_i can be substituted by Dini-pseudoconvexity of $f_\lambda = \sum_{i=1}^{p} \lambda_i f_i$. If in this theorem we want to obtain a Pareto minimum, either we have to consider $\lambda > 0$ instead of $\lambda \neq 0$ requiring f to be Dini-quasiconvex, or to postulate f to be strictly Dini-pseudoconvex, in this case it is enough to have a nonzero multiplier.

Theorem 3. *Under the hypotheses of Theorem 2 it follows that* $x_0 \in \mathrm{Min}(f, S \cap Q)$, *with each one of the following (independent) modifications:*
(*i*) *Supposing, in addition, that* f *is Dini-quasiconvex at* x_0 *and changing in* (*c*2) $\lambda \neq 0$ *by* $\lambda > 0$.
(*ii*) f *is strictly Dini-pseudoconvex at* x_0 *(instead of Dini-pseudoconvex at* x_0*).*
(*iii*) *Changing* (*b*) *by*
(*b*′) f, g_{J_0}, h *are strictly Dini-pseudoconvex at* x_0,
and requiring that the multipliers fulfil $(\lambda, \mu, \nu) \neq 0$ *instead of* $\lambda \neq 0$.

Proof. (i) It is similar to the previous one making a refinement for the index of I that correspond to components of $f(x)$ strictly less than $f(x_0)$. In fact, let us suppose that $x_0 \notin \text{Min}(f, S \cap Q)$. Then there exists $x \in S \cap Q$ such that $f(x) \leq f(x_0)$ and $f(x) \neq f(x_0)$. Let $I_1 = \{i \in I : \ f_i(x) < f_i(x_0)\}$ and $I_0 = \{i \in I : \ f_i(x) = f_i(x_0)\}$. Obviously $I = I_1 \cup I_0$ with $I_1 \neq \emptyset$.
By Dini-pseudoconvexity, (Definition 4(b)), $Df_i(x_0, x - x_0) < 0 \ \forall i \in I_1$ and by Dini-quasiconvexity, $Df_i(x_0, x - x_0) \leq 0 \ \forall i \in I_0$.
As in the previous proof $Dg_j(x_0, x - x_0) \leq 0 \ \forall j \in J_0$ and $Dh_k(x_0, x - x_0) \leq 0 \ \forall k \in K$. Therefore, the system

$$\begin{cases} Df_i(x_0, v) < 0 \ \forall i \in I_1 \\ Df_i(x_0, v) \leq 0 \ \forall i \in I_0 \\ Dg_j(x_0, v) \leq 0 \ \forall j \in J_0 \\ Dh_k(x_0, v) \leq 0 \ \forall k \in K \\ v \in Q - x_0 \end{cases}$$

is compatible (it has at least the solution $v = x - x_0$).
By the generalized Motzkin theorem, there exist no multipliers $(\lambda_{I_1}, \lambda_{I_0}, \mu, \nu) \in \mathbb{R}^{|I_1|} \times \mathbb{R}^{|I_0|} \times \mathbb{R}^{|J_0|} \times \mathbb{R}^r$ such that $(\lambda_{I_1}, \lambda_{I_0}, \mu, \nu) \geq 0, \ \lambda_{I_1} \neq 0$ and $(c3)$ is satisfied. This is a contradiction with hypothesis (c), because $I_1 \subset I$ and the fact that there are not multipliers with $\lambda_{I_1} \neq 0$, contradicting the existence of $\lambda \in \mathbb{R}^p$ with $\lambda_i > 0 \ \forall i \in I = I_1 \cup I_0$.
(ii) It can be proved analogously, taking into account that $f_i(x) \leq f_i(x_0)$ implies $Df_i(x_0, x - x_0) < 0 \ \forall i \in I$.
(iii) It is similar, since under these conditions it follows that

$$Df(x_0, x - x_0) < 0, \ Dg_{J_0}(x_0, x - x_0) < 0, \ Dh(x_0, x - x_0) < 0, \ x - x_0 \in Q - x_0$$

and by the generalized Motzkin theorem we obtain a contradiction. □

Remark 3. Part (i) of this theorem generalizes Theorem 4.1 of Islam [8] who supposes that all the functions are convex, and does not consider equality constraints. In fact, since the functions are convex, they are Dini-pseudoconvex, quasiconvex and Dini differentiable at each point with convex derivative.

This result also extends Theorem 3.2 of Singh [16], who considers convex differentiable functions and does not use a set constraint.

In the same way, it is possible to extend Theorems 3.3 and 3.4 of Singh into this context. Next we give the generalization of the first result of Singh.

Theorem 4. *Let $Q \subset \mathbb{R}^n$, $x_0 \in S \cap Q$, and let us suppose the following:*
(a) f, g_j, $j \in J_0$, h are Dini differentiable at x_0.
(b) There exist multipliers $(\lambda, \mu, \nu) \in \mathbb{R}^p \times \mathbb{R}^{|J_0|} \times \mathbb{R}^r$, $\lambda > 0$, $\mu \geq 0$ such that $DL(x_0, x - x_0) \geq 0$ for each $x \in S \cap Q$, L being the Lagrangian function:

$$L = \sum_{i=1}^{p} \lambda_i f_i + \sum_{j \in J_0} \mu_j g_j + \sum_{k=1}^{r} \nu_k h_k. \tag{5}$$

(c) L is Dini-pseudoconvex at x_0.
Then $x_0 \in \mathrm{Min}(f, S \cap Q)$.

Proof. Let us suppose that $x_0 \notin \mathrm{Min}(f, S \cap Q)$. Then there exists $x \in S \cap Q$ such that $f(x) \leq f(x_0)$ and $f(x) \neq f(x_0)$. Since $x \in S$, we have $g_j(x) \leq g_j(x_0)$ $\forall j \in J_0,\ h(x) = h(x_0)$, and consequently

$$\sum_{i=1}^{p} \lambda_i f_i(x) + \sum_{j \in J_0} \mu_j g_j(x) + \sum_{k=1}^{r} \nu_k h_k(x)$$
$$< \sum_{i=1}^{p} \lambda_i f_i(x_0) + \sum_{j \in J_0} \mu_j g_j(x_0) + \sum_{k=1}^{r} \nu_k h_k(x_0),$$

that means, $L(x) < L(x_0)$. Since L is Dini-pseudoconvex at x_0, it follows that $DL(x_0, x - x_0) < 0$, contradicting hypothesis (b). □

In the following result we extend Theorem 3.4 of Singh [16]. Its proof is similar and so is omitted.

Theorem 5. *Let $Q \subset \mathbb{R}^n$, $x_0 \in S \cap Q$, and let us suppose the following:*
(a) $f,\ g_j,\ j \in J_0,\ h$ are Dini differentiable at x_0.
(b) There exist multipliers $(\lambda, \mu, \nu) \in \mathbb{R}^p \times \mathbb{R}^{|J_0|} \times \mathbb{R}^r$, $\lambda > 0$, $\mu \geq 0$ such that
(b1) $f_\lambda = \sum_{i=1}^{p} \lambda_i f_i$ is Dini-pseudoconvex at x_0.
(b2) $g_\mu = \sum_{j \in J_0} \mu_j g_j$ and $h_\nu = \sum_{k=1}^{r} \nu_k h_k$ are Dini-quasiconvex at x_0.
(c) $DL(x_0, x - x_0) \geq 0$ for each $x \in S \cap Q$, where L is the Lagrangian function given by (5).
Then $x_0 \in \mathrm{Min}(f, S \cap Q)$.

If all the functions involved in Problem (MP) are Hadamard differentiable, the convexity requirements can be suppressed and we get a local minimum. The following result generalizes Theorem 5.2.II(i) of Di [1].

Theorem 6. *Let $Q \subset \mathbb{R}^n$, $x_0 \in S \cap Q$ and let us suppose the following conditions are verified:*
(a) For each $j \in J_0$, g_j is either Hadamard differentiable at x_0 or Dini-quasiconvex at x_0 with Dini derivative $Dg_j(x_0, \cdot)$ continuous.
(b) For each $k \in K$, h_k is either Hadamard differentiable at x_0 or Dini-quasilinear at x_0 with Dini derivative $Dh_k(x_0, \cdot)$ continuous.
(c) There exist multipliers $(\lambda, \mu, \nu) \in \mathbb{R}^p \times \mathbb{R}^{|J_0|} \times \mathbb{R}^r$, $(\lambda, \mu) \geq 0$, $\lambda \neq 0$ such that
(c1) $f_\lambda = \sum_{i=1}^{p} \lambda_i f_i$ is Hadamard differentiable at x_0.
(c2) $df_\lambda(x_0, v) + Dg_\mu(x_0, v) + Dh_\nu(x_0, v) > 0 \quad \forall v \in T(S \cap Q, x_0) \setminus \{0\}$, being $g_\mu = \sum_{j \in J_0} \mu_j g_j$ and $h_\nu = \sum_{k=1}^{r} \nu_k h_k$.
Then $x_0 \in \mathrm{LMin}(f, S \cap Q)$.

Proof. We have $T(S \cap Q, x_0) \subset T(S, x_0) \cap T(Q, x_0) \subset C(S) \cap T(Q, x_0)$, where the last inclusion is true by [6, Proposition 3.2]. Then, for each $v \in T(S \cap Q, x_0)$, we have that $v \in C(S)$, so that, $Dg_j(x_0, v) \leq 0\ \forall j \in J_0$ and $Dh_k(x_0, v) = 0\ \forall k \in K$. Therefore, $Dg_\mu(x_0, v) \leq 0$ and $Dh_\nu(x_0, v) = 0$. From here, by $(c2)$ it follows that $df_\lambda(x_0, v) > 0\ \forall v \in T(S \cap Q, x_0)$. Using [17, Corollary 2.1], x_0 is a strict local minimum of order 1 of f_λ on $S \cap Q$.

Suppose that $x_0 \notin \mathrm{LMin}(f, S \cap Q)$, then there exists a sequence $x_n \in S \cap Q$, $x_n \to x_0$ such that $f(x_n) \leq f(x_0)$, $f(x_n) \neq f(x_0)$. Therefore, $\sum_{i=1}^{p} \lambda_i f_i(x_n) \leq \sum_{i=1}^{p} \lambda_i f_i(x_0)$, that means, $f_\lambda(x_n) \leq f_\lambda(x_0)$, contradicting the fact that x_0 is a strict local minimum of f_λ. □

Remark 4. Obviously, if f is $\mathbb{R}$-valued $(p = 1)$, the conclusion means that x_0 is a strict local minimum of f on $S \cap Q$. Note that this theorem is almost a converse of Theorem 4.1 in [6], since condition (ii) of this theorem, by Theorem 1(i), is equivalent to:

$$df_\lambda(x_0, v) + Dg_\mu(x_0, v) + Dh_\nu(x_0, v) \geq 0 \quad \forall v \in Q - x_0,$$

and thus, the inequality is true $\forall v \in T(Q, x_0)$.

Note the difficulty in Theorem 5 of proving (c) (since S must be determined). It can be given alternatives using $T(S \cap Q, x_0)$, instead of $S \cap Q$, since, if a constraint qualification holds, this cone is easily obtained as $C(S) \cap T(Q, x_0)$. These alternatives require stronger hypotheses. Thus, we can finally give the following intermediate sufficient conditions between Theorems 2, 5 and 6.

Theorem 7. *Let $Q \subset \mathbb{R}^n$ be a convex set, $x_0 \in S \cap Q$, and let us suppose the following:*
(a) *g_j, $j \in J_0$ are quasiconvex at x_0 and Dini differentiable at x_0.*
(b) *h is quasilinear at x_0 and Dini differentiable at x_0.*
(c) *There exist multipliers $(\lambda, \mu, \nu) \in \mathbb{R}^p \times \mathbb{R}^{|J_0|} \times \mathbb{R}^r$, $(\lambda, \mu) \geq 0$, $\lambda > 0$ such that*
$(c1)$ *$f_\lambda = \sum_{i=1}^{p} \lambda_i f_i$ is Dini-pseudoconvex at x_0.*
$(c2)$ *$Df_\lambda(x_0, v) + Dg_\mu(x_0, v) + Dh_\nu(x_0, v) \geq 0 \quad \forall v \in T(S \cap Q, x_0)$.*
Then $x_0 \in \mathrm{Min}(f, S \cap Q)$.

Proof. Let us suppose that there exists $x \in S \cap Q$ such that $f(x) \leq f(x_0)$ and $f(x) \neq f(x_0)$. Then $f_\lambda(x) < f_\lambda(x_0)$. Since $x \in S$, $g_j(x) \leq 0 = g_j(x_0)\ \forall j \in J_0$, and since g_j is Dini-quasiconvex at x_0 because is quasiconvex at x_0, it follows that $Dg_j(x_0, x - x_0) \leq 0$. Analogously for h, we have $Dh_k(x_0, x - x_0) = 0$. Consequently, multiplying each one of these inequalities by the corresponding multiplier and adding, it follows that $Df_\lambda(x_0, x - x_0) + Dg_\mu(x_0, x - x_0) + Dh_\nu(x_0, x - x_0) < 0$. Let us see that $x - x_0 \in T(S \cap Q, x_0)$, obtaining a contradiction, by condition $(c2)$.
By [9, Lemma 2.6], there exists a ball B_0 centered at x_0 such that

$T(S \cap Q, x_0) = \mathrm{cl\,cone}\,(S \cap Q \cap B_0 - x_0)$ and $g_j(x) < 0\ \forall x \in B_0,\ \forall j \in J \setminus J_0$.

Let $x_\alpha = \alpha x + (1-\alpha)x_0$ with $\alpha \in [0,1]$. Since Q is convex, $x_\alpha \in Q$. As g_j, $j \in J_0$, is quasiconvex at x_0, $g_j(x_\alpha) \leq 0$. By continuity, there exists $\alpha_1 \in (0,1]$ such that $x_\alpha \in B_0$ $\forall \alpha \in [0,\alpha_1]$ and by election of B_0, $g_j(x_\alpha) < 0$ $\forall \alpha \in [0,\alpha_1]$, $\forall j \in J \setminus J_0$. Since h is quasilinear, $h(x_\alpha) = 0$ $\forall \alpha \in [0,1]$. Then $x_\alpha \in S \cap Q \cap B_0$ $\forall \alpha \in [0,\alpha_1]$, in particular $x_1 = x_{\alpha_1} = \alpha_1 x + (1-\alpha_1)x_0 \in S \cap Q \cap B_0$, following

$$x - x_0 = \alpha_1^{-1}(x_1 - x_0) \in \operatorname{cone}(S \cap Q \cap B_0 - x_0) \subset T(S \cap Q, x_0).$$

□

Theorem 8. *Under the hypotheses of Theorem 7, if f_λ is Dini-quasiconvex at x_0, then, substituting the weak inequality in $(c2)$ with a strict one for $v \neq 0$, we obtain the same conclusion, i.e., $x_0 \in \operatorname{Min}(f, S \cap Q)$.*

Proof. If conclusion were not be true (see proof of Theorem 7), would exists $x_0 \in S \cap Q$ such that $f_\lambda(x) < f_\lambda(x_0)$, $Dg_\mu(x_0, x-x_0) \leq 0$, $Dh_\nu(x_0, x-x_0) = 0$ and $x - x_0 \in T(S \cap Q, x_0)$. By Dini-quasiconvexity of f_λ, $Df_\lambda(x_0, x-x_0) \leq 0$. Then $Df_\lambda(x_0, x-x_0) + Dg_\mu(x_0, x-x_0) + Dh_\nu(x_0, x-x_0) \leq 0$, contradicting $(c2)$. □

Remark 5. As a final remark to the present section, clearly we can give other combinations of the hypotheses used in the last theorems in order to obtain other sufficient conditions. For example, if there exist $(\lambda, \mu, \nu) \in \mathbb{R}^p \times \mathbb{R}^{|J_0|} \times \mathbb{R}^r$, $(\lambda, \mu) \geq 0$, such that f_λ, g_μ and h_ν are Dini-quasiconvex at x_0 and $DL(x_0, x - x_0) > 0$ $\forall x \in S \cap Q$, $x \neq x_0$, then $x_0 \in \operatorname{Min}(f, S \cap Q)$.

4 Duality

In this section we consider two types of dual problems for (MP): a Wolfe-type dual problem and a Mond-Weir-type dual problem (see [19] and [14] for the original formulations). Duality is a very common topic in multiobjective optimization under generalized convexity (see for example [3, 11, 13, 15, 18]).

For simplicity, in (MP) $Q \subset \mathbb{R}^n$ is considered an open set, so $T(Q,x) = \mathbb{R}^n$ and $N(Q,x) = \{0\}$ for all $x \in Q$. In this section, we assume that f and g are Dini differentiable with convex derivative and h is Dini differentiable with linear derivative.

We begin with a Wolfe-type dual problem for (MP).

(WD) Max $\left(f_1(y) + \mu^T g(y) + \nu^T h(y), \ldots, f_p(y) + \mu^T g(y) + \nu^T h(y)\right)$
subject to:

$$0 \in \sum_{i=1}^{p} \lambda_i \partial_D f_i(y) + \sum_{j=1}^{m} \mu_j \partial_D g_j(y) + \sum_{k=1}^{r} \nu_k \partial_D h_k(y), \tag{6}$$

$$y \in \mathbb{R}^n,\ \lambda \in \mathbb{R}^p,\ \lambda \geq 0,\ \lambda^T e = 1,\ \mu \in \mathbb{R}^m,\ \mu \geq 0,\ \nu \in \mathbb{R}^r$$

(here $e = (1, \ldots, 1)^T \in \mathbb{R}^p$). We now obtain weak and strong duality results for (MP) and (WD).

Theorem 9. *(Weak duality) For all feasible points x for (MP) and all feasible (y, λ, μ, ν) for (WD), if $\lambda^T f + \mu^T g + \nu^T h$ is Dini-pseudoconvex, then*

$$f(x) \not< f(y) + \mu^T g(y)e + \nu^T h(y)e.$$

(Here $x \not< y$ is the negation of $x < y$).

Proof. Suppose that $f(x) < f(y) + \mu^T g(y)e + \nu^T h(y)e$. As x is a feasible point of (MP) we have

$$f(x) + \mu^T g(x)e + \nu^T h(x)e \leq f(x) < f(y) + \mu^T g(y)e + \nu^T h(y)e.$$

Multiplying by $\lambda \geq 0$, ($\lambda \neq 0$ and $\lambda^T e = 1$ by assumption) we deduce that $\lambda^T f(x) + \mu^T g(x) + \nu^T h(x) < \lambda^T f(y) + \mu^T g(y) + \nu^T h(y)$. As $\lambda^T f + \mu^T g + \nu^T h$ is Dini-pseudoconvex we obtain that

$$D(\lambda^T f + \mu^T g + \nu^T h)(y, x - y) < 0. \tag{7}$$

On the other hand, from equation (6) it follows that there exist

$$\begin{gathered} a_i \in \partial_D f_i(y),\ i = 1, \ldots, p,\ b_j \in \partial_D g_j(y),\ j = 1, \ldots, m, \\ c_k \in \partial_D h_k(y),\ k = 1, \ldots, r \end{gathered} \tag{8}$$

such that

$$0 = \sum_{i=1}^{p} \lambda_i a_i + \sum_{j=1}^{m} \mu_j b_j + \sum_{k=1}^{r} \nu_k c_k. \tag{9}$$

By the definition of Dini subdifferential, from (8) we derive

$$\begin{aligned} a_i^T u &\leq Df_i(y, u) \quad \forall u \in \mathbb{R}^n,\ i = 1, \ldots, p, \\ b_j^T u &\leq Dg_j(y, u) \quad \forall u \in \mathbb{R}^n,\ j = 1, \ldots, m, \\ c_k^T u &= Dh_k(y, u) \quad \forall u \in \mathbb{R}^n,\ k = 1, \ldots, r. \end{aligned}$$

Multiplying these inequalities by λ_i, μ_j and ν_k, respectively, adding up, and taking into account (9) we obtain

$$0 = \left(\sum_{i=1}^{p} \lambda_i a_i + \sum_{j=1}^{m} \mu_j b_j + \sum_{k=1}^{r} \nu_k c_k \right)^T u \leq D(\lambda^T f + \mu^T g + \nu^T h)(y, u),$$

which, considering $u = x - y$, contradicts (7). □

Remark 6. We can obtain the same thesis under other assumptions. For example, we can assume that the function $f(\cdot) + \mu^T g(\cdot)e + \nu^T h(\cdot)e$ is Dini-pseudoconvex.

Remark 7. If we replace the Dini-pseudoconvexity hypothesis by strict Dini-pseudoconvexity (or, e.g., assuming that $f(\cdot) + \mu^T g(\cdot)e + \nu^T h(\cdot)e$ is strictly Dini-pseudoconvex), or assuming that $\lambda > 0$, we can assert that the following cannot hold:

$$f_i(x) \leq f_i(y) + \mu^T g(y) + \nu^T h(y), \quad \forall i \in I = \{1, 2, \ldots, p\},\ i \neq s,$$
$$f_s(x) < f_s(y) + \mu^T g(y) + \nu^T h(y), \quad \text{for some } s \in I.$$

In [6], necessary optimality conditions are given that ensure $\lambda > 0$.

Corollary 1. *Asume that (x_0, λ, μ, ν) is feasible for (WD) with $\mu^T g(x_0) = 0$ and x_0 is feasible for (MP). If $\lambda^T f + \mu^T g + \nu^T h$ is Dini-pseudoconvex, or $f(\cdot) + \mu^T g(\cdot)e + \nu^T h(\cdot)e$ is Dini-pseudoconvex, then x_0 is a weak Pareto minimum of (MP) and (x_0, λ, μ, ν) is a weak Pareto solution of (WD).*

Proof. Suppose that x_0 is not a weak Pareto minimum of (MP), then there exists a feasible point x such that

$$f(x) < f(x_0). \tag{10}$$

As $\mu^T g(x_0) = 0$ (by assumption) and $\nu^T h(x_0) = 0$ (since x_0 is feasible), it follows that $f(x) < f(x_0) + \mu^T g(x_0)e + \nu^T h(x_0)e$. This contradicts Theorem 9 because (x_0, λ, μ, ν) is feasible for (WD).

Now suppose that (x_0, λ, μ, ν) is not a weak Pareto solution of (WD). Then there exists a feasible point $(y, \bar{\lambda}, \bar{\mu}, \bar{\nu})$ such that

$$f(x_0) + \mu^T g(x_0)e + \nu^T h(x_0)e < f(y) + \bar{\mu}^T g(y)e + \bar{\nu}^T h(y)e.$$

As $\mu^T g(x_0) = 0$ and $\nu^T h(x_0) = 0$, we deduce that $f(x_0) < f(y) + \bar{\mu}^T g(y)e + \bar{\nu}^T h(y)e$, but this contradicts Theorem 9. □

Theorem 10. *(Strong duality) Let x_0 be a weak Pareto minimum for (MP) at which a constraint qualification holds [9]. Then there exist $\lambda \in \mathbb{R}^p_+$, $\mu \in \mathbb{R}^m_+$ and $\nu \in \mathbb{R}^r$ such that (x_0, λ, μ, ν) is feasible for (WD). If, in addition, $\lambda^T f + \mu^T g + \nu^T h$ is Dini-pseudoconvex, or $f(\cdot) + \mu^T g(\cdot)e + \nu^T h(\cdot)e$ is Dini-pseudoconvex, then (x_0, λ, μ, ν) is a weak Pareto solution of (WD) and the optimal values of (MP) and (WD) are equal.*

Proof. From Theorem 4.1 (or Theorem 4.5) in [9] we see that there exist (λ, μ, ν) such that $\lambda \geq 0$, $\lambda \neq 0$, $\mu \geq 0$, $\mu^T g(x_0) = 0$ and (6) holds. Without loss of generality we can assume that $\lambda^T e = 1$. Thus, (x_0, λ, μ, ν) is feasible for (WD). Now we get the conclusion by Corollary 1 (the optimal values are equal because $f(x_0) = f(x_0) + \mu^T g(x_0)e + \nu^T h(x_0)e$). □

We now prove weak and strong duality results between (MP) and the following Mond-Weir-type dual problem:

(MWD) Max $\big(f_1(y), \ldots, f_p(y)\big)$
subject to:

$$0 \in \sum_{i=1}^{p} \lambda_i \partial_D f_i(y) + \sum_{j=1}^{m} \mu_j \partial_D g_j(y) + \sum_{k=1}^{r} \nu_k \partial_D h_k(y),$$

$$\mu^T g(y) + \nu^T h(y) \geq 0,$$

$$y \in \mathbb{R}^n,\ \lambda \in \mathbb{R}^p,\ \lambda \geq 0,\ \lambda^T e = 1,\ \mu \in \mathbb{R}^m,\ \mu \geq 0,\ \nu \in \mathbb{R}^r.$$

Theorem 11. *(Weak duality) Let x be feasible for (MP) and (y, λ, μ, ν) be feasible for (MWD). If f is Dini-pseudoconvex and $\mu^T g + \nu^T h$ is Dini-quasiconvex, then it cannot be*

$$f(x) < f(y).$$

Proof. Since x is feasible for (MP) and (y, λ, μ, ν) is feasible for (MWD) we have

$$\mu^T g(x) + \nu^T h(x) \leq \mu^T g(y) + \nu^T h(y).$$

As $\mu^T g + \nu^T h$ is Dini-quasiconvex we derive that

$$\sum_{j=1}^{m} \mu_j D g_j(y, x-y) + \sum_{k=1}^{r} \nu_k D h_k(y, x-y) \leq 0. \tag{11}$$

Suppose that $f(x) < f(y)$. As f is Dini-pseudoconvex it follows that $Df(y, y-x) < 0$, and so $\sum_{i=1}^{p} \lambda_i D f_i(y, x-y) < 0$. Adding this to (11) it results that

$$\sum_{i=1}^{p} \lambda_i D f_i(y, x-y) + \sum_{j=1}^{m} \mu_j D g_j(y, x-y) + \sum_{k=1}^{r} \nu_k D h_k(y, x-y) < 0.$$

Now we continue as in the proof of Theorem 9. □

For this theorem we can make the same remarks made for Theorem 9.

The proofs of Corollary 2 and Theorem 12 are similar to that Corollary 1 and Theorem 10, respectively, and are omitted.

Corollary 2. *Under the hypothesis of Theorem 11, assume that x_0 is feasible for (MP) and (x_0, λ, μ, ν) is feasible for (MWD). Then x_0 is a weak Pareto minimum of (MP) and (x_0, λ, μ, ν) is a weak Pareto solution of (MWD).*

Theorem 12. *(Strong duality) Let x_0 be a weak Pareto minimum point of (MP) at which a constraint qualification holds [9]. Then there exist $\lambda \in \mathbb{R}^p_+$, $\mu \in \mathbb{R}^m_+$ and $\nu \in \mathbb{R}^r$ such that (x_0, λ, μ, ν) is feasible for (MWD). If, in addition, f is Dini-pseudoconvex and $\mu^T g + \nu^T h$ is Dini-quasiconvex, then (x_0, λ, μ, ν) is a weak Pareto solution of (MWD) and the optimal values of (MP) and (MWD) are equal.*

About another necessary conditions and constraint qualifications, the reader is referred to [5, 6].

References

1. Di S (1996) Classical optimality conditions under weaker assumptions SIAM J. Optim. 6:178–197
2. Di S, Poliquin R (1994) Contingent cone to a set defined by equality and inequality constraints at a Fréchet differentiable point J. Optim. Theory Appl. 81:469–478
3. Egudo RR (1989) Efficiency and generalized convex duality in multiobjective programs J. Math. Anal. Appl. 138:84–94
4. Giorgi G, Guerraggio A (1994) First order generalized optimality conditions for programming problems with a set constraint. In: Komlósi S, Rapcsák T, Schaible S (eds.) Generalized convexity, Lecture Notes in Econom. and Math. Systems, 405:171–185, Springer-Verlag, New York
5. Giorgi G, Jiménez B, Novo V (2004) Minimum principle-type optimality conditions for Pareto problems Int. J. Pure Appl. Math. 10(1):51–68
6. Giorgi G, Jiménez B, Novo V (2004) On constraint qualification in directionally differentiable multiobjective optimization problems RAIRO Operations Research, 38:255–274
7. Giorgi G, Komlósi S (1992) Dini derivatives in optimization. Part II Riv. Mat. Sci. Econom. Social., Anno 15(2):3–24
8. Islam M (1994) Sufficiency and duality in nondifferentiable multiobjective programming Pure Appl. Math. Sci. 39:31–39
9. Jiménez B, Novo V (2002) Alternative theorems and necessary optimality conditions for directionally differentiable multiobjective programs J. Convex Anal. 9(1):97–116
10. Kim DS, Lee GM, Lee BS, Cho SJ (2001) Counterexample and optimality conditions in differentiable multiobjective programming J. Optim. Theory Appl. 109(1):187–192
11. Kim DS, Schaible S (2004) Optimality and duality for invex nonsmooth multiobjective programming Optimization 53(2):165–176
12. Majumdar AAK (1997) Optimality conditions in differentiable multiobjective programming J. Optim. Theory Appl. 92:419–427
13. Mishra SK, Wang SY, Lai KK (2005) Optimality and duality for multiple-objective optimization under generalized type I univexity J. Math. Anal. Appl. 303(1):315–326
14. Mond B, Weir T (1981) Generalized concavity and duality. In: Schaible S, Ziemba WT (eds.) Generalized concavity in optimization and Economics, 263–270, Academic Press. New York
15. Preda V (1992) On efficiency and duality for multiobjective programs J. Math. Anal. Appl. 166(2):365–377
16. Singh C (1987) Optimality conditions in multiobjective differentiable programming J. Optim. Theory Appl. 53:115–123
17. Studniarski M (1986) Necessary and sufficient conditions for isolated local minima of nonsmooth functions SIAM J. Control Optim. 24:1044–1049
18. Weir T, Mond B (1989) Generalized convexity and duality in multiple programming Bull. Austral. Math. Soc. 39(2):287–299
19. Wolfe P (1961) A duality theorem for non-linear programming Quart. Appl. Math. 19(3):239–244

Optimality Conditions for Tanaka's Approximate Solutions in Vector Optimization*

César Gutiérrez[1], Bienvenido Jiménez[2], and Vicente Novo[3]

[1] Departamento de Matemática Aplicada, E.T.S.I. Informática, Universidad de Valladolid, Edificio de Tecnologías de la Información y las Telecomunicaciones, Campus Miguel Delibes, s/n, 47011 Valladolid, Spain (cesargv@mat.uva.es)
[2] Departamento de Matemática Aplicada, E.T.S.I. Industriales, Universidad Nacional de Educación a Distancia, c/ Juan del Rosal, 12, Ciudad Universitaria, 28040 Madrid, Spain (bjimen1@encina.pntic.mec.es)
[3] Departamento de Matemática Aplicada, E.T.S.I. Industriales, Universidad Nacional de Educación a Distancia, c/ Juan del Rosal, 12, Ciudad Universitaria, 28040 Madrid, Spain (vnovo@ind.uned.es)

Summary. In this work, approximate solutions of vector optimization problems in the sense of Tanaka [18] are characterized via scalarization. Necessary and sufficient conditions are obtained using a new order representing property and a new monotonicity concept, respectively. A family of gauge functions defined by generalized Chebyshev norms and verifying both properties is introduced in order to characterize approximate solutions of vector optimization problems via approximate solutions of several scalarizations.

Key words: Vector optimization, ε-efficient solutions, scalarization, gauge function, generalized Chebyshev norms.

1 Introduction and Preliminaries

Approximate solutions of vector optimization problems, known as ε-efficient solutions, are important from both the practical and theoretical points of view because they exist under very mild hypotheses and a lot of solution methods (for example, iterative and heuristic methods) obtain this kind of solutions.

The first and more widely used ε-efficiency concept was introduced by Kutateladze [11]. This notion is useful to approximate the weak efficiency set (see Definition 1), but not for the efficiency set (see, for example, [9, Section 3.1]). In [19] Vályi defined an ε-efficiency concept based on a previously fixed scalar function that allows us to approximate the efficiency set (see, for

*This research was partially supported by the Ministerio de Ciencia y Tecnología (Spain), project BFM2003-02194.

example, [9, Section 3.2]). However, there are a lot of problems in which it is difficult to choose such a scalar function. In [18] Tanaka introduced an ε-efficiency concept that approximates the efficiency set and is independent of any scalar function. In Pareto optimization problems, this notion is equivalent to a previous concept defined by White [20] (see [22, Proposition 3.2]). In [20, 9, 22, 23, 5, 8] the reader can find more relations concerning these ε-efficiency concepts.

When a vector optimization problem is solved using an associated scalar optimization problem (i.e., via a scalarization process), it is important to know if improved ε-efficient solutions are obtained when the scalar objective decreases and viceversa (see [20, Section 1] for a complete discussion of this motivation). In [5, Theorems 3.1.3 and 3.1.4] Göpfert et al. prove relations between ε-efficient solutions following a similar definition as that introduced by Kutateladze and approximate solutions obtained in a scalarization process given by a gauge functional.

In this work, Tanaka's concept is analyzed from this point of view. Specifically, two conditions (see Definitions 4 and 6) are introduced that allow us to obtain relations between the sets of approximate solutions of both problems.

The work is structured as follows. In Section 2, the vector optimization problem is fixed. Moreover, several notations and some preliminary results are given. In Section 3, necessary and sufficient conditions for Tanaka's approximate solutions are established via scalarization. A new monotonicity notion is introduced in obtaining sufficient conditions, while necessary conditions are attained by using a generalized order representing property in nonconvex vector optimization problems and by separation theorems in cone-subconvexlike vector optimization problems. In Section 4, from the previous necessary and sufficient conditions and by using gauge functions and generalized Chebyshev norms, a characterization for ε-efficient solutions is obtained, which attains the same precision in the vector problem as in the scalar problem.

2 Approximate Solutions of Vector Optimization Problems

In the sequel, we denote the interior, the closure and the complement of a set A of $\mathbb{R}^p$ by $\text{int}(A)$, $\text{cl}(A)$ and A^c, respectively. We say that A is solid if $\text{int}(A) \neq \emptyset$. We write the nonnegative orthant of $\mathbb{R}^p$ by $\mathbb{R}^p_+$.

In this paper, the following vector optimization problem is considered:

$$\text{Min}\{f(x) : x \in S\}, \tag{1}$$

where $f : \mathbb{R}^n \to \mathbb{R}^p$ and $S \subset \mathbb{R}^n$, $S \neq \emptyset$. As it is usual in solving (1), a partial order is introduced in the final space as follows, which models the preferences of the decision-maker. Consider a convex cone $D \subset \mathbb{R}^p$, which is pointed ($D \cap (-D) = \{0\}$) and solid. Then,

$$y,\, z \in \mathbb{R}^p,\, y \leq z \iff y - z \in -D. \tag{2}$$

If $D = \mathbb{R}^p_+$ then problem (1) is called Pareto problem.

Let us recall the notions of efficient and weak efficient solution of (1).

Definition 1. *A point $x \in S$ is an efficient (resp. weak efficient) solution of* (1) *if $(f(x) - D\backslash\{0\}) \cap f(S) = \emptyset$ (resp. $(f(x) - \operatorname{int}(D)) \cap f(S) = \emptyset$).*

We denote the set of efficient (resp. weak efficient) solutions of (1) by $\mathrm{E}(f, S)$ (resp. $\mathrm{WE}(f, S)$).

The following ε-efficient concept was introduced by Tanaka [18]. We denote the closed unit ball of $\mathbb{R}^p$ by $\mathbb{B}$ and we consider $\varepsilon \geq 0$.

Definition 2. *A point $x \in S$ is an ε-efficient (resp. weak ε-efficient) solution of* (1) *if $(f(x) - D) \cap f(S) \subset f(x) + \varepsilon\mathbb{B}$ (resp. $(f(x) - \operatorname{int}(D)) \cap f(S) \subset f(x) + \varepsilon\mathbb{B}$).*

The set of ε-efficient (resp. weak ε-efficient) solutions of (1) is denoted by $\mathrm{AE}(f, S, \varepsilon)$ (resp. $\mathrm{WAE}(f, S, \varepsilon)$). Let us observe that Definition 2 depends on the norm considered in $\mathbb{R}^p$ by means of the closed unit ball $\mathbb{B}$. Moreover it is clear that if $0 \leq \varepsilon_1 < \varepsilon_2$ then

$$\mathrm{AE}(f, S, \varepsilon_1) \subset \mathrm{AE}(f, S, \varepsilon_2)\,, \tag{3}$$

and for $\varepsilon = 0$ we recover the sets of efficient and weak efficient solutions: $\mathrm{AE}(f, S, 0) = \mathrm{E}(f, S)$ and $\mathrm{WAE}(f, S, 0) = \mathrm{WE}(f, S)$.

Next, Tanaka's ε-efficient solutions are considered as approximations to the efficient and weak efficient sets.

Theorem 1. *Let $f : \mathbb{R}^n \to \mathbb{R}^p$ be a continuous function at $x_0 \in S$ and let $(\varepsilon_n) \subset \mathbb{R}_+$, $(x_n) \subset S$ be such that $\varepsilon_n \downarrow 0$ and $x_n \to x_0$.*

(i) If $x_n \in \mathrm{WAE}(f, S, \varepsilon_n)$ for each n, then $x_0 \in \mathrm{WE}(f, S)$.
(ii) If $(f(x_n))$ is a nonincreasing sequence (i.e., $f(x_m) \in f(x_n) - D$, $\forall\, m > n$), $x_n \in \mathrm{AE}(f, S, \varepsilon_n)$ for each n and D is closed, then $x_0 \in \mathrm{E}(f, S)$.
(iii) Suppose that $f(S)$ is externally stable with respect to the efficient set (i.e. $f(S) \subset f(\mathrm{E}(f, S)) + D$). If $x_n \in \mathrm{AE}(f, S, \varepsilon_n)$ for each n, then $f(x_0) \in \mathrm{cl}(f(\mathrm{E}(f, S)))$.

Proof. (i) Suppose that $x_0 \notin \mathrm{WE}(f, S)$. Then, there exists $x \in S$ such that $f(x_0) - f(x) \in \operatorname{int}(D)$. As f is continuous at x_0 and $x_n \to x_0$, it follows that $f(x_n) - f(x_0) \to 0$ and so, there exists $m \in \mathbb{N}$ such that $f(x_n) - f(x) \in \operatorname{int}(D)$, $\forall\, n \geq m$. Then, using that x_n is a weak ε_n-efficient solution we deduce that $\|f(x_n) - f(x)\| \leq \varepsilon_n$, $\forall\, n \geq m$. Thus, $f(x_n) \to f(x)$, which is a contradiction since $f(x) \neq f(x_0)$.

(ii) Consider $x \in S$ such that

$$f(x) \in f(x_0) - D. \tag{4}$$

As $f(x_n) \to f(x_0)$, $(f(x_n))$ is a nonincreasing sequence and D is closed we deduce that $f(x_0) \in f(x_n) - D$, $\forall\, n$. From (4) we see that

$$f(x) - f(x_n) = (f(x) - f(x_0)) + (f(x_0) - f(x_n)) \in -D - D = -D, \quad \forall\, n$$

and using that x_n is an ε_n-efficient solution we have that $\|f(x_n) - f(x)\| \leq \varepsilon_n$, $\forall\, n$. Therefore, $f(x_n) \to f(x)$ and so $f(x) = f(x_0)$. It follows that $(f(x_0) - D) \cap f(S) = \{f(x_0)\}$ and we conclude that $x_0 \in \mathrm{E}(f,S)$.

(*iii*) As $f(S)$ is externally stable with respect to $\mathrm{E}(f,S)$ we deduce that there exists a sequence $(z_n) \subset \mathrm{E}(f,S)$ such that $f(z_n) - f(x_n) \in -D$, $\forall\, n$, and using that x_n is an ε_n-efficient solution we see that $\|f(x_n) - f(z_n)\| \leq \varepsilon_n$. Therefore,

$$\begin{aligned} \|f(x_0) - f(z_n)\| &\leq \|f(x_0) - f(x_n)\| + \|f(x_n) - f(z_n)\| \\ &\leq \|f(x_0) - f(x_n)\| + \varepsilon_n \end{aligned}$$

and it follows that $f(x_0) \in \mathrm{cl}(f(\mathrm{E}(f,S)))$, since $f(x_n) \to f(x_0)$ and $\varepsilon_n \downarrow 0$.

□

Let us observe that Kutateladze's ε-efficiency concept gives approximate solutions which are not metrically consistent, i.e., it is possible to obtain feasible sequences (x_n) such that $x_n \to x_0$, x_n is an ε_n-efficient solution in the sense of Kutateladze, $\varepsilon_n \to 0$ and the image $f(x_0)$ is far from the set of efficient objectives $f(\mathrm{E}(f,S))$ (see [9, Example 3.2]). Theorem 1 shows under very mild hypotheses that Tanaka's ε-efficiency solutions are metrically consistent.

3 Conditions for ε-Efficient Solutions via Scalarization

Let $\varphi : \mathbb{R}^p \to \mathbb{R}$ be a scalar function. The scalar optimization problem

$$\mathrm{Min}\{(\varphi \circ f)(x) : x \in S\} \tag{5}$$

is called a scalarization for (1) if solutions of problem (5) are also efficient solutions of problem (1). Following this idea, in this section approximate solutions of (1) and (5) are related.

Definition 3. *A point* $x_0 \in S$ *is an approximate solution of* (5) *(with precision* $\varepsilon \geq 0$*) if*

$$(\varphi \circ f)(x_0) - \varepsilon \leq (\varphi \circ f)(x), \quad \forall\, x \in S.$$

The set of approximate solutions of (5) is denoted by $\mathrm{AMin}(\varphi \circ f, S, \varepsilon)$. Then, our goal in the sequel is to relate the sets $\mathrm{AMin}(\varphi \circ f, S, \varepsilon)$, $\mathrm{AE}(f,S,\varepsilon)$ and $\mathrm{WAE}(f,S,\varepsilon)$.

3.1 Necessary Optimality Conditions

In obtaining necessary conditions for weak ε-efficient solutions, the following approximate order representing property is considered.

Definition 4. *A function $\varphi : \mathbb{R}^p \to \mathbb{R}$ satisfies the approximate order representing property (AORP) at $y_0 \in \mathbb{R}^p$ if*

$$\{y \in \mathbb{R}^p : \varphi(y) < 0\} = y_0 - (\operatorname{int}(D) \cap (\varepsilon\mathbb{B})^c).$$

In Section 4, a wide class of functions satisfying this property is considered. Let us note that the usual order representing property (see, for example, [21, Section 5]) is obtained from property (AORP) taking $\varepsilon = 0$. In the literature, the order representing property has been used to prove necessary conditions for weak efficient solutions via scalarization (see, for example, [21, Theorem 10]). Next, these necessary conditions are extended to weak ε-efficient solutions via property (AORP).

Theorem 2. *Let $x_0 \in S$ and let $\varphi : \mathbb{R}^p \to \mathbb{R}$ be a function verifying property (AORP) at $f(x_0)$. If $x_0 \in \mathrm{WAE}(f, S, \varepsilon)$ then $x_0 \in \mathrm{AMin}(\varphi \circ f, S, \varphi(f(x_0)))$.*

Proof. Consider $x_0 \in \mathrm{WAE}(f, S, \varepsilon)$. From Definition 2 it follows that

$$[f(x_0) - (\operatorname{int}(D) \cap (\varepsilon\mathbb{B})^c)] \cap f(S) = \emptyset,$$

and by property (AORP) at $f(x_0)$ we have that $\varphi(f(x)) \geq 0$, $\forall\, x \in S$. In particular, $\varphi(f(x_0)) \geq 0$,

$$\varphi(f(x_0)) - \varphi(f(x_0)) = 0 \leq \varphi(f(x)), \quad \forall\, x \in S$$

and we conclude that $x_0 \in \mathrm{AMin}(\varphi \circ f, S, \varphi(f(x_0)))$. □

Let us observe that when $\varepsilon = 0$ and φ is a continuous function at $f(x_0)$ verifying property (AORP) at $f(x_0)$ it follows that $\varphi(f(x_0)) = 0$ and so, if $x_0 \in \mathrm{E}(f, S)$ then x_0 is an exact solution of problem (5). This necessary condition is well-known (see, for instance, [21, Theorem 10] and [14, Corollary 1.7]).

Under subconvexlikeness hypotheses and by using separation theorems, a necessary condition for weak ε-efficient solutions can be obtained. Similar results on Kutateladze's ε-efficiency concept are [15, Theorem 3], [1, Theorem 2.1] and [2, Theorem 2.1]. The (positive) polar and strict polar cone of D are

$$D^+ = \{l \in \mathbb{R}^p : \langle l, d\rangle \geq 0, \forall\, d \in D\}$$

and

$$D^{+s} = \{l \in \mathbb{R}^p : \langle l, d\rangle > 0, \forall\, d \in D\backslash\{0\}\},$$

respectively. We denote the dual norm of $\| \;\|$ by $\| \;\|_*$, i.e., $\|l\|_* = \sup\limits_{\|y\| \leq 1} \{|\langle l, y\rangle|\}$.

Definition 5. *[3, Definition 2.4] It is said that a function $f : \mathbb{R}^n \longrightarrow \mathbb{R}^p$ is D-subconvexlike on a nonempty set $S \subset \mathbb{R}^n$ if $f(S) + \mathrm{int}(D)$ is a convex set.*

Theorem 3. *If problem (1) is externally stable with respect to the weak efficiency set (i.e., $f(S) \subset f(\mathrm{WE}(f, S)) + (\mathrm{int}(D) \cup \{0\})$) and the objective function f is D-subconvexlike on the feasible set S then*

$$\mathrm{WAE}(f, S, \varepsilon) \subset \bigcup_{l \in D^+, \|l\|_* = 1} \mathrm{AMin}(\langle l, f(\cdot)\rangle, S, \varepsilon)\,, \quad \forall\, \varepsilon \geq 0.$$

Proof. Consider $x_0 \in \mathrm{WAE}(f, S, \varepsilon)$. As (1) is externally stable with respect to the weak efficiency set, we deduce that there exist $x \in \mathrm{WE}(f, S)$ and $d \in \mathrm{int}(D) \cup \{0\}$ such that

$$f(x_0) = f(x) + d. \tag{6}$$

By Definition 2 we deduce that $\|d\| \leq \varepsilon$. From Definition 1 and by using that $\mathrm{int}(D)$ is a convex cone it follows that

$$(f(x) - \mathrm{int}(D)) \cap (f(S) + \mathrm{int}(D)) = \emptyset.$$

As f is a D-subconvexlike function on the feasible set S then $f(S) + \mathrm{int}(D)$ is a convex set and by applying the Separation Theorem (see, for example, [10, Theorem 3.14]) it follows that there exists $l \in D\backslash\{0\}$ such that

$$\langle l, f(x) - d_1\rangle \leq \langle l, f(z) + d_2\rangle, \quad \forall\, z \in S, \forall\, d_1,\, d_2 \in \mathrm{int}(D)\,. \tag{7}$$

We can suppose that $\|l\|_* = 1$ since $l \neq 0$ and by continuity we can extend (7) to vectors d_1, $d_2 \in D$. Taking $z = x$ and $d_1 = 0$ we deduce that $l \in D^+$. Moreover, from (6) and taking $d_1 = d_2 = 0$ it follows that

$$\langle l, f(x_0) - d\rangle \leq \langle l, f(z)\rangle, \quad \forall\, z \in S.$$

As $|\langle l, d\rangle| \leq \|l\|_* \|d\| \leq \varepsilon$ we see that

$$\langle l, f(x_0)\rangle - \varepsilon \leq \langle l, f(x_0) - d\rangle \leq \langle l, f(z)\rangle, \quad \forall\, z \in S,$$

and the proof is completed. □

3.2 Sufficient Optimality Conditions

It is well-known for practitioners and researchers in vector optimization that one can obtain sufficient conditions on efficient solutions of problem (1) via solutions of problem (5) when the scalar function $\varphi : \mathbb{R}^p \to \mathbb{R}$ is monotone with respect to the order considered in the final space (see relation (2)). Following this idea, we obtain sufficient conditions for ε-efficient solutions of (1) through approximate solutions of scalarizations whose scalar functions satisfy a new monotonicity concept introduced in Definition 6(*ii*).

Definition 6.

(*i*) *A function* $\varphi : \mathbb{R}^p \to \mathbb{R}$ *is* D*-monotone on a set* $F \subset \mathbb{R}^p$ *if* $\varphi(y) \leq \varphi(z)$, $\forall\, y, z \in F, y - z \in -D$.

(*ii*) *A function* $\varphi : \mathbb{R}^p \to \mathbb{R}$ *is strictly local* D*-monotone (SLM) at* $y_0 \in \mathbb{R}^p$ *(respect to the norm* $\| \|$*) with constants* $\alpha > 0$ *and* $\rho > 0$ *if it is* D*-monotone on* $\mathbb{R}^p$ *and*

$$\varphi(y_0) \geq \varphi(y) + \alpha\|y - y_0\|, \quad \forall\, y \in y_0 - (D \cap \mathrm{int}(\rho\mathbb{B})). \tag{8}$$

Let us observe that a D-monotone function $\varphi : \mathbb{R}^p \to \mathbb{R}$ verifies property (SLM) at y_0 if and only if y_0 is a strict local solution of first order of the scalar problem

$$\mathrm{Max}\{\varphi(y) : y \in y_0 - D\}.$$

Example 1. Let us suppose that the cone D is closed. For each $y_0 \in \mathbb{R}^p$ and $\rho > 0$ it is easy to prove that any $l \in D^{+s}$ is strictly local D-monotone at y_0 with constants ρ and

$$\alpha = \min_{d \in D, \|d\|=1} \{\langle l, d\rangle\}.$$

By the Weierstrass theorem, it is clear that $\alpha > 0$. In Section 4, another wide class of strictly local D-monotone functions is considered.

Theorem 4. *Let* $x_0 \in S$ *and let* $\varphi : \mathbb{R}^p \to \mathbb{R}$ *be a strictly local* D*-monotone function at* $f(x_0)$ *with constants* α *and* ρ*. If* $0 \leq \delta < \alpha\rho$ *and* $x_0 \in \mathrm{AMin}(\varphi \circ f, S, \delta)$ *then* $x_0 \in \mathrm{AE}(f, S, \delta/\alpha)$.

Proof. Let us suppose that $x_0 \notin \mathrm{AE}(f, S, \delta/\alpha)$. Consequently, there exists $x \in S$ such that

$$f(x) \in f(x_0) - D \tag{9}$$

and $\|f(x) - f(x_0)\| > \delta/\alpha$. As $\delta < \alpha\rho$ we can select $\nu > 0$ verifying

$$\frac{\delta + \nu}{\alpha\|f(x) - f(x_0)\|} < 1, \qquad \delta + \nu < \alpha\rho. \tag{10}$$

Let us consider the point

$$y := f(x) + \left(1 - \frac{\delta + \nu}{\alpha\|f(x) - f(x_0)\|}\right)(f(x_0) - f(x)) \tag{11}$$

$$= f(x_0) + \frac{\delta + \nu}{\alpha\|f(x) - f(x_0)\|}(f(x) - f(x_0)). \tag{12}$$

From (9)-(11) we see that $y \in f(x) + D$, and as $x_0 \in \mathrm{AMin}(\varphi \circ f, S, \delta)$ and φ is D-monotone on $\mathbb{R}^p$ it follows that

$$\varphi(f(x_0)) - \delta \le \varphi(f(x)) \le \varphi(y). \tag{13}$$

By (10) and (12) we deduce that

$$\|y - f(x_0)\| = \frac{\delta + \nu}{\alpha} < \rho, \tag{14}$$

and from (9) and (12) we have that $y \in f(x_0) - D$. Therefore, $y \in f(x_0) - (D \cap \mathrm{int}(\rho\mathbb{B}))$, and so

$$\varphi(f(x_0)) \ge \varphi(y) + \alpha\|y - f(x_0)\|, \tag{15}$$

since φ is strictly local D-monotone at $f(x_0)$ with constants α and ρ. From (14) and (15) it follows that $\varphi(f(x_0)) > \varphi(y) + \delta$, contrary to (13). □

4 Characterization of ε-Efficient Solutions

In this section, a class of functions verifying properties (AORP) and (SLM) is obtained. Next, by means of these functions and the results attained in Section 3 we characterize the ε-efficiency set of (1) through approximate solutions of scalar problems as (5).

The following result due to Gerth and Weidner and Lemma 2 are used to obtain functions with the property (AORP).

Lemma 1. *[4, Theorem 2.1] Let* $C \subset \mathbb{R}^p$ *be a solid set such that* $C \ne \mathbb{R}^p$ *and* $\mathrm{cl}(C) + \mathrm{int}(D) \subset \mathrm{int}(C)$. *For a fixed vector* $q \in \mathrm{int}(D)$, *the function* $\varphi : \mathbb{R}^p \to \mathbb{R}$ *defined by*

$$\varphi(y) = \inf\{s \in \mathbb{R} : y \in sq - \mathrm{cl}(C)\} \tag{16}$$

is continuous, D-monotone on $\mathbb{R}^p$ *and*

$$\{y \in \mathbb{R}^p : \varphi(y) < 0\} = -\mathrm{int}(C)\,. \tag{17}$$

The function φ was introduced by Rubinov [16] and is called "the smallest strictly monotonic function" by Luc [14] due to property (17).

Lemma 2.

(*i*) $\mathrm{cl}(D \cap (\varepsilon\mathbb{B})^c) = \mathrm{cl}(D) \cap (\mathrm{int}(\varepsilon\mathbb{B}))^c$, $\forall\, \varepsilon > 0$.
(*ii*) *If the norm* $\|\ \|$ *is D-monotone on D then*

$$\mathrm{cl}(D \cap (\varepsilon\mathbb{B})^c) + \mathrm{int}(D) \subset \mathrm{int}(D \cap (\varepsilon\mathbb{B})^c)\,, \quad \forall\, \varepsilon \ge 0. \tag{18}$$

Proof. (*i*) For each $\varepsilon > 0$ it is clear that $\mathrm{cl}(D \cap (\varepsilon\mathbb{B})^c) \subset \mathrm{cl}(D) \cap (\mathrm{int}(\varepsilon\mathbb{B}))^c$. In proving the reciprocal inclusion, let us take a point $y \in \mathrm{cl}(D) \cap (\mathrm{int}(\varepsilon\mathbb{B}))^c$. As $y \ne 0$ we deduce that there exists a sequence $(d_n) \subset D\backslash\{0\}$ such that $d_n \to y$. Therefore,

$$\left(\frac{\|y\|}{\|d_n\|} + \frac{1}{n}\right) d_n \in D \cap (\varepsilon\mathbb{B})^c, \qquad \left(\frac{\|y\|}{\|d_n\|} + \frac{1}{n}\right) d_n \to y$$

and so $y \in \mathrm{cl}(D \cap (\varepsilon\mathbb{B})^c)$.

(ii) As D is a solid convex cone it follows that

$$\mathrm{cl}(D) + \mathrm{int}(D) = \mathrm{int}(D)$$

(see, for example, [14, Chapter 1, Proposition 1.4]), and so (18) holds for $\varepsilon = 0$.

Let $\varepsilon > 0$ and $M := D \cap (\varepsilon\mathbb{B})^c$. It is easy to see that

$$M + \mathrm{int}(D) \subset M, \tag{19}$$

because if $y \in M$ and $d \in \mathrm{int}(D)$ then $\|y + d\| \geq \|y\| > \varepsilon$ since the norm $\| \, \|$ is D-monotone on D, and so $y + d \in M$. Also we have that

$$\mathrm{cl}(M) + \mathrm{int}(D) \subset \mathrm{cl}(M) . \tag{20}$$

Indeed, let $y \in \mathrm{cl}(M)$ and $d \in \mathrm{int}(D)$. Then there exists a sequence $(y_n) \subset M$ such that $y_n \to y$, and therefore $y_n + d \to y + d$. But $y_n + d \in M$ by (19), and consequently $y + d \in \mathrm{cl}(M)$.

Now, as the set $\mathrm{cl}(M) + \mathrm{int}(D) = \bigcup_{y \in \mathrm{cl}(M)} (y + \mathrm{int}(D))$ is open, from (20) it follows that

$$\mathrm{cl}(M) + \mathrm{int}(D) \subset \mathrm{int}(\mathrm{cl}(M)) ,$$

and we obtain the conclusion if we see that

$$\mathrm{int}(\mathrm{cl}(M)) = \mathrm{int}(M) .$$

As a matter of fact, $\mathrm{cl}(M) = \mathrm{cl}(D) \cap (\mathrm{int}(\varepsilon\mathbb{B}))^c$ by part (i), and $\mathrm{int}(\mathrm{cl}(D)) = \mathrm{int}(D)$ since D is a solid convex set. Hence

$$\mathrm{int}(\mathrm{cl}(M)) = \mathrm{int}(\mathrm{cl}(D)) \cap (\varepsilon\mathbb{B})^c = \mathrm{int}(D) \cap (\varepsilon\mathbb{B})^c = \mathrm{int}(M) .$$

□

Proposition 1. *Let $y_0 \in \mathbb{R}^p$, $q \in \mathrm{int}(D)$, $\varepsilon \geq 0$, $M = D \cap (\varepsilon\mathbb{B})^c$ and the function*

$$\varphi_{\varepsilon,0}(y) = \inf\{s \in \mathbb{R} : y \in sq - \mathrm{cl}(M)\}. \tag{21}$$

Then the function

$$\varphi_{\varepsilon,y_0}(y) := \varphi_{\varepsilon,0}(y - y_0) \tag{22}$$

is continuous, D-monotone on $\mathbb{R}^p$ and verifies property (AORP) at y_0.

Proof. It is clear that $\varphi_{\varepsilon,0}$ is the function defined in (16) with $C := M$. Moreover, M is a solid set, $M \neq \mathbb{R}^p$ and from Lemma 2(*ii*) we see that $\mathrm{cl}(M)+\mathrm{int}(D) \subset \mathrm{int}(M)$. Then, by Lemma 1 it follows that the function $\varphi_{\varepsilon,0}$ is continuous, D-monotone on $\mathbb{R}^p$ and

$$\{y \in \mathbb{R}^p : \varphi_{\varepsilon,0}(y) < 0\} = -\mathrm{int}(M) = -\mathrm{int}(D) \cap (\varepsilon\mathbb{B})^c. \tag{23}$$

From here it is clear that the function $\varphi_{\varepsilon,y_0}$ is continuous, D-monotone on $\mathbb{R}^p$ and by (23) we conclude that

$$\begin{aligned}\{y \in \mathbb{R}^p : \varphi_{\varepsilon,y_0}(y) < 0\} &= \{y \in \mathbb{R}^p : \varphi_{\varepsilon,0}(y - y_0) < 0\} \\ &= y_0 - (\mathrm{int}(D) \cap (\varepsilon\mathbb{B})^c).\end{aligned}$$

□

Let us observe that function $\varphi_{\varepsilon,0}$ is different from the gauge functional considered in [5, Section 3.1] (for example, $0 \notin \mathrm{cl}(M)$ when $\varepsilon > 0$).

In the sequel, we assume that the cone D is closed and $q \in \mathrm{int}(D)$. In order to obtain functions satisfying property (SLM), the following generalized Chebyshev norm is considered (see [10, Lemma 1.45] for more details on these norms):

$$\|y\|_q = \inf\{s > 0 : y \in (-sq + D) \cap (sq - D)\}.$$

We denote

$$\mathcal{R}_q(y) = \{s > 0 : y \in (-sq + D) \cap (sq - D)\}$$

and we write $\mathbb{B}_q$ to denote the closed unit ball defined by the norm $\| \, \|_q$. The function $\varphi_{\varepsilon,y_0}$ is denoted by $\varphi_{q,\varepsilon,y_0}$ when the closed unit ball considered in (21) via the set M is $\mathbb{B}_q$.

Next, we prove that the function $\varphi_{q,\varepsilon,y_0}$ verifies property (SLM). The following lemma is necessary.

Lemma 3.

(*i*) *If* $d \in D$ *then* $\mathcal{R}_q(d) = \{s > 0 : d \in sq - D\}$.
(*ii*) $\| \, \|_q$ *is a* D*-monotone function on* D.
(*iii*) $\|q\|_q = 1$.
(*iv*) *If* $d \in D$, $t \in \mathbb{R}$ *and* $d + tq \in D$ *then* $\|d + tq\|_q = \|d\|_q + t$.

Proof. (*i*) As D is a convex cone it is clear that

$$d = -sq + (d + sq) \in -sq + D, \quad \forall\, d \in D, \forall\, s > 0$$

and the result follows.

(*ii*) The result is easily deduced from the inclusion

$$\mathcal{R}_q(d_1 + d_2) \subset \mathcal{R}_q(d_1), \quad \forall\, d_1,\, d_2 \in D. \tag{24}$$

To prove (24), let us take any $d_1, d_2 \in D$ and any $s \in \mathcal{R}_q(d_1+d_2)$. Then $d_1+d_2 \in sq-D$ and so $d_1 \in sq-(d_2+D) \subset sq-D$, since D is a convex cone. Hence, by part (i) it follows that $s \in \mathcal{R}_q(d_1)$ and relation (24) holds.

(iii) For each $s \geq 1$ it is clear that

$$q = sq-(s-1)q \in sq-D.$$

Then, by part (i) it follows that $[1,\infty) \subset \mathcal{R}_q(q)$. Conversely, if $s \in (0,1)$ and $q \in sq-D$ then $q \in D \cap (-D)$, which is a contradiction since D is pointed and $q \neq 0$. Therefore $\mathcal{R}_q(q) = [1,\infty)$ and so $\|q\|_q = 1$.

(iv) Let $d \in D$, $t \in \mathbb{R}$ and $d+tq \in D$. If $d=0$ then $tq \in D$ and so $t \geq 0$, since D is pointed. Thus, by part (iii) we see that $\|tq\|_q = t\|q\|_q = t$ and the result follows for $d=0$.

Suppose that $d \neq 0$. In this case we have that

$$\mathcal{R}_q(d+tq) = \mathcal{R}_q(d)+t. \tag{25}$$

Indeed, if $s \in \mathcal{R}_q(d+tq)$ then $s>0$ and $d \in (s-t)q-D$. As D is pointed and $d \neq 0$ it is clear that $s-t>0$ and so $s-t \in \mathcal{R}_q(d)$. Therefore we have that

$$\mathcal{R}_q(d+tq) \subset \mathcal{R}_q(d)+t. \tag{26}$$

From (26) we see also that $\mathcal{R}_q((d+tq)-tq) \subset \mathcal{R}_q(d+tq)-t$. Then (25) follows and so

$$\|d+tq\|_q = \inf\{\mathcal{R}_q(d+tq)\} = \inf\{\mathcal{R}_q(d)+t\} = \inf\{\mathcal{R}_q(d)\}+t = \|d\|_q+t.$$

□

Proposition 2. *Consider* $\varepsilon > 0$*.*

(i) $\varphi_{q,\varepsilon,y_0}(y_0) = \varphi_{q,\varepsilon,y_0}(y_0-d)+\|d\|_q,\ \forall\, d \in D \cap \operatorname{int}(\varepsilon\mathbb{B}_q)$.

(ii) The function $\varphi_{q,\varepsilon,y_0}$ *verifies property (SLM) at* y_0 *with constants* $\alpha = 1$ *and* $\rho = \varepsilon$*.*

Proof. (i) Let $d \in D \cap \operatorname{int}(\varepsilon\mathbb{B}_q)$. From (21) and (22) we have that

$$\varphi_{q,\varepsilon,y_0}(y_0-d) = \varphi_{q,\varepsilon,0}(-d) = \inf\{s \in \mathbb{R} : d+sq \in D \cap (\operatorname{int}(\varepsilon\mathbb{B}_q))^c\}. \tag{27}$$

It follows that

$$\{s \in \mathbb{R} : d+sq \in D \cap (\operatorname{int}(\varepsilon\mathbb{B}_q))^c\} = [\varepsilon-\|d\|_q, \infty). \tag{28}$$

Indeed, let $s \in \mathbb{R}$ such that $d+sq \in D \cap (\operatorname{int}(\varepsilon\mathbb{B}_q))^c$. By Lemma 3$(iv)$ we see that $\|d+sq\|_q = \|d\|_q+s$ and $s \geq \varepsilon-\|d\|_q$, since $d+sq \in (\operatorname{int}(\varepsilon\mathbb{B}_q))^c$. Conversely, if $s \geq \varepsilon-\|d\|_q$ then $s>0$ since $d \in \operatorname{int}(\varepsilon\mathbb{B}_q)$, and so $d+sq \in D$. Moreover, by Lemma 3(iv) we deduce that $\|d+sq\|_q = \|d\|_q+s$ and as $s \geq \varepsilon-\|d\|_q$ we deduce that $d+sq \in (\operatorname{int}(\varepsilon\mathbb{B}_q))^c$.

In particular, taking $d = 0$ in (28) we see that

$$\varphi_{q,\varepsilon,y_0}(y_0) = \varepsilon. \tag{29}$$

By (27) and (28) we conclude that

$$\varphi_{q,\varepsilon,y_0}(y_0 - d) + \|d\|_q = \varepsilon - \|d\|_q + \|d\|_q = \varepsilon = \varphi_{q,\varepsilon,y_0}(y_0).$$

(ii) From Proposition 1 we have that $\varphi_{q,\varepsilon,y_0}$ is a D-monotone function on $\mathbb{R}^p$ and by part (i) we deduce that condition (8) holds for $\alpha = 1$ and $\rho = \varepsilon$. Then, it follows that the function $\varphi_{q,\varepsilon,y_0}$ is strictly local D-monotone at y_0 respect to the norm $\| \|_q$ with constants $\alpha = 1$ and $\rho = \varepsilon$. □

Next, a characterization for Tanaka's ε-efficient solutions is obtained. Let us observe that in both conditions, the precision attained is equal to the precision assumed as hypothesis. Similar results for various Kutateladze type weak ε-efficient solutions have been proved in convex (resp. nonconvex) Pareto problems [1, Theorem 2.1], [2, Theorem 2.1] (resp. [12, Theorem 1]) and in convex (resp. nonconvex) vector optimization problems [15, Theorem 3] (resp. [17, Corollary 1]).

Theorem 5. *Consider $\varepsilon \geq 0$, $q \in \operatorname{int}(D)$ and assume that $\mathbb{R}^p$ is normed by the generalized Chebyshev norm $\| \|_q$.*

(i) *If $x_0 \in \mathrm{AE}(f, S, \varepsilon)$ then $x_0 \in \mathrm{AMin}\big(\varphi_{q,\varepsilon,f(x_0)} \circ f, S, \varepsilon\big)$.*
(ii) *If $\varepsilon > 0$, $x_0 \in \mathrm{AMin}\big(\varphi_{q,\varepsilon,f(x_0)} \circ f, S, \delta\big)$ and $0 \leq \delta < \varepsilon$ then $x_0 \in \mathrm{AE}(f, S, \delta)$.*

Proof. (i) Consider $x_0 \in \mathrm{AE}(f, S, \varepsilon)$. By Proposition 1 we have that the function $\varphi_{q,\varepsilon,f(x_0)}$ verifies the property (AORP) at $f(x_0)$. Hence, from Theorem 2 we deduce that $x_0 \in \mathrm{AMin}\big(\varphi_{q,\varepsilon,f(x_0)} \circ f, S, \varphi_{q,\varepsilon,f(x_0)}(f(x_0))\big)$, since $\mathrm{AE}(f, S, \varepsilon) \subset \mathrm{WAE}(f, S, \varepsilon)$. Then, the result follows if we show that

$$\varphi_{q,\varepsilon,f(x_0)}(f(x_0)) = \varepsilon.$$

Indeed,

$$\varphi_{q,\varepsilon,f(x_0)}(f(x_0)) = \varphi_{q,\varepsilon,0}(0) = \inf\{s \in \mathbb{R} : sq \in \mathrm{cl}(M)\},$$

where $M = D \cap (\varepsilon \mathbb{B}_q)^c$. If $\varepsilon = 0$ then $\mathrm{cl}(M) = D$ and

$$\varphi_{q,0,f(x_0)}(f(x_0)) = \inf\{s \in \mathbb{R} : sq \in D\} = 0.$$

If $\varepsilon > 0$ then by (29) we have that $\varphi_{q,\varepsilon,f(x_0)}(f(x_0)) = \varepsilon$ and the result holds.

(ii) Suppose that $x_0 \in \mathrm{AMin}\big(\varphi_{q,\varepsilon,f(x_0)} \circ f, S, \delta\big)$ and $0 \leq \delta < \varepsilon$. By Proposition 2$(ii)$ we see that the function $\varphi_{q,\varepsilon,f(x_0)}$ satisfies property (SLM) at $f(x_0)$ with constants $\alpha = 1$ and $\rho = \varepsilon$. Then, from Theorem 4 it follows that $x_0 \in \mathrm{AE}(f, S, \delta)$. □

Remark 1. Let us consider $\varepsilon > 0$. From (3) and under the hypotheses of Theorem 5 it follows that

$$\bigcup_{0\le\delta<\varepsilon} \{x \in \mathbb{R}^n : x \in \mathrm{AMin}\big(\varphi_{q,\varepsilon,f(x)} \circ f, S, \delta\big)\} \subset \mathrm{AE}(f, S, \varepsilon)\,.$$

This inclusion is in general strict (see Example 3).

Example 2. Let us obtain the expression of the function $\varphi_{q,\varepsilon,y_0}$ in Pareto problems. Consider $D = \mathbb{R}^p_+$, $q = (q_1, q_2, \ldots, q_p) \in \mathrm{int}\big(\mathbb{R}^p_+\big)$, $y_0 = (y^0_1, y^0_2, \ldots, y^0_p)$, $\varepsilon \ge 0$ and $M = \mathbb{R}^p_+ \cap (\varepsilon\mathbb{B}_q)^c$. Firstly, let us calculate the expression of the generalized Chebyshev norm $\|\ \|_q$. For each $s > 0$ and $y = (y_1, y_2, \ldots, y_p) \in \mathbb{R}^p$ it is clear that

$$y \in (-sq + \mathbb{R}^p_+) \cap (sq - \mathbb{R}^p_+) \iff s \ge \pm\, y_i/q_i, \quad i = 1, 2, \ldots, p$$

and so

$$\|y\|_q = \max_{1\le i\le p}\{|y_i|/q_i\}.$$

Let us observe that $\|\ \|_q$ is the l_∞ norm when $q = (1, 1, \ldots, 1)$.

Moreover, if $\varepsilon = 0$ then $y - y_0 \in sq - \mathrm{cl}(M) = sq - \mathbb{R}^p_+$ if and only if

$$s \ge \max_{1\le i\le p}\left\{\frac{y_i - y^0_i}{q_i}\right\}, \tag{30}$$

and if $\varepsilon > 0$ then $y - y_0 \in sq - \mathrm{cl}(M) = sq - (\mathbb{R}^p_+ \cap (\mathrm{int}(\varepsilon\mathbb{B}_q))^c)$ if and only if (30) holds and

$$\|y - y_0 - sq\|_q = \max_{1\le i\le p}\left\{\frac{|y_i - y^0_i - sq_i|}{q_i}\right\} \ge \varepsilon.$$

Therefore, for each $y \in \mathbb{R}^p$ and $\varepsilon > 0$ it follows that

$$y - y_0 \in sq - \mathrm{cl}(M) \iff$$
$$s \ge \max\left\{\max_{1\le i\le p}\left\{\frac{y_i - y^0_i}{q_i}\right\}, \varepsilon + \min_{1\le i\le p}\left\{\frac{y_i - y^0_i}{q_i}\right\}\right\}. \tag{31}$$

Hence, by (30) and (31) we conclude that

$$\varphi_{q,\varepsilon,y_0}(y) = \max\left\{\max_{1\le i\le p}\left\{\frac{y_i - y^0_i}{q_i}\right\}, \varepsilon + \min_{1\le i\le p}\left\{\frac{y_i - y^0_i}{q_i}\right\}\right\}, \ \forall\,\varepsilon \ge 0. \tag{32}$$

Example 3. Let us characterize the ε-efficiency set of the problem (1) given by the following data: $n = p = 2$, $D = \mathbb{R}^2_+$, $q = (1, 1)$, $0 < \varepsilon < 1$, $f(x_1, x_2) = (x_1, x_2)$ and

$$S = \{(x_1, x_2) \in \mathbb{R}^2 : \ x_1 \ge 0,\ x_2 \ge 0,\ \max\{x_1, x_2\} \ge 1\}.$$

By applying Theorem 5(i) and (32) it follows that if $x_0 = (x_1^0, x_2^0) \in S$ is an ε-efficient solution of this Pareto problem then x_0 is an approximate solution with precision ε of the scalarization (5) given by the same feasible set S and the following objective function:

$$\begin{aligned}&(\varphi_{q,\varepsilon,f(x_0)} \circ f)(x)\\ &= \max\left\{\max_{i=1,2}\left\{x_i - x_i^0\right\}, \varepsilon + \min_{i=1,2}\left\{x_i - x_i^0\right\}\right\}, \quad \forall\, x = (x_1, x_2) \in \mathbb{R}^2.\end{aligned}$$

As $(\varphi_{q,\varepsilon,f(x_0)} \circ f)(x_0) = \varepsilon$ then $x_0 \in \operatorname{AMin}(\varphi_{q,\varepsilon,f(x_0)} \circ f, S, \varepsilon)$ if and only if

$$\max\left\{\max_{i=1,2}\left\{x_i - x_i^0\right\}, \varepsilon + \min_{i=1,2}\left\{x_i - x_i^0\right\}\right\} \geq 0, \quad \forall\, x = (x_1, x_2) \in S. \tag{33}$$

It is clear that

$$\max_{i=1,2}\left\{x_i - x_i^0\right\} < 0 \iff (x_1, x_2) \in x_0 - \operatorname{int}(\mathbb{R}_+^2)$$

and so (33) holds if and only if $x_0 \in \operatorname{WE}(f, S)$ or $x_0 \notin \operatorname{WE}(f, S)$ and

$$\varepsilon + \min_{i=1,2}\left\{x_i - x_i^0\right\} \geq 0, \quad (x_1, x_2) \in (x_0 - \operatorname{int}(\mathbb{R}_+^2)) \cap S.$$

Let us denote $g_{x_0}(x_1, x_2) = \varepsilon + \min_{i=1,2}\left\{x_i - x_i^0\right\}$ and consider that $x_0 \notin \operatorname{WE}(f, S)$. Easy calculations give that

$$\begin{aligned}&\inf_{(x_1,x_2)\in(x_0-\operatorname{int}(\mathbb{R}_+^2))\cap S}\{g_{x_0}(x_1, x_2)\}\\ &= \begin{cases} \varepsilon + 1 - x_2^0 & \text{if } 0 < x_1^0 \leq 1 \text{ and } x_2^0 \geq x_1^0 + 1\\ \varepsilon - x_1^0 & \text{if } 0 < x_1^0 \leq 1 \text{ and } 1 < x_2^0 < x_1^0 + 1\\ \varepsilon - x_2^0 & \text{if } 1 < x_1^0 \quad\quad \text{ and } x_2^0 \geq x_1^0\\ \varepsilon - x_1^0 & \text{if } 1 < x_2^0 \quad\quad \text{ and } x_2^0 < x_1^0\\ \varepsilon - x_2^0 & \text{if } 0 < x_2^0 \leq 1 \text{ and } 1 < x_1^0 < x_2^0 + 1\\ \varepsilon + 1 - x_1^0 & \text{if } 0 < x_2^0 \leq 1 \text{ and } x_1^0 \geq x_2^0 + 1 \end{cases}\end{aligned}$$

Then, as $\varepsilon < 1$ it follows that (33) holds if and only if $(x_1^0, x_2^0) \in \operatorname{WE}(f, S)$, or

$$\begin{aligned}(x_1^0, x_2^0) \in \{&(x_1, x_2) \in \mathbb{R}^2 : 0 \leq x_1 \leq \varepsilon,\, x_1 + 1 \leq x_2 \leq 1 + \varepsilon\}\\ &\cup\{(x_1, x_2) \in \mathbb{R}^2 : 0 \leq x_1 \leq \varepsilon,\, 1 \leq x_2 < x_1 + 1\}\\ &\cup\{(x_1, x_2) \in \mathbb{R}^2 : 0 \leq x_2 \leq \varepsilon,\, 1 \leq x_1 < x_2 + 1\}\\ &\cup\{(x_1, x_2) \in \mathbb{R}^2 : 0 \leq x_2 \leq \varepsilon,\, x_2 + 1 \leq x_1 \leq 1 + \varepsilon\}\\ &= [((0,1) + \varepsilon\mathbb{B}_q) \cup ((1,0) + \varepsilon\mathbb{B}_q)] \cap S,\end{aligned}$$

where $\mathbb{B}_q$ is the closed unit ball with the l_∞ norm. Therefore, the feasible points satisfying the necessary condition proved in Theorem 5(i) are

$$\mathrm{WE}(f,S) \cup ((\{(1,0),\,(0,1)\} + \varepsilon\mathbb{B}_q) \cap S). \tag{34}$$

To obtain ε-efficient solutions through the sufficient condition proved in Theorem 5(*ii*) (see Remark 1), we look for feasible points $x_0 \in S$ such that

$$\max\left\{\max_{i=1,\,2}\left\{x_i - x_i^0\right\},\, \varepsilon + \min_{i=1,\,2}\left\{x_i - x_i^0\right\}\right\} > 0, \quad \forall\, x = (x_1,x_2) \in S. \tag{35}$$

Following a reasoning as the previous one used in applying the necessary condition it is easy to prove that the feasible points satisfying (35) are $(x_1^0, x_2^0) \in \mathrm{E}(f,S)$, or

$$\begin{aligned}(x_1^0,x_2^0) \in \{(x_1,x_2) \in \mathbb{R}^2 : 0 \le x_1 \le \varepsilon,\, x_1+1 \le x_2 < 1+\varepsilon\}\\ \cup\{(x_1,x_2) \in \mathbb{R}^2 : 0 \le x_1 < \varepsilon,\, 1 \le x_2 < x_1+1\}\\ \cup\{(x_1,x_2) \in \mathbb{R}^2 : 0 \le x_2 < \varepsilon,\, 1 \le x_1 < x_2+1\}\\ \cup\{(x_1,x_2) \in \mathbb{R}^2 : 0 \le x_2 \le \varepsilon,\, x_2+1 \le x_1 < 1+\varepsilon\}.\end{aligned}$$

This set of ε-efficient solutions can be written as

$$(\{(1,0),\,(0,1)\} + \mathrm{int}(\varepsilon\mathbb{B}_q)) \cap S. \tag{36}$$

From (34), (36) and Definition 2 is easy to check that

$$\mathrm{AE}(f,S,\varepsilon) = (\{(1,0),\,(0,1)\} + \varepsilon\mathbb{B}_q) \cap S.$$

Under subconvexlikeness hypotheses, the ε-efficient solutions of a vector optimization problem can be characterized via linear functions. This fact is shown in the following theorem, whose proof is direct from Theorem 3, Example 1 and Theorem 4.

Theorem 6. *Let us consider that problem* (1) *is externally stable with respect to the weak efficiency set and the objective function f is D-subconvexlike on the feasible set S. Then,*

(*i*) *If $x_0 \in \mathrm{AE}(f,S,\varepsilon)$ then there exists $l \in D^+$, $\|l\|_* = 1$ such that $x_0 \in \mathrm{AMin}(\langle l, f(\cdot)\rangle, S, \varepsilon)$.*
(*ii*) *Let $l \in D^{+s}$. If $x_0 \in \mathrm{AMin}(\langle l, f(\cdot)\rangle, S, \varepsilon)$ then $x_0 \in \mathrm{AE}(f,S,\varepsilon/\alpha)$, where*

$$\alpha = \min_{d\in D, \|d\|=1}\{\langle l, d\rangle\}.$$

5 Conclusions

In this paper, approximate solutions of vector optimization problems are analyzed via an ε-efficiency concept introduced by Tanaka [18]. This kind of solutions is important from both the practical and theoretical points of view since they exist under very mild hypotheses and they are obtained by a lot of solution methods (for example, by iterative and heuristic methods).

An interesting problem concerning ε-efficient solutions is to relate the approximate solutions of a vector optimization problem with approximate solutions obtained by solution methods based on scalarization processes. This question has been widely studied in the literature [2, 5, 6, 7, 12, 13], but using ε-efficiency concepts based on a previously fixed scalar function or via not metrically consistent ε-efficiency notions.

In this work, Tanaka's concept is analyzed from this point of view. Specifically, necessary and sufficient conditions for these approximate solutions are established via properties (AORP) and (SLM). Property (AORP) extends the classical order representing property and (SLM) is a new generalized monotonicity notion.

As Tanaka's ε-efficiency concept is metrically consistent, properties (AORP) and (SLM) ensure that improved ε-efficient solutions are obtained when the scalar objective decreases and reciprocally. Moreover, from these properties it is possible to estimate the precision ε of an approximate efficient solution obtained by scalarization, through the precision of this solution in the scalar problem.

By using gauge functions and generalized Chebyshev norms, we get a family of functions satisfying properties (AORP) and (SLM), and from the previous results we obtain a characterization for ε-efficient solutions which attains the same precision in the vector problem as in the scalarization. Also, ε-efficient solutions are characterized by separation theorems in cone-subconvexlike vector optimization problems.

Acknowledgements

The authors are grateful to professor Dinh The Luc and an anonymous referee for their helpful comments and suggestions.

References

1. Deng S (1997) On approximate solutions in convex vector optimization. SIAM J. Control Optim. 35(6):2128–2136
2. Dutta J, Vetrivel V (2001) On approximate minima in vector optimization. Numer. Funct. Anal. Optim. 22(7&8):845–859
3. Frenk JBG, Kassay G (1999) On classes of generalized convex functions, Gordan-Farkas type theorems, and Lagrangian duality. J. Optim. Theory Appl. 102(2):315–343
4. Gerth C, Weidner P (1990) Nonconvex separation theorems and some applications in vector optimization. J. Optim. Theory Appl. 67(2):297–320
5. Göpfert A, Riahi H, Tammer C, Zălinescu C (2003) Variational methods in partially ordered spaces. CMS Books in Mathematics. Springer, New York

6. Govil MG, Mehra A (2004) ε-optimality for multiobjective programming on a Banach space. European J. Oper. Res. 157:106–112
7. Gutiérrez C, Jiménez B, Novo V (2005) A chain rule for ε-subdifferentials with applications to approximate solutions in convex Pareto problems. J. Math. Anal. Appl. 310:309–327
8. Gutiérrez C, Jiménez B, Novo V (2005) Multiplier rules and saddle-point theorems for Helbig's approximate solutions in convex Pareto problems. J. Global Optim. 32:367–383
9. Helbig S, Pateva D (1994) On several concepts for ε-efficiency. OR Spektrum 16:179–186
10. Jahn J (2004) Vector optimization. Theory, applications, and extensions. Springer, Berlin
11. Kutateladze SS (1979) Convex ε-programming. Soviet Math. Dokl. 20(2):391–393
12. Li Z, Wang S (1998) ε-approximate solutions in multiobjective optimization. Optimization 44:161–174
13. Liu JC, Yokoyama K (1999) ε-optimality and duality for multiobjective fractional programming. Comput. Math. Appl. 37:119–128
14. Luc DT (1989) Theory of vector optimization. Lecture Notes in Econom. and Math. Systems 319, Springer-Verlag, Berlin
15. Rong WD (1997) Proper ε-efficiency in vector optimization problems with cone-subconvexlikeness. Acta Sci. Natur. Univ. NeiMongol 28:609–613
16. Rubinov A (1977) Sublinear operators and their applications. Russian Math. Surveys 32:115–175
17. Tammer C (1994) Stability results for approximately efficient solutions. OR Spektrum 16:47–52
18. Tanaka T (1995) A new approach to approximation of solutions in vector optimization problems. In: Fushimi M, Tone K (eds) Proceedings of APORS, 1994, pages 497–504. World Scientific Publishing, Singapore
19. Vályi I (1985) Approximate solutions of vector optimization problems. In: Sydow A, Thoma M, Vichnevetsky R (eds) Systems analysis and simulation, pages 246–250. Akademie-Verlag, Berlin
20. White DJ (1986) Epsilon efficiency. J. Optim. Theory Appl. 49(2):319–337
21. Wierzbicki AP (1986) On the completeness and constructiveness of parametric characterizations to vector optimization problems. OR Spektrum 8:73–87
22. Yokoyama K (1996) Epsilon approximate solutions for multiobjective programming problems. J. Math. Anal. Appl. 203:142–149
23. Yokoyama K (1999) Relationships between efficient set and ε-efficient set. In: Proceedings of the International Conference on Nonlinear Analysis and Convex Analysis, 1999, pages 376–380. World Scientific Publishing, New York

On the Work of W. Oettli in Generalized Convexity and Nonconvex Optimization – a Review and Some Perspectives

Joachim Gwinner

Universität der Bundeswehr München, Department of Aerospace Engineering, D-85577 Neubiberg/Munich, Germany Joachim.Gwinner@Unibw-Muenchen.de

More than five years ago, Professor Oettli died prematurely at the age of 62. He was a highly productive scholar. He authored about one hundred publications in nonlinear analysis and optimization, devoted to topics such as convex and variational analysis, duality and minimax theory, quadratic programming, optimization methods, nonconvex and global optimization, vector and set optimization, mathematical economics, networks, variational inequalities and equilibria, generalized convexity and generalized monotonicity. As an active researcher in generalized convexity, he participated in several of the International Conferences on Generalized Convexity/Generalized Monotonicity, including the first one in Vancouver in 1980.
It is the purpose of this paper to give a review of the work of the late W. Oettli in generalized convexity and nonconvex optimization. So only a small part of the whole ouvre will be covered. In particular we do not enter his work in generalized monotonicity. On the other side, we also discuss some possible extensions and sketch out various relations to the more recent literature in generalized convexity. Thus we intend to give some perspectives of his work.

We shall subdivide this paper in 4 sections. In section 1 we begin with papers of W. Oettli together with co-authors that combine optimization theory and optimization methods for the solution of some difficult nonconvex optimization problems. In section 2 we turn to his work on numerical methods for various nonconvex optimization problems, in particular to the versatile joint work with Le Dung Muu on branch and bound algorithms in global optimization. Then we switch to the achievements of W. Oettli in nonconvex and global optimization theory. In section 3 we focus on the seminal work of W. Oettli and co-authors, in particular with V. Jeyakumar on solvability theorems for nonconvex functions and their applications in quasiconvex and dc (difference convex) programming. Section 4 concludes with conjugate duality and optimality conditions in nonconvex optimization, including his profound work on abstract quasiconvex programming via a perturbation approach and the important joint work with Flores-Bazán that presents an axiomatic approach to nonconvex vector optimization.

1 Solution of Some Nonconvex Optimization Problems

1.1 A Quasiconvex Problem from Information Theory

In their joint paper [21] B. Meister and W. Oettli treat an optimization problem from information theory that originated from their then work at IBM Zürich Research Laboratory. They are concerned with the relative capacity of a discrete constant channel which is given as the maximal transmission rate over all admissible input distributions. To attack this problem by methods of mathematical programming, they rewrite this problem as the optimization problem

$$\max T(z) := \frac{\langle a, x\rangle - \langle y, \log y\rangle}{\langle t, x\rangle} \equiv \frac{f(z)}{g(z)}$$

with given vectors a and $t > 0$ (i.e. all components $t_i > 0$) on the compact convex polyhedron

$$Z := \Big\{z = (x, y) | x \geq 0, \sum_j x_j = 1, y = Px\Big\}.$$

Here $(\log y)_i = \log y_i$, $\langle \cdot, \cdot \rangle$ denotes the scalar product, and P is a transition matrix with conditional probabilities as entries where P can be assumed to contain no zero now. The functions f and g are positive in Z, further the target function T is continuous in Z and continuously differentiable in the nonempty set

$$Z^0 := \big\{z \in Z | y > 0\big\}.$$

The numerator function f is known to be concave. However due to the linear denominator g, T is only a quasiconcave function in Z.
In order to build an iterative solver for quasiconcave maximization, suitable for target functions of the form $T = \dfrac{f}{g}$ with differentiable f and g (not necessarily linear), they linearize numerator and denominator, separately, and introduce the function

$$\tau_{z^1}(z) = \frac{f(z^1) + \langle f'(z^1), z - z^1\rangle}{g(z^1) + \langle g'(z^1), z - z^1\rangle}\,.$$

Thus they are able to characterize (Theorem 1) an optimal solution $\hat{z}$ of the problem by

$$\hat{z} \in Z^0 \text{ and } \max_{z \in Z} \tau_{\hat{z}}(z) = \tau_{\hat{z}}(\hat{z})\,.$$

Then they propose the following iterative solver (that modifies the well-known Frank-Wolfe algorithm for quadratic programming):

1. Start with $z^1 \in Z^0$ arbitrarily
2. Given $z^k \in Z^0$, determine $\tilde{z}^k \in Z$ such that

$$\tau_{z^k}(\tilde{z}^k) = \max_{z \in Z} \tau_{z^k}(z)$$

3. Determine z^{k+1} in the line segment $[\tilde{z}^k, z^k]$ such that

$$T(z^{k+1}) = \max_{z \in [\tilde{z}^k, z^k]} T(z).$$

Note that the auxiliary problem in step 2 is a fractional linear programming problem which is directly solved in [21]. Theorem 2 of this paper and its longer proof tell us that the proposed method converges to an optimal solution of the problem, with the extra benefit of converging lower and upper bounds for the sought capacity. Moreover, it is remarked that the characterization result, the iterative solver, and its convergence proof extend to a larger class of nonlinear denominators.

1.2 A Reverse Convex Optimization Problem

In their joint work [35] W. Oettli together with P.T. Thach and R.E. Burkard consider the optimization problem

$$(P) \quad \min f(x) \text{ such that } x \in G \text{ and } Tx \notin \text{ int } D,$$

where

$G \neq \emptyset$ compact (and convex) $\subset \mathbb{R}^n$
$f : G \to \mathbb{R}$ lower semicontinuous (and convex)
$T : \mathbb{R}^n \to \mathbb{R}^d$ continuous (but not necessarily linear)
$D \neq \emptyset$ closed convex $\subset \mathbb{R}^d$ with int $D \neq \emptyset$.

The difficulty of (P) comes from the last constraint $Tx \notin$ int D, "*a reverse convex constraint*". Therefore (P) is a hard optimization problem; e.g. with G a polytype, D the unit ball in $\mathbb{R}^n$ $(T = 1, d = n)$, (P) becomes a set containment problem, known to be NP-complete.
The size of reverse convex programs that are solvable to optimality is very limited. Let us take from [35] the subsequent *example.*
Let G be a convex subset of $\mathbb{R}^n_+$; $T = (T_1, T_2)$ with $T_1(x) = c^T x, T_2(x) = d^T x; c, d \in \mathbb{R}^n_+$, $D =$ epi φ for $\varphi(v_1) = \dfrac{1}{v_1}, v_1 > 0$. Then $Tx \notin$ int D is equivalent to $(c^T x) \cdot (d^T x) \leq 1$. Such constraints involving the product of two linear functions occur in various applications. Nonconvex optimization problems of this kind have been considered by several authors (see the references in [35] and the more recent papers [13, 31, 20, 17, 14, 32]).

The key for the effective solution of (P) is a reduction to a quasiconcave program as follows. Assume that the feasible set $\{x \in G : T(x) \notin \text{ int } D\} \neq \emptyset$.

Let $w \in \text{argmin}\,\{f(x) : x \in G\}$ (Note that the latter is a convex program and there exists a minimizer w). If $T(w) \notin \text{int}\ D$, then w solves (P), stop.
Therefore consider in the following the case $T(w) \in \text{int}\ D$. Let $V := D - T(w)$. Then V is convex with $0 \in \text{int}\ D$ and the polar $E := V^0$ is $\neq \emptyset$, convex and compact.
Now rewrite the reverse convex constraint by the separation theorem: $T(x) \notin \text{int}\ D \Leftrightarrow\ \exists\, u \in E : \langle u, T(x) - T(w)\rangle \geq 1$. Thus

$$\begin{aligned}\inf(P) &= \inf\Big\{f(x) : x \in G;\ \exists\, u \in E : \langle u, T(x) - T(w)\rangle \geq 1\Big\}\\ &= \inf_{u\in E}\inf_{x}\Big\{f(x) : x \in G, \langle u, T(x) - T(w)\rangle \geq 1\Big\} = \inf\Big\{h(u) : u \in E\Big\},\end{aligned}$$

where for $u \in \mathbb{R}^d$

$$h(u) := \inf\Big\{f(x) : x \in G, \langle u, T(x) - T(w)\rangle \geq 1\Big\} =: \inf(Q_u)\,.$$

Note that (Q_u) has an optimal solution because of continuity and compactness; moreover (Q_u) is a convex program, provided T is linear. Then with $\inf\emptyset = +\infty$, $h : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ and is seen to be *quasiconcave*. In addition, h is lower semicontinuous. Hence

$$(\widetilde{P}) \quad \min h(u) \text{ such that } u \in E$$

has an optimal solution. Altogether [35] concludes that $\inf(P) = \inf(\widetilde{P})$; if u^* is an optimal solution of $(\widetilde{P})$ and x^* solves (Q_{u^*}), then x^* solves (P).
The solution method in [35] uses the concept of approximate solutions. Let $\{D_\varepsilon : \varepsilon \geq 0\}$ be a family of subsets $\subset \mathbb{R}^d$ such that $D_0 = \text{int}\ D$; for $\varepsilon > 0, D_\varepsilon$ is closed $\subset \text{int}\ D$. Then the constraint $T(x) \notin D_\varepsilon$ is a relaxation of $T(x) \notin \text{int}\ D$. In this sense $x \in \mathbb{R}^n$ is called (η, ε)–optimal (with $\eta > 0, \varepsilon > 0$), if $x \in G, T(x) \notin D_\varepsilon$ and $f(x) < \inf(P) + \eta$.
The algorithm in [35] is essentially an outer approximation method based on cutting planes specialized to $(\widetilde{P})$ and is proved to terminate for any given $\eta > 0, \varepsilon > 0$ after finitely many iterations with an (η, ε)-optimal solution.

Possible Extensions and Some Relations to More Recent Work

There is a *duality theory* due to Thach [34] between

quasiconvex minimization with a reverse convex constraint
and quasiconvex maximization over a convex set.

In this vein, (P) can be considered as a primal program, $(\widetilde{P})$ as its dual.

More recently, in [15] Lemaire develops a *duality theory* for the general problem of minimizing an extended real-valued convex function on a locally convex linear space under a reverse convex constraint. More precisely, he considers

$$(P_{\mathcal{L}}) \quad \min g_{\mathcal{L}}(x) \text{ such that } h_{\mathcal{L}}(x) > 0,$$

where $g_{\mathcal{L}}$ and $h_{\mathcal{L}}$ are extended real-valued and convex. The original program (P) can be rewritten as a program of the form $(P_{\mathcal{L}})$ under extra conditions as follows. Assume that the implicit constraint $x \in G$ is subsumed in the objective function f and that $\inf(P)$ is finite; however, we have to suppose that f is upper semicontinuous. Moreover, suppose that T is linear.
Then using the above construction involving the minimizer w and the polar $E = V^{\circ}$, we obtain that

$$(P) \quad \min f(x) \text{ such that } x \in \mathbb{R}^n \text{ and } Tx \notin \text{ int } D$$

is equivalent to

$$(P_{\geq}) \quad \min f(x) \text{ such that } h_{\mathcal{L}}(x) \geq 0,$$

where $h_{\mathcal{L}}(x) = \max\{\langle u, Tx - Tw\rangle : u \in E\} - 1$ is convex. Now $h_{\mathcal{L}}(w) = -1 < 0$ and by continuity and convexity arguments, inf $(P_{\geq}) =$ inf $(P_{>})$ follows with

$$(P_{>}) \min f(x) \text{ such that } h_{\mathcal{L}}(x) > 0.$$

Indeed, inf $(P_{>}) \geq$ inf $(P_{\geq}) =$ inf $(P) \in \mathbb{R}$. Assume $\delta =$ inf $(P_{>}) -$ inf $(P_{\geq}) > 0$. We find some x_1 with $h_{\mathcal{L}}(x_1) \geq 0$ and $f(x_1) <$inf $(P_{\geq}) + \frac{\delta}{2}$. If $h_{\mathcal{L}}(x_1) > 0$, then inf $(P_{>}) \leq f(x_1) <$ inf $(P_{\geq}) + \delta =$ inf $(P_{>})$ and a contradiction is reached. If otherwise $h_{\mathcal{L}}(x_1) = 0$, consider $x_{\lambda} = \lambda w + (1-\lambda)x_1$ for $\lambda < 0$. Then by convexity of $h_{\mathcal{L}}, h_{\mathcal{L}}(x_{\lambda}) > 0$ for $\lambda < 0$. Finally, since f is usc, $f(x_{\lambda}) <$ inf $(P_{\geq}) + \delta$ for $x_{\lambda} = x_1 + \lambda(w - x_1)$ and $|\lambda| > 0$ small enough leading to a contradiction.
On the other hand, when one changes the original constraint $T(x) \notin$ int D in (P) to $Tx \notin D$ (with T linear), then the duality theory of [15] directly applies. Then one chooses

$$g_{\mathcal{L}}(x) = \begin{cases} f(x) \in G; \\ +\infty \text{ else} \end{cases}$$
$$h_{\mathcal{L}}(x) = \max\big\{\langle T^{*}u, x - w\rangle : u \in E\big\} - 1$$

and thus obtains

$$(P_{\mathcal{L}}) \quad \min f(x) \text{ subject to } x \in G,\ Tx - Tw \notin V.$$

Then by the classical minimax theorem, one computes

$$h_{\mathcal{L}}^{*}(y) = \begin{cases} \langle y, w\rangle + 1 \text{ if } y \in T^{*}E; \\ +\infty \qquad \text{else} \end{cases}$$
$$Y = \big\{y \in \mathbb{R}^n : \langle y, x\rangle > h_{\mathcal{L}}^{*}(y) \text{ for some } x \in G\big\}$$
$$= \big\{T^{*}u : \langle T^{*}u, x - w\rangle > 1 \text{ for some } x \in G\big\}$$

and

$$\sup_{\lambda \geq 0} \left\{ \lambda h^*_{\mathcal{L}}(y) - g^*_{\mathcal{L}}(\lambda y) \right\} = \inf \left\{ f(x) : x \in G, \langle y, x - w \rangle \geq 1 \right\}$$

and by Theorem 4.1 in [15] one arrives at the duality relation $\inf(P_{\mathcal{L}}) = \inf(D_{\mathcal{L}})$ for the dual problem

$$\begin{aligned} (D_{\mathcal{L}}) \ \ & \min \ \inf \left\{ f(x) : x \in G, \langle y, x - w \rangle \geq 1 \right\} \\ & \text{subject to } y \in Y. \end{aligned}$$

Note that $(D_{\mathcal{L}})$ is equivalent to

$$\begin{aligned} & \min h(u) = \ \inf \left\{ f(x) : x \in G, \langle T^* u, x - w \rangle \geq 1 \right\} \\ & \text{subject to } T^* u \in Y \end{aligned}$$

which coincides with $(\widetilde{P})$ apart from the strict inequality in the definition of the set Y.

Concerning the *numerical solution* note that the dual $(\widetilde{P})$ can be solved much easier than the primal (P), if $d << n$. There are algorithms for linearly constrained quasiconvave minimization subproblems in d dimensions at each iteration (see the monograph [7] of Horst and Tuy). As an alternative to the outer approximation method employed in [35] for the solution of $(\widetilde{P})$ one can think of branch and bound methods which was also a field of research of W. Oettli together with Le Dung Muu as we shall see in the next section.

2 Some Numerical Methods for Various Nonconvex Optimization Problems

2.1 Decomposition Methods for Saddle Points of Quasi-convex-concave Functions

In [27] Oettli presents decomposition methods for finding saddle points of a quasi-convex-concave function $\varphi : X \times Y \to \mathbb{R}$ (i.e. $\varphi(\cdot, y) : X \to \mathbb{R}$ is quasiconvex for all $y \in Y$ and $\varphi(x, \cdot) : Y \to \mathbb{R}$ is quasiconcave for all $x \in X$), where X, Y are nonvoid closed convex sets in some linear normed spaces. Here decomposition methods consist in an alternating succession of master programs and subprograms.

In the master programs, the proper iteration points are determined as approximate saddle points over a subset $X^n \times Y^n$ of the given domain $X \times Y$. In the subprograms auxiliary points are calculated that serve to update the subset under consideration. For certain structured problems the subprograms may decompose; this fact accounts for the name and enhances its practical usage of decomposition methods, but this is not the impetus of the convergence theory as presented in [27].

By this paper Oettli succeeds in unifying and extending prior decomposition methods that are in particular due to Auslender, Cohen, Dantzig, Huard and

Zangwill. Here starting from Sion's celebrated minimax theorem, the compactness assumption for the underlying domain is relaxed and also regularizations in solving the subproblems are admitted. The decomposition principle is described in terms of a variational inequality problem and thus the extension to Nash equilibrium points in n-person games becomes straightforward.
Moreover under the extra hypothesis that $\varphi(x, \cdot)$ is unimodal on Y for all $x \in X$, versions of the original method are obtained that do not need the storage of the auxiliary points and instead allow for the deletion of auxiliary points. By this hypothesis, the master program can be drastically simplified towards a method of feasible directions. Under additional compactness assumptions (but not requiring $X \times Y$ to be compact) every cluster point of the sequence generated by the algorithm is proved to be a saddle point of φ on $X \times Y$. Even, in a more specialized version, estimates of the rate of convergence can be established.

2.2 Some Branch and Bound Methods in Global Optimization

In this subsection we summarize joint work of W. Oettli with Le Dung Muu in [22, 23, 24, 25] on branch and bound algorithms [7] for solving various classes of global optimization problems.
In [22] the authors present a new branch and bound method for minimizing an indefinite quadratic function

$$f(x, y) = p^T x + x^T M y + q^T y$$

on a given closed convex non-empty set $S \subset R^n \times R^m$, where $p \in R^n$ and $q \in R^m$ are given vectors and M is a given $n \times m$ matrix. The branching here is a simplex bisection and the bounding is obtained by the solution of $(m+1)$ convex subprograms and in the case if S is a polyhedron these subprograms are even linear.
In [23] the authors propose a branch and bound method for minimizing a convex-concave function over a convex set. The bounding operation is essentially the same as in the previous paper. The difference here is the branching operation that is based on bisection of rectangles, taking into account the current iteration point obtained by the bounding operation. An important special case is the minimization of a dc-function (i.e. a function representable as the difference of two convex functions). In this case, the subproblems occurring in the bounding operation can be solved effectively as shown by a numerical example.
In [24] the authors exhibit unified branch and bound and cutting plane algorithms for global minimization of a function $f(x, y)$ over a certain closed set. By formulating the problem in terms of two groups of variables and two groups of constraints they arrive at new relaxation bounding and adaptive branching operations. The branching operation takes place in y-space only and uses the iteration points obtained through the bounding operation. The cutting

is performed in parallel with the branch and bound procedure. The method can be applied implementably for a certain class of nonconvex programming problems.
Finally in this subsection we review more detailed the paper [25] that gives an interesting link between global optimization and vector optimization. First generalizing the well-known notion of Pareto efficient points the authors call a point x in some given set X an equilibrium point iff, for some $\lambda \in \Lambda$ (a nondegenerate p–simplex in $\mathbb{R}^p$)

$$c(\lambda, x) \geq c(\lambda, y), \ \forall\, y \in X,$$

where $c : \Lambda \times X \to \mathbb{R}$ is given. Note that even if X is convex and $c(\lambda, \cdot)$ is linear the set of equilibrium points of X is generally not convex. Therefore the problem of maximizing a function f over the set of equilibrium points leads to a difficult problem of global optimization.
Prior work on maximization over the Pareto efficient points was done under the assumption that essentially f is quasiconvex, thus obtaining the maximum at an extreme point. Here the authors drop this assumption and present a branch and bound method in the criteria space for approximately solving the problem

$$P \equiv P(\Lambda) \quad \max\Big\{f(\lambda, x) | \lambda \in \Lambda, x \in X, c(\lambda, x) \geq c(\lambda, y)\ \forall\, y \in X\Big\}.$$

To describe the method, let $\nu(\Lambda)$ denote the optimal value of problem $P(\Lambda)$, likewise $\nu(S)$ the optimal value of problem $P(S)$, where Λ is replaced by a p–simplex $S \subseteq \Lambda$. In iteration k, we have a family Γ_k of subsimplices $S \subset \Lambda$ and some (ω_k, z_k) feasible for P such that $(\omega_k, z_k) \cup \bigcup\{S : S \in \Gamma_k\}$ contains a solution of P. The value $\beta_k := f(\omega_k, z_k) \leq \nu(\Lambda)$ is the best lower bound for $\nu(\Lambda)$ available in the current step. For all $S \in \Gamma_k$ let there be given an upper bound $\alpha(S) \geq \nu(S)$. Set $\alpha_k := \sup\Big\{\alpha(S) : S \in \Gamma_k\Big\}$. Then $\max\{\beta_k, \alpha_k\} \geq \nu(\lambda)$. If $\alpha_k \leq \beta_k$, then $\beta_k = \nu(\lambda)$, hence (ω_k, z_k) solves (P), and the method terminates. Otherwise, if $\alpha_k > \beta_k$, then delete from Γ_k all $S \in \Gamma_k$ with $\alpha(S) \leq \beta_k$. This gives a reduced family $R_k \subseteq \Gamma_k, R_k \neq \emptyset$, of subsimplices $S \subseteq \Lambda$. Select those $S_k \in R_k$ such that $\alpha(S_k) = \alpha_k \geq \nu(\Lambda)$. Then bisect those S_k into two subsimplices $S_{k,1}$ and $S_{k,2}$. Determine upper bounds $\alpha(S_{k,i}) \geq \nu(S_{k,i})\,(i = 1, 2)$ such that $\alpha(S_{k,i}) \leq \alpha_k$ and determine (ω_{k+1}, z_{k+1}) feasible for P such that $\beta_{k+1} := f(\omega_{k+1}, z_{k+1}) \geq \beta_k$. Finally delete S_k from R_k and add $S_{k,1}$ and $S_{k,2}$, thus obtaining the family Γ_{k+1} for the next iteration, with $\alpha_{k+1} \leq \alpha_k$.
For bisection the authors employ a simplex bisection devised by Horst and later refined by Tuy. For bounding the authors suggest the solution of two simpler structured optimization problems, a "feasibility problem" (to construct (ω_{k+1}, z_{k+1}) above) and a "relaxed subproblem" for each bisected $S_{k,i}$.
In this general setting the authors prove that for any $\varepsilon > 0$ the method terminates after a finite number of steps either with an exact solution or with

an ε–solution (that is ε–feasible and ε–optimal), if $\varepsilon = 0$ and there is no termination, then a suitable subsequence of the generated sequence converges to a solution.
The paper concludes with the discussion of an implementable version under additional assumptions: X is convex, $f(\lambda, x) = f(x)$ is concave, $c(\lambda, x)$ is affine in λ and concave in x. Then the feasibility problem becomes a standard concave maximization problem and the relaxed subproblem reduces to $p+1$ standard concave maximization problems.
Still the global optimization problem over efficient sets is in the fore of research. Let us only mention the recent paper of Le Thi Hoai An, Le Dung Muu, and Pham Dinh Tao [16] who formulate optimization problems over efficient and weakly efficient sets as dc problems over a simplex and develop a decomposition algorithm using an adaptive simplex subdivision.

3 Nonconvex Solvability Theorems with Applications in Quasiconvex and DC Programming

Solvability theorems (also called transposition theorems or theorems of the alternative) have been established by the use of the Hahn-Banach theorem (or its equivalents) or of an appropriate minimax theorem. These theorems have become an important tool to derive various results in optimization theory, e.g. the existence of Lagrange multipliers and first order F. John or Kuhn Tucker optimality conditions, duality results, scalarization of vector-valued objectives. Here we focus on the work of W. Oettli and co-authors on solvability theorems for nonconvex functions and their applications to some classes of nonconvex optimization problems and review the papers [4, 5, 12].

3.1 Towards Optimality Conditions in Quasiconvex Optimization

Here we summarize the joint work of W. Oettli with V. Jeyakumar and M. Natividad in [12]. In the adopted setting of topological vector spaces X, Y with $C \subseteq X$ convex and $P \subseteq Y$ closed convex cone, the notion of a quasiconvex (real-valued) function on C extends readily to a quasiconvex map $f : C \to Y$ by imposing $\forall\, x_1, x_2 \in C, y \in Y$

$$f(x_1) \in y - P \wedge f(x_2) \in y - P \Rightarrow f(\xi) \in y - P,\ \forall\, \xi \in [x_1, x_2]\,.$$

However, as the authors show by a counterexample, already in finite dimensions with a polyhdreal cone, quasiconvex maps do not satisfy a solvability theorem that typically states that *either* the system

$$i) \qquad x \in C, \quad \forall \lambda \in P^* \backslash \{0\} \quad \langle \lambda, f(x) \rangle < 0$$

has a solution, *or*

$$ii) \qquad \exists\, y^* \in P^*\backslash\{0\},\; \forall\, x \in C \;\; \langle y^*, f(x)\rangle \geq 0\,.$$

Therefore the authors have to identify a suitable subclass of quasiconvex maps. They introduce socalled "$*$-quasiconvex" ($*$-qv) maps f by imposing the function $\langle y^*, f(\cdot)\rangle$ to be quasiconvex for all $y^* \in P^*$. Moreover, they assume that the map $f : C \to Y$ is "$*$-lsc", i.e. $\forall\, y^* \in P^*$, the function $\langle y^*, f(\cdot)\rangle$ is lsc (lower semicontinuous). For this subclass, under the assumption that the nonnegative polar P^+ admits a $\sigma(Y^*, Y)$-compact base B (i.e. $0 \notin B, P^* = \mathbb{R}_+ B$), the solvability theorem is shown to be true. Its proof is based on the application of the celebrated minimax theorem of Sion to the function $\varphi(x, y^*) = \langle y^*, f(x)\rangle$ on $C \times B$.
The solvability theorem obtained is used in two ways. First in the global theory considering the constrained problem

$$\text{(CP)} \;\; \min f_0(x) \text{ subject to } x \in C, g(x) \in -P$$

assume that $(f_0, g) : C \to \mathbb{R} \times Y$ be $*$-qv with respect to $\mathbb{R}_+ \times P^*$, moreover let f_0 lsc and g $*$-lsc. Suppose, inf (CP) is finite. Then under a Slater type constraint condition it is shown that there exists a Lagrange multiplier $\lambda^* \in P^*$ that satisfies

$$\inf \text{(CP)} \;\leq f_0(x) + \langle \lambda^*, g(x)\rangle,\; \forall\, x \in C\,.$$

Secondly, necessary optimality conditions for a local minimum of (CP) can be presented. To this end, consider $h : C \to Y; a \in C$; then a map $\varphi : C \to Y$ is called a "$*$–upper approximation" to h at a, if $\varphi(a) = h(a)$ and $\forall\, x \in C\; \exists\, o(\cdot) : [0,1] \to \mathbb{R}$:

$$\langle y^*, h(\tau x + (1-\tau)a)\rangle \leq \tau\langle y^*, \varphi(x)\rangle + (1-\tau)\langle y^*, \varphi(a)\rangle + o(\tau), \forall \tau \in [0,1], y^* \in B\,.$$

By this generalized differentiability approach the authors arrive at the following F. John type result: Assume (CP) attains a local minimum at $a \in C$. Suppose, $\varphi_0 : C \to \mathbb{R}$ is lsc and an upper approximation to f_0 at a. Suppose, $\varphi : C \to Y$ is $*$-lsc and a $*$-upper approximation to g at a. If (φ_0, φ) is $*$-qv, then $\exists\, (\lambda_0, \lambda) \in \mathbb{R} \times P^*, (\lambda_0, \lambda) \neq 0$ such that

$$\lambda_0 f_0(a) = \lambda_0 \varphi_0(a) \leq \lambda_0 \varphi_0(x) + \langle \lambda^*, \varphi(x)\rangle,\; \forall\, x \in C.$$

Clearly, the multiplier λ_0 can be shown to be nonzero under an appropriate Slater type constraint qualification.

Some Comments on Quasiconvexity for Vector-Valued Functions and Related Recent Work

Note that there are quasiconvex maps that are not $*$-qv , but satisfy the solvability theorem. This comes out from the following simple

Example. Let $Y = \mathbb{R}^2, C = \mathbb{R}_+, P = P^* = \mathbb{R}^2_+$. Consider $f = (f_1, f_2), f_1(x) = x, f_2(x) = -x^2$. Then f_1 and f_2 are qv and so f is qv. But f is not $*$-qv (consider $y^* = (2,1)$) On the other hand, i) is not true and ii) holds with $y^* = (1,0)$.

Thus it may be worthwhile to revisit optimality in quasiconvex optimization. Another generalization of the classical notion of quasiconvex real-valued functions to the vector-valued case is the geometrically appealing concept of cone-quasiconvexity, introduced by Dinh The Luc, which states that for all $y \in Y$ the level set $\{x \in C | f(x) \in -P\}$ is convex. The difficulty in scalarizing such vector-valued functions lies in the fact that this concept is not stable under composition with linear forms from the nonnegative polar of P (what per definitionem holds true for $*$-quasiconvex functions). Nevertheless, in the particular case where P is generated by an algebraic base in $\mathbb{R}^n$, Dinh The Luc [18] could characterize cone-quasiconvexity by the stability property that the scalar function $l \circ f$ is quasiconvex for every extreme direction l of P^+. This characterization has been recently extended to closed convex cones in a Banach space by Benoist, Borwein, and Popovici [1] assuming that $(Y, \leq_P)$ is directed (i.e. $(y_1 + P) \cap (y_2 + P) \neq \emptyset$ for all $y_1, y_2 \in Y$) and that the polar P^+ coincides with the weak-star closed convex hull of the set of its extreme directions.

3.2 Towards Optimality Conditions in DC Optimization

In [4] the authors lift the class of almost dc (difference convex) functions due to Gorokhovik [6] and Margolis [19] connected with the order structure of a normed Riesz space to a higher level of abstraction and introduce the class of almost S-DC functions with respect to a order cone S in a general locally convex topological vector space. These almost S-DC functions encompass differences $f = p - q$, where p is S-convex and q is regularly S-sublinear,but nonconvex functions that are the difference of S-convex functions may not be almost S-DC. Solvability theorems for this new class of nonconvex functions are formulated in terms of subdifferentials and established both without assuming any regularity condition and under a Slater-type regularity condition, using the Ky Fan minimax theorem.

In [5] the authors present new versions of theorems of the alternative, including a generalized Farkas lemma, for systems of functions bounded above by sublinear mappings and in particular for (possibly infinite) systems of difference sublinear functions. Here (possibly infinite dimensional) vector functions $g : X \to Y$ are considered such that $\lambda g : X \to \mathbb{R}$ is bounded above by some sublinear mapping p_λ, respectively is a difference sublinear function for any λ in the dual cone of the ordering cone in Y. This scalarization permits to express their theorems of the alternative in terms of the subdifferential $\partial p_\lambda(0)$, respectively in terms of the subdifferential $\underline{\partial}(\lambda g)(0)$ and the superdifferential $\overline{\partial}(\lambda g)(0)$.

In the second part of [5] the authors apply their new theorems of the alternative and derive first order optimality conditions of Karush-Kuhn-Tucker type for weak minima of general quasidifferentiable vector optimization problems with vector-valued quasidifferentiable objective and constraint functions. They illustrate their theory by some short examples discussing some special problems of dc programming, (finite dimensional) quasidifferentiable programming, infinite dimensional concave minimization,and nonsmooth semi-infinite programming. In addition they study their presupposed constraint qualification of local solvability in the setting of finite-dimensional locally Lipschitz and directionally differentiable functions and are able to extend the well-known Robinson stability condition to this nonsmooth setting.

Some Relations to More Recent Work

As the authors already remark, the stability investigations in [5] are essentially finite-dimensional, since they rely critically on the Clarke generalized derivative. The extension to some infinite dimensional (Banach) space setting seems to be still open. In this context we refer to the recent deep study of Jeyakumar and Yen [11] on stability theory of nonsmooth continuous systems in finite dimensions where the recent theory of approximate Jacobians is employed and a new extension of the Robinson regularity condition is used.

On the other hand, the approach to optimality conditions in global optimization via generalized theorems of the alternative, due to W. Oettli, B.M. Glover and V. Jeyakumar, has been extended to general dc minimization in various ways by V. Jeyakumar, A.M. Rubinov, B.M. Glover, Y. Ishizuka [10], respectively by V. Jeyakumar and B.M. Glover [8, 9].

4 Conjugate Duality and Optimality in Nonconvex Optimization

Here we report on the work of W. Oettli in [26], moreover on his more recent work together with Schläger in [28] and with Flores-Bazán in [3].

4.1 Optimization with Quasiconvex Maps via a Perturbation Approach

In [26] the abstract program

$$(P) \quad \min f(x) \text{ subject to } 0_Y \in \Gamma(x), x \in X$$

is investigated, where X, Y are real topological vector spaces, $f : X \to \mathbb{R} \cup \{+\infty\}, \Gamma : X \rightrightarrows Y$ is a multimap. With $\Gamma(x) := \{G(x)\} - C$, $x \in X$ for a map $G : X \to Y$ and a convex cone C in Y, one obtains ordinary programs with operator constraints.

Here, a perturbation approach to optimality conditions and conjugate duality is given using the perturbation function

$$\sigma(y) := \inf \Big\{ f(x) : x \in X, y \in \Gamma(x) \Big\},\ y \in Y\,.$$

Recall that f is qv (quasiconvex), iff the strict level sets $\Big\{ x : f(x) < \alpha \Big\}$ are convex for all $\alpha \in \mathbb{R}$. Here a multimap $\Gamma : X \rightrightarrows Y$ is called qv (quasiconvex), iff

$$\lambda\Gamma(x_1) + (1-\lambda)\Gamma(x_2) \subset \bigcup \Big\{ \Gamma(\xi) : \xi \in [x_1, x_2] \Big\}$$

for all $x_1, x_2 \in X; \lambda \in [0,1]$. This definition not only includes convex multimaps (i.e. multimaps with a convex graph), but is also the "right" extension, since the perturbation function σ is shown to be qv for f qv and Γ qv.
Then under convexity and interior point conditions with respect to the convex core topology, the separation theorem is applied to achieve the extension of the classical optimality condition of Luenberger in finite dimensional quasiconvex programming to general programs of the form (P): There exists a multiplier $l \in Y^*$ such that

$$\text{(OC)} \qquad \sigma(0) \le \sigma(y) \text{ for all } y \in Y \text{ satisfying } l(y) \le 0\,.$$

The required assumptions are in particular satisfied for qv f with $f|_{\mathrm{dom}\ f}$ upper semicontinuous along lines and qv Γ with int $\Gamma(x) \neq \emptyset$ under a Slater type condition ($\exists x^0 \in X$ such that $0 \in \operatorname{int} \Gamma(x^0)$). Note that (OC) relaxes the Kuhn-Tucker condition for (P)

$$\sigma(0) \le \sigma(y) + l(y),\ \forall\, x \in \mathrm{dom}\ f \cap \mathrm{dom}\ \Gamma, y \in \Gamma(x)\,.$$

The optimality condition (OC) motivates to introduce the dual functional $\sigma^* : Y^* \to \bar{\mathbb{R}}$ by

$$\begin{aligned} \sigma^*(y^*) &:= \inf \Big\{ \sigma(y) : y \in Y, \langle y^*, y \rangle \le 0 \Big\} \\ &= \inf \Big\{ f(x) : x \in X, y \in \Gamma(x), \langle y^*, y \rangle \le 0 \Big\}\,. \end{aligned}$$

Indeed, [26] proves the duality result that (OC) holds, iff $\sigma(0) = \sigma^*(l) = \max \Big\{ \sigma^*(y^*) : y^* \in Y^* \Big\}$.
Clearly, σ^* is positively homogeneous (of degree 0), and moreover shown to be quasiconcave.
In a further step towards biduality, Y^* is endowed with some vector space topology that is finer than $\sigma(Y^*, Y)$ and the bidual function $\sigma^{**} : Y^{**} = \mathcal{L}(Y^*, \mathbb{R}) \to \bar{\mathbb{R}}$ is defined by

$$\sigma^{**}(y^{**}) := \sup \Big\{ \sigma^*(y^*) : y^* \in Y^*, \langle y^{**}, y^* \rangle \le 0 \Big\}\,.$$

From above, σ^{**} is qv and positively homogeneous of degree 0.
With $Y \subset Y^{**}$, one is led to compare σ and σ^{**} on Y and [26] proves the biduality result that $\sigma^{**}(y) \le \sigma(y),\ \forall\, y \in Y$; moreover, if (OC) holds, then

$\sigma^{**}(0) = \sigma(0)$. This situation can be interpreted as follows: $\sigma^{**}|_Y$ is a qv and positively homogeneous (of zero degree) support functional to σ at 0. Thus, $\sigma^{**}|_Y$ can be considered as a qv "subgradient" of σ.
Finally again employing the separation theorem, now assuming that the qv function $\sigma : Y \to \bar{\mathbb{R}}$ is lower semicontinuous at $y = 0$ in the locally convex space Y, [26] establishes that $\sigma^{**}(0) = \sigma(0)$.

Possible Extensions and Some Relations to More Recent Work

At the first sight, some concepts and results above (in particular definition and quasiconcavity of σ^*) extend to more general coupling functions φ on $Y^* \times Y$ instead of the natural coupling $\langle y^*, y\rangle$. This gives a connection to the modern theory of Abstract Convexity in the treatises of Rubinov [30] and of Singer [33].
However, as already the definition of a qv multimap indicates, [26] works with the *inner definition* of convexity based on the notion of convex combination. On the contrary, the *outer definition* of convexity based on the separation property leads to the modern abstract convexity theory of Rubinov ([30]). This might be the key for the extension of some of the hard results (proved in [26] by the separation theorem) to the more general general setting of abstract convexity with the potential of further applications to global optimization.
Another direction of extension is to consider perturbations in the constraints by subsets of a given set Y instead by elements of Y. Thus instead of (P) one can study perturbed programs and associated perturbation functions

$$\sigma_{\cap}(B) := \inf\left\{f(x) : x \in X, B \cap \Gamma(x) \neq \emptyset\right\},\ B \in \mathcal{B};$$

$$\sigma_{\subset}(B) := \inf\left\{f(x) : x \in X, B \subset \Gamma(x)\right\},\ B \in \mathcal{B}$$

with $\mathcal{B}$ a given system of subsets of Y. Clearly, the constraint $B \cap \Gamma(x) \neq \emptyset$ rewrites $0_Y \in \tilde{\Gamma}(x)$, the constraint of (P), with the modified multimap $\tilde{\Gamma}(x) = \Gamma(x) - B$. This reformulation, however, hides the dependence of the perturbed program and its value on the perturbation B.
Instead of $\mathcal{B} = \{\{y\}, y \in Y\}$ one can consider $\mathcal{B} = 2^Y$, the set of subsets of Y, or more generally a complete lattice $\mathcal{B} \subset 2^Y$ ordered by containment. The introduction of a complete lattice $\mathcal{B}$ gives a connection to the concepts of abstract (quasi)convexity in [33].

More recently Penot and Volle [29] studied the more conventional quasiconvex program of minimizing a quasiconvex function on a convex subset in a Banach space under convex operator constraints. They were able to relax the Slater condition to the now standard constraint qualification of convex optimization and obtain a "surrogate" multiplier, i.e. a multiplier that satisfies (OC).

4.2 Axiomatic Approach to Optimality and Conjugate Duality in Global Optimization

In [28] the authors study conjugate duality with arbitrary coupling functions. Their only tool is a certain support property, which is automatically fulfilled in the two most widely used special cases, namely the case where the underlying space is a topological vector space and the coupling functions are the continuous linear ones, and the case where the underlying space is a metric space and the coupling functions are the continuous ones. They obtain thereby a simultaneous axiomatic extension of these two classical models. Also included is a condition for global optimality, which requires only the mentioned support property.

The axiomatic approach to optimality conditions in global optimization in [28] is extended and deepened in [3]. There nonconvex vector minimization problems $\min\Big\{g(x) - h(x) : x \in X\Big\}$ are considered, where g and h are functions defined on an arbitrary set X and taking values in an topological space Z, ordered by some convex cone P with nonempty interior. To mimic the standard case $Z = \mathbb{R}, P = \mathbb{R}_+$, $(Z, \geq_P)$ is assumed to be order-complete (i.e. every nonempty subset of Z that has an upper bound, also has a supremum) and two artificial elements $+\infty$ and $-\infty$ are adjoined to Z.
Further g and h are allowed to be proper extended-valued, i.e. $g, h : X \to Z \cup \{+\infty\}$ with dom $g = \Big\{x \in X : g(x) \in Z\Big\} \neq \emptyset$, dom $h \neq \emptyset$. To extend the classical subdifferential in convex analysis, a nonempty family $\varPhi$ of functions from X to Z are fixed. Then following earlier work by Flores-Bazán and Martinez-Legaz, for any $\bar{x} \in$ dom g and $\varepsilon \in Z$ with $\varepsilon \geq 0$,

$$\partial_\varepsilon g(\bar{x}) = \Big\{\varphi \in \varPhi : g(x) - g(\bar{x}) \geq \varphi(x) - \varphi(\bar{x}) - \varepsilon,\ \forall\, x \in X\Big\}.$$

Since this definition makes sense only when $\bar{x} \in$ dom g, minorants besides subgradients are introduced by

$$\mu_\alpha f(\bar{x}) = \Big\{\varphi \in \varPhi : g(x) \geq \alpha + \varphi(x) - \varphi(\bar{x}),\ \forall\, x \in X\Big\}$$

for any $\bar{x} \in X$ and $\alpha \in Z$ with $\alpha \leq g(\bar{x})$.
In this order-theoretic setting, the authors put the following assumptions into play:

$$\text{(H1)} \quad \forall\, \bar{x} \in X, \quad \alpha \in Z \wedge h(\bar{x}) - \alpha \in \text{ int } P \Rightarrow \mu_\alpha h(\bar{x}) \neq \emptyset\,,$$
$$\text{(H2)} \quad \forall\, \bar{x} \in X, \quad \alpha \in Z \wedge h(\bar{x}) - \alpha \in P \Rightarrow \mu_\alpha h(\bar{x}) \neq \emptyset\,.$$

These hypotheses are verified in various examples. Namely, with $Z = \mathbb{R}$, convex analysis guarantees (H1), respectively (H2) for convex proper functions h with $\varPhi = X^*$, provided h is lower semicontinuous, respectively continuous. Classic analysis provides the stronger condition (H2) for proper lower

semicontinuous functions on a metric space with Φ given as the family of all continuous functions on X. These instances are extended to general ordered spaces in further examples.
Thus unifying previous work and proceeding to a new level of abstraction the authors arrive at the following global optimality condition:
Assume (H1), let $x^\circ \in \text{dom}\, g \cap \text{dom}\, h$. Then x° solves the above minimization problem in the sense that

$$g(x) - g(x^\circ) \geq_P h(x) - h(x^\circ),\ \forall\, x \in X,$$

if and only if

$$(*) \qquad \partial_\varepsilon h(x^\circ) \subset \partial_\varepsilon g(x^\circ),\ \forall\, \varepsilon \geq_P 0\,.$$

If (H1) is replaced by (H2) and Φ satisfies an extra condition, then the optimality condition $(*)$ can be simplified to

$$(**) \qquad \partial_0 h(x^\circ) \subset \partial_0 g(x^\circ).$$

Furthermore the authors are able to reformulate the optimality conditions $(*)$ and $(**)$ as duality results of Singer-Toland type involving Φ-conjugate functions (compare also with [33], especially chapters 8, 10).
The paper under review gives an essential contribution to the axiomatic optimality and duality theory in nonconvex vector optimization. It clearly demonstrates the role that lattice theory and sandwich theorems for vector functions [2] play in this field.

References

1. Benoist J, Borwein JM, Popovici N (2003) A characterization of quasiconvex vector-valued functions. Proc. Am. Math. Soc. 131: 1109–1113
2. Borwein JM, Théra M(1992) Sandwich theorems for semicontinuous operators. Can. Math. Bull. 35: 463–474
3. Flores-Bazán F, Oettli W (2001) Simplified optimality conditions for minimizing the difference of vector-valued functions. J. Optimization Theory Appl. 108: 571–586
4. Glover BM, Jeyakumar V, Oettli W (1994) Solvability theorems for classes of difference convex functions. Nonlinear Anal., Theory Methods Appl. 22: 1191–1200
5. Glover BM, Jeyakumar V, Oettli W (1994) A Farkas lemma for difference sublinear systems and quasidifferentiable programming. Math. Program., Ser. A 63: 109–125
6. Gorokhovik VV (1986) ϵ-quasidifferentiability of real-valued functions and optimality conditions in extremal problems. Math. Program. Study 29, 203–218
7. Horst R, Tuy Hoang (1996) Global optimization. Deterministic approaches.3rd rev. a. enl. ed. Springer, Berlin
8. Jeyakumar V, Glover BM (1995) Nonlinear extensions of Farkas' lemma with applications to global optimization and least squares. Math. Oper. Res. 20: 818–837

9. Jeyakumar V, Glover BM (1996) Characterizing global optimality for DC optimization problems under convex inequality constraints. J. Glob. Optim. 8: 171–187
10. Jeyakumar V, Rubinov AM, Glover BM, Ishizuka Y (1996) Inequality systems and global optimization. J. Math. Anal. Appl. 202: 900–919
11. Jeyakumar V, Yen ND (2004) Solution stability of nonsmooth continuous systems with applications to cone-constrained optimization. SIAM J. Optim. 14: 1106–1127
12. Jeyakumar V, Oettli W, Natividad M (1993) A solvability theorem for a class of quasiconvex mappings with applications to optimization. J. Math. Anal. Appl. 179: 537–546
13. Konno H, Kuno T (1992) Linear multiplicative programming. Math. Program., Ser. A 56: 51–64
14. Kuno T (2001) A finite branch-and-bound algorithm for linear multiplicative programming. Comput. Optim. Appl. 20: 119–135
15. Lemaire B (1998) Duality in reverse convex optimization. SIAM J. Optim. 8: 1029–1037
16. Le Thi HA, Pham Dinh T, Muu LD (2003) Simplicially-constrained DC optimization over efficient and weakly efficient sets. J. Optimization Theory Appl. 117: 503–531
17. Liu XJ, Umegaki T, Yamamoto Y (1999) Heuristic methods for linear multiplicative programming. J. Glob. Optim. 15: 433–447
18. Luc DT (1988) Theory of vector optimization, Lect. Notes Econ. Math. Sci. 319. Springer, Berlin
19. Margolis B (1990) Quasi-differentiability: A formal approach. First and second gradient classes. Optimization 21: 649–667
20. Matsui T(1996) NP-hardness of linear multiplicative programming and related problems. J. Glob. Optim. 9: 113–119
21. Meister B, Oettli W (1967) On the capacity of a discrete, constant channel Inf. Control 11: 341-351
22. Muu LD, Oettli W (1991) An algorithm for indefinite quadratic programming with convex constraints. Oper. Res. Lett. 10: 323–327
23. Muu LD, Oettli W (1991) Method for minimizing a convex-concave function over a convex set. J. Optimization Theory Appl. 70: 377–384
24. Muu LD, Oettli W (1993) Combined branch-and-bound and cutting plane methods for solving a class of nonlinear programming problems. J. Glob. Optim. 3: 377–391
25. Muu LD, Oettli W (2001) Optimization over equilibrium sets. Optimization 49: 179–189
26. Oettli W (1981) Optimality conditions involving generalized convex mappings. In: Schaible S, Ziemba WT (eds) Generalized Concavity in Optimization and Economics 227–238
27. Oettli W (1989) Decomposition schemes for finding saddle points of quasi-convex-concave functions. In: Kall P et al. (eds) Quantitative Methoden in den Wirtschaftswissenschaften 31–51
28. Oettli W, Schläger D (1998) Conjugate functions for convex and nonconvex duality. J. Glob. Optim. 13: 337–347
29. Penot J-P, Volle M (2004) Surrogate programming and multipliers in quasi-convex programming. SIAM J. Control Optimization 42: 1994–2003

30. Rubinov A (2000) Abstract convexity and global optimization. Kluwer Academic Publishers, Dordrecht
31. Schaible S, Sodini C (1995) Finite algorithm for generalized linear multiplicative programming. J. Optimization Theory Appl. 87: 441–455
32. Scott CH, Jefferson Th R (2005) Duality for linear multiplicative programs. ANZIAM J. 46: 393–397
33. Singer I (1997) Abstract convex analysis. John Wiley & Sons, New York
34. Thach PT (1991) Quasiconjugates of functions, duality relationship between quasiconvex minimization under a reverse convex constraint and quasiconvex maximization under a convex constraint, and applications. J. Math. Anal. Appl. 159: 299–322
35. Thach PT; Burkard RE, Oettli W (1991) Mathematical programs with a two-dimensional reverse convex constraint. J. Glob. Optim. 1: 145–154

Local and Global Consumer Preferences

Reinhard John*

Department of Economics, University of Bonn, Adenauerallee 24-42, 53113 Bonn, Germany rjohn@uni-bonn.de

Summary. Several kinds of continuous (generalized) monotone maps are characterized by partial gradient maps of skew-symmetric real-valued bifunctions displaying corresponding (generalized) concavity-convexity properties. As an economic application, it is shown that two basic approaches explaining consumer choice are behaviorally equivalent.

Key words: Consumer preference, equilibrium problem, generalized concave-convex bifunction, generalized monotonicity, variational inequality.

1 Introduction

In order to explain consumer behavior, two basic models coexist in the economic literature: the local and the global theory [10].
The global theory, described in more detail in Section 2, is the standard approach which assumes that a consumer is able to rank any two alternative commodity bundles in a convex consumption set $X \subseteq \mathbb{R}^n$. By contrast with the usual textbook version, we do not suppose this ranking to be transitive. According to Shafer[9], it is formalized by a continuous and complete binary relation R on X that can be numerically represented by a continuous and skew-symmetric real-valued bifunction r on X, i.e. xRy if and only if $r(x, y) \geq 0$ where this is interpreted as "x is at least as good as y". We call R a global preference on X.
Assume now that the consumer faces a nonempty set of feasible alternatives $Y \subseteq X$ from which he is allowed to choose. Then, the interpretation of R implies a choice of an $x \in Y$ such that x is as least as good as any other

*I would like to thank Nicolas Hadjisavvas and Juan Enrique Martínez-Legaz for helpful discussions. In particular, I am grateful to N. Hadjisavvas for a valuable suggestion leading to an improvement over an earlier version.

alternative $y \in Y$. In terms of the numerical representation r of R, the choice set assigned to Y in the global theory is given by

$$C^r(Y) = \{x \in Y | \forall y \in Y : r(x, y) \geq 0\},$$

i.e. $C^r(Y)$ is the solution set of the equilibrium problem $\mathrm{EP}(r, Y)$ defined by r on Y.

The local theory has been introduced by Allen[1] and further developed by Georgescu-Roegen[4],[5] and Katzner[6]. By contrast with the global approach, a local preference only requires that the consumer is able to rank alternatives in a small neighborhood of a given bundle relative to that bundle. As will be explained in Section 3, this idea can be represented by a continuous function $g : X \to \mathbb{R}^n$ such that y in a neighborhood of x is interpreted to be better than x if and only if $g(x)^T(y - x) > 0$. For $Y \subseteq X$ the choice set assigned to Y in the local theory is then characterized by

$$C^g(Y) = \{x \in Y | \forall y \in Y : g(x)^T(y - x) \leq 0\},$$

i.e. $C^g(Y)$ is the solution set of the Stampacchia variational inequality problem $\mathrm{VI}(g, Y)$ defined by g on Y.[2] Furthermore, as argued by Allen and Georgescu-Roegen, stable choices require g to be at least pseudomonotone.

Conceptually, the local approach seems to be weaker than the global one. However, as shown in Section 5 (Theorem 2), both are *behaviorally equivalent* if the global preferences are represented by a class of bifunctions which are defined as *diagonally pseudo-concave-convex* in Section 4. It means that any such global preference representation r yields a local preference representation g_r such that $C^{g_r}(Y) = C^r(Y)$ for every convex $Y \subseteq X$ and that, conversely, for any given local preference defined by g there is a diagonally pseudo-concave-convex global preference representation r^g such that $g_{r^g} = g$. Put differently, for every convex $Y \subseteq X$ the solutions to $\mathrm{EP}(r, Y)$ and $\mathrm{VI}(g_r, Y)$ coincide as well as those to $\mathrm{VI}(g, Y)$ and $\mathrm{EP}(r^g, Y)$. Since g_r is obtained as the partial gradient map given by $g_r(x) = \nabla_1 r(x, x)$, this result confirms that "VI can be viewed as the differential form of EP" (I. Konnov in [7], p. 560).

Taking this view for granted, it is natural to extend the well known relationship between (generalized) convexity properties of real-valued functions and cyclic (generalized) monotonicity properties of their gradients to the case of non-gradient maps. In Section 4 several kinds of diagonally (generalized)-concave-convex skew-symmetric bifunctions are introduced (Definition 1). It is shown that their partial gradient mappings characterize continuous maps with corresponding (generalized) monotonicity properties (Theorem 1). From a purely mathematical viewpoint, this section contains the main results. They provide the essential steps for the intended application to consumer theory in Section 5.

[2] Notice the reverse inequality compared to the usual definition. This convention is employed throughout the paper.

2 Global Preferences

Consider an economy with a finite number n of commodities. A vector $x = (x_i)_{i=1}^n \in \mathbb{R}^n$ is called a consumption bundle and interpreted as the consumption of x_i units of commodity i for $i = 1, ..., n$ where positive components represent inputs (e.g.consumption goods) and negative components represent outputs (e.g.labor services). In general, not all consumption bundles are physically possible. We assume that the feasible ones are given by an open convex subset X of $\mathbb{R}^n$, the consumer's *consumption set*. For example, $X = \mathbb{R}^n_{++}$ in case that only consumption goods are considered.

A consumer's taste is described by a binary relation R on X where xRy is interpreted as "x is at least as good as y" or as "x is (weakly) preferred to y". x is called strictly preferred (resp. indifferent) to y if xRy but not yRx (resp. xRy and yRx). A basic assumption in the global theory is the *completeness* of R, i.e.

$$xRy \vee yRx \quad \text{for all } x, y \in X.$$

Equivalently, any two consumption bundles x and y can be ranked in the sense that either x is strictly preferred to y or y is strictly preferred to x or x and y are indifferent.

In contrast to the standard textbook model (see e.g. [8]) we do not assume that R is transitive, i.e. that R is a total preorder. However, as usual in that model, we always require R to be *continuous*, i.e. R is a closed subset of $X \times X$, and call a complete and continuous relation R a *global preference* on X.

In case that R is transitive, it is well known (see e.g. [3]), that R can be represented by a continuous *utility function* $u : X \to \mathbb{R}$ in the sense that for all $x, y \in X$

$$xRy \Leftrightarrow u(x) \geq u(y).$$

In the general case, Shafer[9] has shown that R has a continuous *numerical representation*, i.e. a real-valued function r defined on $X \times X$ such that for all $x, y \in X$

$$\begin{aligned} xRy &\Leftrightarrow r(x, y) \geq 0 \\ r(x, y) &= -r(y, x) . \end{aligned}$$

Observe that a utility function yields such a representation by defining $r(x, y) = u(x) - u(y)$.

Conversely, it is obvious that any continuous skew-symmetric real-valued bifunction r on X defines a global preference that is represented by r and, therefore, is called a *global preference representation.*

Let us now describe a consumer's behavior. If the consumer is allowed to make a choice between the alternatives in a nonempty subset Y of X, it is natural to assume that he will choose a bundle $x \in Y$ such that xRy for all $y \in Y$. If R is represented by r this is obviously equivalent to $r(x, y) \geq 0$ for all $y \in Y$. Put differently, the choice set

$$C^r(Y) = \{x \in Y \,|\, \forall y \in Y : r(x, y) \geq 0\}$$

consists of all solutions to the equilibrium problem EP(r, Y) defined by the representing bifunction r on Y.

It is clear that, in general, the existence of such a bundle is not guaranteed. In order that the concept of a preference can be considered as an operational definition of a consumer, it should meet at least the requirement that the choice set is nonempty for all nonempty, compact, and convex subsets Y of X, in particular, for all compact budget sets.

In the transitive case, this property is implied by continuity. In general, an additional assumption is needed. As it is well known, a sufficient condition is that the representation r is quasiconvex in the second variable or, equivalently, quasiconcave in the first variable (see e.g. [7], Theorem 13.1). At the same time, this condition also ensures the convexity of $C^r(Y)$ which is another indispensable property for developing a meaningful equilibrium theory. For that reason, the utility function u in the standard model is usually also assumed to be quasiconcave.

Most textbook presentations actually suppose that u is differentiable and pseudoconcave in order to derive the well known first order characterization

$$x \in C^r(Y) \Leftrightarrow \forall y \in Y : \nabla u(x)^T (y - x) \leq 0.$$

In the general case with a differentiable and pseudoconcave-pseudoconvex representation r one obtains the analogous characterization (see [7], Theorem 13.10)

$$\begin{aligned} x \in C^r(Y) &\Leftrightarrow \forall y \in Y : \nabla_1 r(x, x)^T (y - x) \leq 0 \\ &\Leftrightarrow \forall y \in Y : \nabla_2 r(x, x)^T (y - x) \geq 0. \end{aligned}$$

These results replace the global viewpoint embodied in the definition of $C^r(Y)$ by a local one in the sense that knowledge of R is only required near x. As will be shown in Section 5, they even hold under weaker conditions on r. Furthermore, they suggest to base a theory of choice directly on a local preference concept to which we turn in the following section.

3 Local Preferences

Consider a consumer with a consumption set X as in the previous section who faces alternative consumption bundles in a neighborhood of some given bundle $x \in X$. Assume that the consumer is able to distinguish three kinds of possible directions in which he can move away from x according to his taste: Preference, nonpreference, and indifference. Allen[1] has formalized this idea by postulating the existence of a vector $g(x) \in \mathbb{R}^n$ such that a direction $v \in \mathbb{R}^n$ is a preference (resp. nonpreference, resp. indifference) direction if and only if

$$g(x)^T v > 0 \text{ (resp. } < 0, \text{resp. } = 0).$$

Observe that multiplying $g(x)$ by some positive number does not change the local preference, i.e. provided that $g(x) \neq 0$ for all $x \in X$ we could normalize g by requiring $||g(x)|| = 1$ for all $x \in X$. However, in order to be as general as possible, we do not exclude the case $g(x) = 0$ which is interpreted as a local satiation point.

According to Allen[1] and Georgescu-Roegen[4],[5], a consumer's choice is described as follows. Assume that x is contained in some given set $Y \subseteq X$ that represents the consumer's feasible bundles. They have called x an "equilibrium position" relative to Y if no direction away from x to any other alternative y in Y is one of preference, i.e. if

$$g(x)^T(y - x) \leq 0 \quad \text{for all } y \in Y.$$

Thus, x is an equilibrium consumption bundle iff x is a solution to the Stampacchia variational inequality problem VI(g, Y) defined by g on Y.

Allen also required the equilibrium to be stable. He added an assumption on g that was made more precise by Georgescu-Roegen who called it the "principle of persisting nonpreference":

If the consumer moves away from an arbitrary bundle x to a bundle $x + \Delta x$ such that Δx is not a preference direction, then Δx is a nonpreference direction at $x + \Delta x$. Formally stated, this principle says

$$g(x)^T \Delta x \leq 0 \quad \text{implies} \quad g(x + \Delta x)^T \Delta x < 0,$$

or, by denoting $x + \Delta x = y$,

$$g(x)^T(y - x) \leq 0 \quad \text{implies} \quad g(y)^T(y - x) < 0.$$

Clearly, this property is now called strict pseudomonotonicity of g, where, in contrast to most of the literature, we use this notion in the sense of generalizing a *decreasing* real valued function of one real variable.

Later, Georgescu-Roegen[5] generalized his principle of persisting nonpreference by only requiring g to be pseudomonotone, i.e.

$$g(x)^T(y - x) \leq 0 \quad \text{implies} \quad g(y)^T(y - x) \leq 0.$$

Not surprisingly, he also assumed continuity of g. Thus, we call a continuous and pseudomonotone mapping $g : X \to \mathbb{R}^n$ a *local preference representation* on X.

It is important to notice that the local approach yields a satisfactory choice behavior. Indeed, continuity and pseudomonotonicity of g imply that for any nonempty, convex, and compact subset Y of X the choice set in the local theory

$$C^g(Y) = \{x \in Y \,|\, \forall y \in Y : g(x)^T(y - x) \leq 0\}$$

is nonempty, convex, and compact (see e.g. [7], Theorem 13.6).

4 Main Results

In the sequel, X denotes a nonempty, open, and convex subset of $\mathbb{R}^n$.
Let $r : X \times X \to \mathbb{R}$ be a bifunction on X and define for arbitrary $x \in X$ and $h \in \mathbb{R}^n$ the single variable real valued functions $r_1(x,h)$ and $r_2(x,h)$ on the set $I_{x,h} = \{t \in \mathbb{R} |\, x + th \in X\}$ by

$$r_1(x,h)(t) = r(x+th, x) \quad \text{and} \quad r_2(x,h)(t) = r(x, x+th).$$

We call r *diagonally differentiable* if for all x, h and t the derivatives $r_1(x,h)'(t)$ and $r_2(x,h)'(t)$ exist and are continuous in x, h, t and if the partial gradients $\nabla_1 r(x,x)$ and $\nabla_2 r(x,x)$ exist for all $x \in X$, i.e. for all x, h

$$r_1(x,h)'(0) = \nabla_1 r(x,x)^T h \quad \text{and} \quad r_2(x,h)'(0) = \nabla_2 r(x,x)^T h.$$

A basic derivation of such a bifunction from a continuous mapping $g : X \to \mathbb{R}^n$ is given in

Proposition 1. *Let $g : X \to \mathbb{R}^n$ be continuous. Then the bifunction r^g : $X \times X \to \mathbb{R}$ defined by*

$$r^g(y,x) = \int_0^1 g\,(x + s(y-x))^T (y-x)\, ds$$

is skew-symmetric, continuous, and diagonally differentiable.[3] *More precisely, for every $x \in X$ and every h*

$$r_1^g(x,h)'(t) = g(x+th)^T h = -r_2^g(x,h)'(t)$$

and, in particular,

$$\nabla_1 r^g(x,x) = g(x) = -\nabla_2 r^g(x,x).$$

Proof. It is easy to see that r^g is skew-symmetric and continuous. Moreover, for x, h and t such that $x + th \in X$ we obtain

$$r_1^g(x,h)(t) = r^g(x+th, x) =$$
$$\int_0^1 g(x+sth)^T th\, ds = \int_0^t g(x+sh)^T h\, ds.$$

Since the derivative of the definite integral with respect to the upper limit of integration is equal to the value of the integrand at that limit, $g(x+th)^T h = r_1^g(x,h)'(t) = -r_2^g(x,h)'(t)$ which is continuous in x, h, t. By definition of the partial gradients, $\nabla_1 r^g(x,x) = g(x) = -\nabla_2 r^g(x,x)$. □

[3]This function was suggested by Nicolas Hadjisavvas.

The next proposition shows that (generalized) monotonicity properties of g imply (generalized) concavity-convexity properties of the bifunction r^g which are introduced in the following definition. We emphasize that all notions of (generalized) monotonicity are defined in the sense of generalizing a *nonincreasing* real valued function of one real variable.

Definition 1. *A skew-symmetric and diagonally differentiable bifunction r on X is called*

(a) *diagonally quasi-concave-convex, if for every $x \in X$ and $h \neq 0$ the function $r_1(x,h)$ is quasiconcave,*
(b) *diagonally (strictly) pseudo-concave-convex, if for every $x \in X$ and $h \neq 0$ the function $r_1(x,h)$ is (strictly) pseudoconcave,*
(c) *diagonally (strictly) concave-convex, if for every $x \in X$ and $h \neq 0$ the function $r_1(x,h)$ is (strictly) concave.*

Proposition 2. *Let $g : X \to \mathbb{R}^n$ be a continuous function.*

(a) *If g is quasimonotone, then r^g is diagonally quasi-concave-convex.*
(b) *If g is (strictly) pseudomonotone, then r^g is diagonally (strictly) pseudo-concave-convex.*
(c) *If g is (strictly) monotone, then r^g is diagonally (strictly) concave-convex.*

Proof. It is well known (see e.g. [2]) that all mentioned (generalized) monotonicity properties of g are characterized by the corresponding property of the single variable functions $g_{x,h} : I_{x,h} \to \mathbb{R}$, defined by $g_{x,h}(t) = g(x+th)^T h$. By Proposition 1, $g_{x,h}$ is the derivative of $r_1^g(x,h)$ which implies the corresponding (generalized) concavity property of $r_1^g(x,h)$ (see e.g. Proposition 2.5 in [2]). □

Conversely, we shall show the (generalized) monotonicity of the partial gradient map $x \mapsto \nabla_1 r(x,x)$ for a skew-symmetric and diagonally (generalized) concave-convex bifunction r. It turns out that this is already implied by weaker (generalized) concavity-convexity conditions on r. These are, in a sense, local versions of the well known characterizations for differentiable functions and introduced in

Definition 2. *Let r be a skew-symmetric bifunction on X. If the partial function $r(\cdot,x)$ is differentiable at $x \in X$ then r is called*

(i) *quasi-concave-convex at x, if for all $y \in X$*

$$\nabla_1 r(x,x)^T(y-x) < 0 \Rightarrow r(y,x) < 0, \tag{1}$$

(ii) *pseudo-concave-convex at x, if, in addition to (1), for all $y \in X$*

$$\nabla_1 r(x,x)^T(y-x) \leq 0 \Rightarrow r(y,x) \leq 0, \tag{2}$$

(iii) *strictly pseudo-concave-convex at x, if for all $y \in X, y \neq x$*

$$\nabla_1 r(x,x)^T(y-x) \leq 0 \Rightarrow r(y,x) < 0, \tag{3}$$

(iv) *(strictly) concave-convex at* x*, if for all* $y \in X, y \neq x$

$$r(y,x) \leq \ (<) \, \nabla_1 \, r(x,x)^T (y-x). \tag{4}$$

Observe that (ii) in Definition 2 also requires (1) to be satisfied. Actually, it is not difficult to show that in general (1) does not follow from (2). On the other hand, (2) trivially implies

$$\nabla_1 \, r(x,x)^T (y-x) < 0 \Rightarrow r(y,x) \leq 0 \tag{5}$$

which means that in general (5) is weaker than (1).
However, in the special case where $r(y,x) = f(y) - f(x)$ such that (2) describes pseudoconcavity of f, (1) and (5) are equivalent and characterize quasiconcavity of f (see e.g. Proposition 2.1 in [2]). For our purpose, the stronger condition (1) is the appropriate one as shown by the next two results. On the one hand, (1) turns out to be necessary for a diagonally quasi-concave-convex r, on the other hand, it is needed in the proof of quasimonotonicity of the partial gradient mapping in Proposition 4.

Proposition 3. *Let* r *be a skew-symmetric and diagonally differentiable bifunction on* X*. Then the partial gradients* $\nabla_1 \, r(x,x)$ *are continuous in* x *and*

(a) *If* r *is diagonally quasi-concave-convex, then* r *is quasi-concave-convex at every* $x \in X$*.*
(b) *If* r *is diagonally (strictly) pseudo-concave-convex, then* r *is (strictly) pseudo-concave-convex at every* $x \in X$*.*
(c) *If* r *is diagonally (strictly) concave-convex, then* r *is (strictly) concave-convex at every* $x \in X$*.*

Proof. The partial gradients are continuous by the definition of diagonal differentiability. We only prove (a) and (c). The statement (b) is shown analogously.
(a): By Definition 1, r is diagonally quasi-concave-convex if for every $x \in X$ and every $h \neq 0$ the function $r_1(x,h)$ is quasiconcave, i.e. for all $t_1, t_2 \in I_{x,h}$ the inequality $r_1(x,h)'(t_1)(t_2 - t_1) < 0$ implies that $r_1(x,h)(t_2) < r_1(x,h)(t_1)$. By setting $h = y - x$ for $y \in X, y \neq x$ and $t_1 = 0, t_2 = 1$ we obtain $r_1(x,h)'(t_1)(t_2 - t_1) = \nabla_1 \, r(x,x)^T (y-x), r_1(x,h)(t_2) = r(y,x)$ and $r_1(x,h)(t_1) = r(x,x) = 0$ which yields (1).
(c): r is diagonally (strictly) concave-convex if for every $x \in X$ and every $h \neq 0$ the function $r_1(x,h)$ is (strictly) concave, i.e. for all $t_1, t_2 \in I_{x,h}$ such that $t_1 \neq t_2$ the inequality $r_1(x,h)(t_2) - r_1(x,h)(t_1) \leq (<) \, r_1(x,h)'(t_1)(t_2 - t_1)$ holds. Setting h, t_1, t_2 as before yields (4). □

Proposition 4. *If the bifunction* r *on* X *is quasi-concave-convex (resp. (strictly) pseudo-concave-convex, resp. (strictly) concave-convex) at every* $x \in X$ *then the gradient mapping* g_r *defined by*

$$g_r(x) = \nabla_1 \, r(x,x)$$

is quasimonotone (resp. (strictly) pseudomonotone, resp. (strictly) monotone).

Proof. If r is quasi-concave-convex at every x, $\nabla_1 r(x,x)^T(y-x) < 0$ implies $r(y,x) \leq 0$ by (1). From skew-symmetry of r it follows that $r(x,y) \geq 0$ and, again by (1), that $\nabla_1 r(y,y)^T(x-y) \geq 0$ or, equivalently, $\nabla_1 r(y,y)^T(y-x) \leq 0$. Hence, g_r is quasimonotone.
Assume now that r is pseudo-concave-convex at every x. If $\nabla_1 r(x,x)^T(y-x) \leq 0$ then, by (2), $r(y,x) \leq 0$ or, equivalently, $r(x,y) \geq 0$. Hence, (1) implies $\nabla_1 r(y,y)^T(x-y) \geq 0$, i.e. $\nabla_1 r(y,y)^T(y-x) \leq 0$. Thus, we have proved that g_r is pseudomonotone.
In the strict case, by (3), $\nabla_1 r(x,x)^T(y-x) \leq 0$ and $x \neq y$ imply $r(y,x) < 0$ or, equivalently, $r(x,y) > 0$. It follows again from (3) that $\nabla_1 r(y,y)^T(x-y) > 0$ or $\nabla_1 r(y,y)^T(y-x) < 0$. Hence, g_r is strictly pseudomonotone.
Finally, assume that r is (strictly) concave-convex at every x. (4) implies the inequalities

$$r(y,x) \leq (<)\ \nabla_1 r(x,x)^T(y-x)$$
$$r(x,y) \leq (<)\ \nabla_1 r(y,y)^T(x-y).$$

Adding these inequalities and using skew-symmetry yields

$$0 \leq (<)\, [\nabla_1 r(x,x) - \nabla_1 r(y,y)]^T(y-x)$$

or, equivalently,

$$[\nabla_1 r(x,x) - \nabla_1 r(y,y)]^T(x-y) \leq (<)\, 0.$$

Thus, g_r is (strictly) monotone. □

The previous results can be summarized by the following

Theorem 1. *Let $g : X \to \mathbb{R}^n$ be defined on an open and convex subset X of $\mathbb{R}^n$. Then the following statements are equivalent:*

(i) *g is continuous and quasimonotone (resp. (strictly) pseudomonotone, resp. (strictly) monotone).*
(ii) *There is a skew-symmetric, continuous, and diagonally quasi-concave-convex (resp. diagonally (strictly) pseudo-concave-convex, resp. diagonally (strictly) concave-convex) bifunction $r : X \times X \to \mathbb{R}$ such that for every $x \in X$*
$$\nabla_1 r(x,x) = g(x).$$
(iii) *There is a skew-symmetric and continuous bifunction $r : X \times X \to \mathbb{R}$ which is quasi-concave-convex (resp. (strictly) pseudo-concave-convex, resp. (strictly) concave-convex) at every $x \in X$ such that $\nabla_1 r(x,x) = g(x)$ is continuous in x.*

Proof. By Propositions 1 and 2, (i) implies (ii) by choosing $r = r^g$. (iii) follows from (ii) by Proposition 3. Finally, by Proposition 4, (iii) implies (i). □

Remark: In an earlier version of this paper the implication (i) $\Rightarrow$ (iii) was obtained more easily by assigning to g the bifunction r_g defined by

$$r_g(y,x) = \frac{1}{2}(g(y)+g(x))^T(y-x).$$

The use of r^g enables the proof of the stronger implication (i) $\Rightarrow$ (ii). The question remains whether (ii) can be further sharpened.

5 Behavioral Equivalence of the Local and the Global Approach

In order to prove the behavioral equivalence of the local and the global theory we need a further step that is provided by the following

Proposition 5. *Let r be a skew-symmetric and continuous bifunction on X that is pseudo-concave-convex at every $x \in X$ and let Y be a convex subset of X with $\bar{x} \in Y$. Then the following statements are equivalent:*

(i) $r(\bar{x}, y) \geq 0$ *for all* $y \in Y$.
(ii) $\nabla_1 r(\bar{x}, \bar{x})^T(y - \bar{x}) \leq 0$ *for all* $y \in Y$.
(iii) $\nabla_2 r(\bar{x}, \bar{x})^T(y - \bar{x}) \geq 0$ *for all* $y \in Y$.

Proof. By (2) and skew-symmetry of r, (i) is immediately implied by (ii).
Assume now that (i) holds. Convexity of Y implies that $\bar{x} + t(y - \bar{x}) \in Y$ for $t \in [0,1]$. Thus, it follows from (i) that $r(\bar{x}, \bar{x} + t(y - \bar{x})) \geq 0$ for all $t \in [0,1]$. Hence, for all $t > 0$

$$\frac{1}{t}[r(\bar{x}, \bar{x} + t(y - \bar{x})) - r(\bar{x}, \bar{x})] \geq 0$$

and, consequently,

$$\nabla_2 r(\bar{x}, \bar{x})^T(y - \bar{x}) = \lim_{t \to 0^+} \frac{1}{t}[r(\bar{x}, \bar{x} + t(y - \bar{x})) - r(\bar{x}, \bar{x})] \geq 0,$$

i.e., we have shown that (i) implies (iii).
Finally, $\nabla_2 r(\bar{x}, \bar{x}) = -\nabla_1 r(\bar{x}, \bar{x})$ by skew-symmetry of r, i.e. (iii) and (ii) are equivalent. □

Assume that the behavior of a consumer with the consumption set X is described by a *choice correspondence* C on X which assigns to each nonempty and convex subset Y of X a (possibly empty) choice set $C(Y) \subseteq Y$. Then the local and the global approach are behaviorally equivalent in the sense of the following

Theorem 2. *If C is a choice correspondence on X, the following conditions are equivalent:*

(i) $C = C^g$ *for some local preference representation* g *on* X.
(ii) $C = C^r$ *for some diagonally pseudo-concave-convex global preference representation* r *on* X.

Moreover, the equivalence also holds for strictly pseudomonotone local representations and diagonally strictly pseudo-concave-convex global representations.

Proof. If (i) holds, then, by Theorem 1, (i) $\Rightarrow$ (ii), there is some diagonally pseudo-concave-convex global preference representation r^g such that $\nabla_1 r^g(x,x) = g(x)$ for every $x \in X$. This implies for every Y that $\bar{x} \in C^g(Y)$ iff $g(\bar{x})^T(y-\bar{x}) \leq 0$ for all $y \in Y$, i.e. $\nabla_1 r^g(\bar{x},\bar{x})^T(y-\bar{x}) \leq 0$ for all $y \in Y$. By Proposition 5, the latter statement is equivalent to $r^g(\bar{x},y) \geq 0$ for all $y \in Y$, i.e. to $\bar{x} \in C^{r^g}(Y)$.
Conversely, if (ii) is satisfied then we obtain for every Y that $\bar{x} \in C^r(Y)$ iff $r(\bar{x},y) \geq 0$ for all $y \in Y$. By Proposition 5, this is equivalent to the inequality $\nabla_1 r(\bar{x},\bar{x})^T(y-\bar{x}) \leq 0$ for all $y \in Y$, i.e. to $\bar{x} \in C^{g_r}(Y)$ for $g_r(x) = \nabla_1 r(x,x)$. By Theorem 1,(ii) $\Rightarrow$ (i), g_r is continuous and pseudomonotone, i.e. a local preference representation.
The strict case is proved analogously. □

References

1. Allen RGD (1932) The foundations of a mathematical theory of exchange. Economica 12:197–226
2. Crouzeix JP (2005) Criteria for generalized convexity and generalized monotonicity in the differentiable case. In: Hadjisavvas N, Komlósi S, Schaible S (eds) Handbook of Generalized Convexity and Generalized Monotonicity. Springer, Berlin Heidelberg New York
3. Debreu G (1959) Theory of value. Wiley, New York
4. Georgescu-Roegen N (1936) The pure theory of consumer's behavior. Quarterly J. of Economics 50:545–593
5. Georgescu-Roegen N (1954) Choice and revealed preference. Southern Economic Journal 21:119–130
6. Katzner D (1971) Demand and exchange analysis in the absence of integrability conditions. In: Chipman JS, Hurwicz L, Richter MK, Sonnenschein HF (eds) Preferences, Utility, and Demand. Harcourt Brace Jovanovich, New York
7. Konnov I (2005) Generalized monotone equilibrium problems and variational inequalities. In: Hadjisavvas N, Komlósi S, Schaible S (eds) Handbook of Generalized Convexity and Generalized Monotonicity. Springer, Berlin Heidelberg New York
8. Mas-Colell A, Whinston MD, Green JR (1995) Microeconomic theory. Oxford University Press, New York
9. Shafer WJ (1974) The nontransitive consumer. Econometrica 42:913–919
10. Shafer WJ (1987) Transitivity. In: Eatwell J, Milgate M, Newman P (eds) The New Palgrave: A Dictionary of Economics. Macmillan Press, London

Optimality Conditions for Convex Vector Functions by Mollified Derivatives

Davide La Torre

Department of Economics, Business and Statistics, University of Milan, Italy.
Email: davide.latorre@unimi.it

Summary. Necessary and sufficient optimality conditions for nonsmooth multiobjective optimization problems and some characterizations of convex vector functions are proved by means of mollified derivatives.

Key words: Mollified derivatives, convexity, nonsmooth optimization

1 Introduction

In this paper we study optimality conditions for vector functions by means of mollified derivatives. This type of generalized derivative was introduced first by Ermoliev, Norkin and Wets (see [8]) in the area of stochastic optimization and then studied in recent papers by Burke, Lewis and Overton [2] and by Crespi, La Torre and Rocca in [6, 7]. The earliest use of these tools in the context of nonsmooth optimization is probably due to the work of Craven [4, 5]. The main idea behind this technique consists of building smooth approximations of nonsmooth data and using these to obtain first and second order generalized derivatives. The smoothness techniques based on mollifiers seems to be a good tools for this purpose; they allow to have sequences of smooth functions with the same regularity of the mollifier. So if it is at least twice differentiable one can consider sets of cluster points of classical derivatives as use these to obtain generalized optimality conditions. Necessary conditions for nonsmooth multiobjective problems have been proved by Crespi, La Torre and Rocca in [7]; here we conclude the analysis of the vector case studying sufficient conditions and some characterizations of convex vector functions. These results follow a componentwise approach; the analysis of the vector case is reduced to the study of the properties of each component. In particular, section 2 recalls the notion of mollifier, and the definitions of first and second order mollified derivatives. Section 3 deals with optimality conditions for nonsmooth vector functions with and without convexity.

2 Preliminaries

We now recall some important definitions and results which will be useful in the following. From now on all the functions considered will be assumed to be locally integrable. The following two definitions recall the notion of mollifier and mollified functions. These are built taking the convolution between the nonsmooth data and the mollifier.

Definition 1. *[1] A sequence of mollifiers is any sequence of functions* $\psi_\epsilon : \mathbb{R}^m \to \mathbb{R}_+$, $\epsilon \downarrow 0$, *such that:*

i) $\text{supp}\ \psi_\epsilon := \{x \in \mathbb{R}^m \mid \psi_\epsilon(x) > 0\} \subseteq \rho_\epsilon cl\mathcal{B},\ \rho_\epsilon \downarrow 0,$
ii) $\int_{\mathbb{R}^m} \psi_\epsilon(x)dx = 1,$

where $\mathcal{B}$ *is the unit ball in* $\mathbb{R}^n$, clX *means the closure of the set* X *and* dx *denotes Lebesgue measure.*

Definition 2. *[1] Given a locally integrable function* $f : \mathbb{R}^m \to \mathbb{R}$ *and a sequence of bounded mollifiers, define the functions* $f_\epsilon(x)$ *through the convolution*

$$f_\epsilon(x) := \int_{\mathbb{R}^m} f(x-z)\psi_\epsilon(z)dz.$$

The sequence $f_\epsilon(x)$ *is said a sequence of mollified functions.*

A crucial point in this technique is the convergence of sequence of mollified functions f_ϵ to f. This is stated in the following results.

Theorem 1. *[1] Let* $f \in \mathcal{C}(\mathbb{R}^m)$. *Then* f_ϵ *converges continuosly to* f, *i.e.* $f_\epsilon(x_\epsilon) \to f(x)$ *for all* $x_\epsilon \to x$. *In fact* f_ϵ *converges uniformly to* f *on every compact subset of* $\mathbb{R}^m$ *as* $\epsilon \downarrow 0$.

Definition 3. *[10] A sequence of functions* $f_n : \mathbb{R}^m \to \mathbb{R}$ epi–converges *to* $f : \mathbb{R}^m \to \mathbb{R}$ *at* x, *if:*

i) $\liminf_{n\to+\infty} f_n(x_n) \geq f(x)$ *for all* $x_n \to x$;
ii) $\lim_{n\to+\infty} f_n(x_n) = f(x)$ *for some sequence* $x_n \to x$.

The sequence f_n *epi–converges to* f *if this holds for all* $x \in \mathbb{R}^m$, *in which case we write* $f = e - \lim f_n$.

Definition 4. *[8] A function* $f : \mathbb{R}^m \to \mathbb{R}$ *is said* strongly lower semicontinuous *(s.l.s.c.) at* x *if it is lower semicontinuous at* x *and there exists a sequence* $x_n \to x$ *with* f *continuous at* x_n *(for all* n*) such that* $f(x_n) \to f(x)$. *The function* f *is strongly lower semicontinuous if this holds at all* x.
The function f *is said* strongly upper semicontinuous *(s.u.s.c.) at* x *if it is upper semicontinuous at* x *and there exists a sequence* $x_n \to x$ *with* f *continuous at* x_n *(for all* n*) such that* $f(x_n) \to f(x)$. *The function* f *is strongly lower semicontinuous if this holds at all* x.

Theorem 2. *[8] Let $\varepsilon_n \downarrow 0$. For any s.u.s.c. function $f : \mathbb{R}^m \to \mathbb{R}$ and any associated sequence f_{ϵ_n} of mollified functions, we have for any $x \in \mathbb{R}^m$:*

i) $\limsup_{n\to+\infty} f_{\epsilon_n}(x_n) \le f(x)$ *for any sequence $x_n \to x$;*
ii) $\lim_{n\to+\infty} f_{\epsilon_n}(x_n) = f(x)$ *for some sequence $x_n \to x$.*

Proposition 1. *[11, 12] Whenever the mollifiers ψ_ϵ are of class $\mathcal{C}^k$, so are the associated mollified functions f_ε.*

The following definitions recall the notion of first order mollified derivatives (upper and lower).

Definition 5. *[8] Let $f : \mathbb{R}^m \to \mathbb{R}$, $\epsilon_n \downarrow 0$ as $n \to +\infty$ and consider the sequence f_{ϵ_n} of mollified functions with associated mollifiers $\psi_{\epsilon_n} \in \mathcal{C}^1$. The* upper mollified derivative *of f at x_0 in the direction $d \in \mathbb{R}^m$, with respect to (w.r.t.) the mollifiers sequence ψ_{ϵ_n} is defined as:*

$$\overline{\mathcal{D}}_\psi f(x_0, d) := \sup_{x_n \to x_0} \limsup_{n\to+\infty} \nabla f_{\epsilon_n}(x_n)^\top d.$$

Definition 6. *Let $f : \mathbb{R}^m \to \mathbb{R}$, $\epsilon_n \downarrow 0$ as $n \to +\infty$ and consider the sequence f_{ϵ_n} of mollified functions with associated mollifiers $\psi_{\epsilon_n} \in \mathcal{C}^1$. The* lower mollified derivative *of f at x_0 in the direction $d \in \mathbb{R}^m$, w.r.t. the mollifiers sequence ψ_{ϵ_n} is defined as:*

$$\underline{\mathcal{D}}_\psi f(x_0, d) := \inf_{x_n \to x_0} \liminf_{n\to+\infty} \nabla f_{\epsilon_n}(x_n)^\top d.$$

In [8] it has been defined also the following generalized gradient

$$\partial_\psi f(x_0) := \left\{ L := \limsup_{n\to+\infty} \nabla f_{\epsilon_n}(x_n), \quad x_n \to x_0 \right\}.$$

The following proposition states a relationship between this and the Clarke's generalized gradient.

Proposition 2. *[8] Let $f : \mathbb{R}^m \to \mathbb{R}$ be locally Lipschitz at x; then $\partial_\psi f(x)$ coincides with Clarke's generalized gradient and $\overline{\mathcal{D}}_\psi f(x_0, d)$ coincides with Clarke's generalized derivative ([3]).*

The following two propositions recall the properties of the first order derivatives.

Proposition 3. *[7] Let $f : \mathbb{R}^m \to \mathbb{R}$ and $x \in \mathbb{R}^m$. Then:*

i) $\overline{\mathcal{D}}_\psi f(\cdot; d)$ *is upper semicontinuous (u.s.c.) at x for all $d \in \mathbb{R}^m$;*
ii) $\underline{\mathcal{D}}_\psi f(\cdot; d)$ *is lower semicontinuous (l.s.c.) at x for all $d \in \mathbb{R}^m$.*

Proposition 4. *[7] $\overline{\mathcal{D}}_\psi f(x; \cdot)$ and $\underline{\mathcal{D}}_\psi f(x; \cdot)$ are positively homogeneous functions. Furthermore, if $\overline{\mathcal{D}}_\psi f(x; \cdot)$ ($\underline{\mathcal{D}}_\psi f(x; \cdot)$ respectively) is finite then it is subadditive (resp. superadditive) and hence convex (resp. concave) as a function of the direction d.*

We now recall the notion of second order mollified derivative. As in the first order, we use the regularity of the mollifier to build these generalized derivatives (upper and lower).

Definition 7. *[7] Let $f : \mathbb{R}^m \to \mathbb{R}$, $\epsilon_n \downarrow 0$ and consider the sequence of mollified functions f_{ϵ_n}, obtained from a family of mollifiers $\psi_{\epsilon_n} \in \mathcal{C}^2$. We define the* second–order upper mollified derivative *of f at x_0 in the directions d and v, w.r.t. to the mollifiers sequence ψ_{ϵ_n}, as:*

$$\overline{\mathcal{D}}^2_\psi f(x; d, v) := \sup_{x_n \to x} \limsup_{n \to +\infty} d^\top H f_{\epsilon_n}(x_n) v,$$

where $Hf_{\epsilon_n}(x)$ is the Hessian matrix of the function $f_{\epsilon_n} \in \mathcal{C}^2$ at the point x.

Definition 8. *[7] Let $f : \mathbb{R}^m \to \mathbb{R}$, $\epsilon_n \downarrow 0$ and consider the sequence of mollified functions f_{ϵ_n}, obtained from a family of mollifiers $\psi_{\epsilon_n} \in \mathcal{C}^2$. We define the* second–order lower mollified derivative *of f at x_0 in the directions d and v, w.r.t. the mollifiers sequence ψ_{ϵ_n}, as:*

$$\underline{\mathcal{D}}^2_\psi f(x; d, v) := \inf_{x_n \to x} \liminf_{n \to +\infty} d^\top H f_{\epsilon_n}(x_n) v.$$

Proposition 5. *[7] Let $f : \mathbb{R}^m \to \mathbb{R}$ and $x \in \mathbb{R}^m$.*

i) If $\lambda > 0$, then:

$$\overline{\mathcal{D}}^2_\psi \lambda f(x; d, d) = \lambda \overline{\mathcal{D}}^2_\psi f(x; d, d);$$
$$\underline{\mathcal{D}}^2_\psi \lambda f(x; d, d) = \lambda \underline{\mathcal{D}}^2_\psi f(x; d, d).$$

Moreover, if $\lambda < 0$ we get:

$$\overline{\mathcal{D}}^2_\psi \lambda f(x; d, d) = \lambda \underline{\mathcal{D}}^2_\psi f(x; d, d).$$

ii) The maps $(d, v) \to \overline{\mathcal{D}}^2_\psi f(x; d, v)$ and $(d, v) \to \underline{\mathcal{D}}^2_\psi f(x; d, v)$ are symmetric (that is $\overline{\mathcal{D}}^2_\psi f(x; d, v) = \overline{\mathcal{D}}^2_\psi f(x; v, d)$ and $\underline{\mathcal{D}}^2_\psi f(x; d, v) = \underline{\mathcal{D}}^2_\psi f(x; v, d)$).
iii) The functions $\overline{\mathcal{D}}^2_\psi f(x; d, \cdot)$ and $\underline{\mathcal{D}}^2_\psi f(x; d, \cdot)$ are positively homogeneous, whenever $d \in \mathbb{R}^m$.
iv) If $\overline{\mathcal{D}}^2_\psi f(x; \cdot, \cdot)$ ($\underline{\mathcal{D}}^2_\psi f(x; \cdot, \cdot)$ resp.) is finite, then it is sublinear (superlinear).
v) $\overline{\mathcal{D}}^2_\psi f(x; d, -v) = -\underline{\mathcal{D}}^2_\psi f(x; d, v)$.
vi) $\overline{\mathcal{D}}^2_\psi f(\cdot; d, v)$ is upper semicontinuous (u.s.c.) at x for every d, $v \in \mathbb{R}^m$.
vii) $\underline{\mathcal{D}}^2_\psi f(\cdot; d, v)$ is lower semicontinuous (l.s.c.) at x for every $d, v \in \mathbb{R}^m$.

Using these notions of derivatives it is possible the following generalized Taylor's formula which will be useful for proving optimality conditions.

Theorem 3. *[7] Let $f : \mathbb{R}^m \to \mathbb{R}$ be a s.l.s.c. (resp. s.u.s.c.) function and let $\epsilon_n \downarrow 0$, $t > 0$ and $d \in \mathbb{R}^m$.*

i) *If* $\psi_{\epsilon_n} \in \mathcal{C}^1$ *is a sequence of mollifiers, there exists a point* $\xi \in \big[x_0, x_0 + td\big]$ *such that:*

$$f(x_0 + td) - f(x_0) \le t\overline{\mathcal{D}}_{\psi} f(\xi; d).$$
$$\big(f(x_0 + td) - f(x_0) \ge t\underline{\mathcal{D}}_{\psi} f(\xi; d)\big)$$

ii) *If* $\psi_{\epsilon_n} \in \mathcal{C}^2$ *is a sequence of mollifiers, there exists* $\xi \in [x_0,\, x_0 + td]$ *such that:*

$$f(x_0 + td) - f(x_0) \le t\overline{\mathcal{D}}_{\psi} f(x_0; d) + \tfrac{t^2}{2}\overline{\mathcal{D}}^{\,2}_{\psi} f(\xi; d)$$
$$\big(f(x_0 + td) - f(x_0) \ge t\underline{\mathcal{D}}_{\psi} f(x_0; d) + \tfrac{t^2}{2}\underline{\mathcal{D}}^{\,2}_{\psi} f(\xi; d)\big)$$

assuming that the righthand sides are well defined, i.e. it does not happen the situation $+\infty - \infty$ *(in which the first term is* $+\infty$ *and the second in* $-\infty$*).*

3 Convex Vector Functions and Optimization

In this section we recall a characterization of convex vector functions and necessary conditions proved in [7] by means of second–order mollified derivatives. Then we study sufficient conditions with and without convexity. The following definition recalls the notion of convexity for vector functions.

Definition 9. *A vector function* $f : \mathbb{R}^m \to \mathbb{R}^l$ *is said to be* $\mathbb{R}^l_+$*-convex if each component* f_i*,* $i = 1 \ldots l$*, is convex.*

Lemma 1. *[9] Let* $f : \mathbb{R}^m \to \mathbb{R}$ *be a continuous function. Then* f *is convex if and only if the mollified functions* f_ε*, obtained from a sequence of mollifiers* ψ_ε*, are convex for every* $\varepsilon > 0$*.*

Lemma 2. *[13] Let* $f : \mathbb{R}^m \to \mathbb{R}$ *be a continuous function. Then* f *is convex if and only if:*

$$\frac{f(x + td) - 2f(x) + f(x - td)}{t^2} \ge 0,$$

$\forall x, d \in \mathbb{R}^m$, $\forall t \in \mathbb{R}$.

Theorem 4. *[7] Let* $f : \mathbb{R}^m \to \mathbb{R}$ *be a continuous function and let* $\varepsilon_n \downarrow 0$ *and* $\psi_{\varepsilon_n} \in \mathcal{C}^2$*. A necessary and sufficient condition for* f *to be convex is that:*

$$\underline{\mathcal{D}}^{\,2}_{\psi} f(x; d) \ge 0, \quad \forall x \in \mathbb{R}^m, \quad \forall d \in \mathbb{R}^m.$$

Corollary 1. *[7] Let* $f : \mathbb{R}^m \to \mathbb{R}^l$ *be a continuous function and let* $\varepsilon_n \downarrow 0$ *and* $\psi_{\varepsilon_n} \in \mathcal{C}^2$*. A necessary and sufficient condition for* f *to be* $\mathbb{R}^l_+$*-convex is that:*

$$\underline{\mathcal{D}}^{\,2}_{\psi} f_i(x; d) \ge 0, \quad \forall x \in \mathbb{R}^m, \forall d \in \mathbb{R}^m, \forall i = 1 \ldots l.$$

Given $f : \mathbb{R}^m \to \mathbb{R}^l$ and a subset $X \subset \mathbb{R}^m$ we now consider the following vector optimization problem:

$$VP) \qquad \min_{x \in X} f(x).$$

We recall that $a \leq_{\mathbb{R}^l} b$ if and only if $b - a \in \mathbb{R}^l_+$. For this type of problem the notion of of (weak) solution is recalled in the following definition.

Definition 10. *$x_0 \in X$ is a local (weak) solution of VP) if there exists a neighborhood U of x_0 such that $[f(U \cap X) - f(x_0)] \cap (-\mathbb{R}^l_+ \backslash \{0\}) = \emptyset$. ($[f(U \cap X) - f(x_0)] \cap (-\text{int}\, \mathbb{R}^l_+) = \emptyset$).*

In the sequel, the following definitions of first order set approximations will be useful.

Definition 11. *Let $x_0 \in \text{cl } X$, where* cl X *is the closure of the set X. We define the following sets*

- $WF(X, x_0) = \{d \in \mathbb{R}^m | \exists t_n \downarrow 0, x_0 + t_n d \in X\}$
- $T(X, x_0) = \{d \in \mathbb{R}^m | \exists t_n \downarrow 0, d_n \to d, x_0 + t_n d_n \in X\}$

which are called, respectively, the cone of weak feasible directions *and the* contingent cone*.*

Theorem 5 (First order necessary condition). *[7] Assume that f_i, $i = 1, \dots, l$, are s.l.s.c. functions. Let $\epsilon_n \downarrow 0$, $\psi_{\epsilon_n} \in C^1$, and $x_0 \in X$ be a weak local solution of VP). Then the following system has no solution on the set $WF(X, x_0)$:*

$$\overline{\mathcal{D}}_{\psi} f_i(x_0; d) < 0, \quad i = 1, \dots, l,$$

that is:

$$\max_{i=1,\dots,l} \overline{\mathcal{D}}_{\psi} f_i(x_0; d) \geq 0, \quad \forall d \in WF(X, x_0).$$

Theorem 6 (First order sufficient condition). *Assume that f_i, $i = 1, \dots, l$, are s.u.s.c. functions and the function $(x, u) \to \underline{\mathcal{D}}_{\psi} f_i(x, u)$ is lower semicontinuous, $\forall i = 1 \dots l$. Let $\epsilon_n \downarrow 0$, $\psi_{\epsilon_n} \in C^1$. If for all $d \in T(X, x_0)$ we have:*

$$\max_{i=1,\dots,l} \underline{\mathcal{D}}_{\psi} f_i(x_0; d) > 0$$

then x_0 is a local minimum point.

Proof. Suppose that x_0 is not a local minimum point; then there exists $x_n \to x_0$ such that $f(x_n) \in f(x_0) - (\mathbb{R}^m_+ \backslash \{0\})$. Then $x_n = x_0 + t_n d_n$ with $t_n \downarrow 0$, $d_n \to d$ and $d \in T(X, x_0)$. So, for all $i = 1 \dots l$, we have

$$0 \geq f_i(x_n) - f_i(x_0) \geq \underline{\mathcal{D}}_{\psi} f_i(\xi_n; d_n)$$

where $\xi_n \in [x_0, x_0 + t_n d_n]$. Taking the limit when $n \to +\infty$ and using the lower semicontinuity property, we obtain

$$0 \geq \underline{\mathcal{D}}_{\psi} f_i(x_0; d)$$

which is absurdo.

Lemma 3. *Let X be a convex subset of $\mathbb{R}^m$ and $f : \mathbb{R}^m \to \mathbb{R}^l$ be a $\mathbb{R}^l_+$ convex function. Then*

$$f(x) - f(y) - \overline{\mathcal{D}}_{\psi} f(y; x - y) \in \mathbb{R}^l_+$$

for all $x, y \in \mathbb{R}^m$.

Proof. By using the previous lemma 1, we obtain

$$f_{i,\varepsilon}(x) - f_{i,\varepsilon}(y) \geq \nabla f_{i,\varepsilon}(y)(x - y).$$

If $y_n \to y$ and $\epsilon_n \downarrow 0$, then

$$\begin{aligned} f_{i,\varepsilon_n}(x) - f_{i,\varepsilon_n}(y_n) &\geq \nabla f_{i,\varepsilon_n}(y_n)(x - y_n) \\ &= \nabla f_{i,\varepsilon_n}(y_n)(x - y) - \nabla f_{i,\varepsilon_n}(y_n)(y_n - y). \end{aligned}$$

Since f_i is convex then it is locally Lipschitz at the point y with a Lipschitz constant K_y. This implies that $\nabla f_{i,\varepsilon_n}(y_n)$ is bounded by K_y when n is sufficiently large. So

$$\lim_{n \to +\infty} \nabla f_{i,\varepsilon_n}(y_n)(y_n - y) = 0$$

and then we prove that

$$f(x) - f(y) - \overline{\mathcal{D}}_{\psi} f(y; x - y) \in \mathbb{R}^l_+.$$

Theorem 7. *Let X be convex subset of $\mathbb{R}^m$ and $x_0 \in \mathbb{R}^m$. If $f : \mathbb{R}^m \to \mathbb{R}^l$ is $\mathbb{R}^l_+$-convex and*

$$\max_{i=1,\ldots,l} \overline{\mathcal{D}}_{\psi} f_i(x_0; d) \geq 0, \quad \forall d \in WF(X, x_0).$$

then x_0 is a weak minimum point.

Proof. The hypothesis implies that

$$\overline{\mathcal{D}}_{\psi} f(x_0; d) \subset (-\operatorname{int} \mathbb{R}^l_+)^c$$

$\forall d \in WF(X, x_0)$ and so using the previous lemma

$$\begin{aligned} f(x) &\in f(x_0) + \overline{\mathcal{D}}_{\psi} f(x_0; x - x_0) + \mathbb{R}^l_+ \\ &\subset f(x_0) + \mathbb{R}^l_+ + (-\operatorname{int} \mathbb{R}^l_+)^c \\ &\subset f(x_0) + (-\operatorname{int} \mathbb{R}^l_+)^c. \end{aligned}$$

Definition 12. *The* sets of the descent directions $\overline{D}_{\leq}$ *and* $\underline{D}_{\leq}$ *for* f *at* x_0 *are:*

$$\overline{D}_{\leq}(f, x_0) = \{d \in \mathbb{R}^m : \overline{\mathcal{D}}_{\psi} f(x_0; d) \in -\mathbb{R}^l_+\},$$

$$\underline{D}_{\leq}(f, x_0) = \{d \in \mathbb{R}^m : \underline{\mathcal{D}}_{\psi} f(x_0; d) \in -\mathbb{R}^l_+\},$$

where $\overline{\mathcal{D}}_{\psi} f(x_0; d) := \{\overline{\mathcal{D}}_{\psi} f_1(x_0; d), \ldots, \overline{\mathcal{D}}_{\psi} f_l(x_0; d)\}$ *and* $\underline{\mathcal{D}}_{\psi} f(x_0; d) := \{\underline{\mathcal{D}}_{\psi} f_1(x_0; d), \ldots, \underline{\mathcal{D}}_{\psi} f_l(x_0; d)\}$.

Theorem 8. *[7] Assume that* f_i, $i = 1, \ldots, l$, *are s.l.s.c. functions,* $\epsilon_n \downarrow 0$, $\psi_{\epsilon_n} \in C^2$. *If* $x_0 \in X$ *is a local weak minimum point then*

$$\max_{i \in \overline{I}(x_0; d)} \overline{\mathcal{D}}^2_{\psi} f_i(x_0; d) \geq 0$$

for all $d \in WF(X, x_0) \cap \overline{D}_{\leq}(f, x_0)$, *where*

$$\overline{I}(x_0; d) = \left\{i : \overline{\mathcal{D}}_{\psi} f_i(x_0; d) = 0, i = 1 \ldots l\right\}.$$

Theorem 9. *Assume that* f_i, $i = 1, \ldots, l$, *are s.u.s.c. functions and* $\epsilon_n \downarrow 0$, $\psi_{\epsilon_n} \in C^2$. *Suppose that the functions* $(x, u) \to \underline{\mathcal{D}}_{\psi} f_i(x, u)$ *and* $(x, u) \to \underline{\mathcal{D}}^2_{\psi} f_i(x, u)$ *are lower semicontinuous,* $\forall i = 1, \ldots, l$. *If for all* $d \in T(X, x_0) \cap \underline{D}_{\leq}(f, x_0)$ *we have:*

$$\max_{i \in \underline{I}(x_0, d)} \underline{\mathcal{D}}^2_{\psi} f_i(x_0; d) > 0,$$

where

$$\underline{I}(x_0; d) = \left\{i : \underline{\mathcal{D}}_{\psi} f_i(x_0; d) = 0, i = 1 \ldots l\right\}.$$

then x_0 *is a local minimum point.*

Proof. Suppose that x_0 is not a local minimum point; then there exists a sequence $x_n \to x_0$ such that $f(x_n) \in f(x_0) - (\mathbb{R}^l_+ \backslash \{0\})$. If we build $d_n = (x_n - x_0)/\|x_n - x_0\|$ then $d_n \in S^1 = \{d \in \mathbb{R}^m : \|d\| = 1\}$ and so $d_n \to d \in S^1$. In other words $x_n = x_0 + t_n d_n$ and $d \in T(X, x_0)$. So

$$f(x_n) - f(x_0) \geq t_n \underline{\mathcal{D}}_{\psi} f(\xi_n; d_n)$$

and using the lower semicontinuity we have $d \in \underline{D}_{\leq}(f, x_0)$. For all $i \in \underline{I}(x_0, d)$, we have

$$0 \geq f_i(x_n) - f_i(x_0) - t_n \underline{\mathcal{D}}_{\psi} f_i(x_0; d) \geq \frac{t_n^2}{2} \underline{\mathcal{D}}^2_{\psi} f_i(\xi_n; d_n)$$

where $\xi_n \in [x_0, x_0 + t_n d_n]$. Taking the limit when $n \to +\infty$ and using the lower semicontinuity property, we obtain

$$0 \geq \underline{\mathcal{D}}^2_{\psi} f_i(x_0; d).$$

References

1. Brezis H (1963) Analysie fonctionelle – Theorie et applications. Masson editeur, Paris
2. Burke V, Lewis AS, Overton ML (2005) A robust gradient sampling algorithm for nonsmooth, nonconvex optimization SIAM Journal on Optimization 15:751–779
3. Clarke FH (1990) Optimization and nonsmooth analysis SIAM Classics in applied mathematics, Philadelphia
4. Craven BD (1986) Nondifferentiable optimization by smooth approximations Journal of Optimization Theory and Applications 17:3–17
5. Craven BD (1989) Nonsmooth multiobjective programming Numerical Functional Analysis and Optimization 10:49–64
6. Crespi G, La Torre D, Rocca M (2003) Mollified derivatives and second-order optimality conditions Journal of Nonlinear and Convex Analysis 4:437–454
7. Crespi G, La Torre D, Rocca M (2005) Secord order optimality conditions for nonsmooth multiobjective optimization problems. In Eberhard A, Hadjisavvas N, Luc DT (Eds.) Generalized convexity and generalized monotonicity. Springer, Berlin Heidelberg New York
8. Ermoliev YM, Norkin VI, Wets RJB (1995) The minimization of semicontinuous functions: mollifier subgradient SIAM Journal on Control and Optimization 33:149–167
9. Evans LC, Gariepy RF (1992) Measure theory and fine properties of functions. CRC Press, New York
10. Rockafellar RT, Wets RJB (1998) Variational Analysis. Springer, Berlin Heidelberg New York
11. Schwartz L (1966) Théorie des distributions. Hermann, Paris
12. Sobolev SL (1988) Some applications of functional analysis in mathematical physics. Nauka, Moscow
13. Zygmund A (1959) Trigonometric series. Cambridge Press, Cambridge

On Arcwise Connected Convex Multifunctions

Davide La Torre

Department of Economics, Business and Statistics, University of Milan, Italy.
Email: davide.latorre@unimi.it

Summary. In this paper arcwise connected convex multifunctions are introduced and studied. Optimality conditions involving this type of data are analyzed.

Key words: Convexity, set valued analysis, optimization

1 Introduction

The notions of convexity and generalized convexity have been studied by many authors in literature for the crucial role they play in analysis and optimization. The present paper generalizes to multifunctions the work by Fu and Wang [3] for vector functions. A multifunction $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is a function from $\mathbb{R}^n$ to the power set $2^{\mathbb{R}^m}$. We introduce a generalized definition of arcwise connected multifunction and we study nonsmooth optimization problems involving this type of data. In the following definitions the notions of local and global (weak) minimum point are recalled. Let $C \subset \mathbb{R}^m$ be a closed convex pointed cone.

Definition 1. *A point (x_0, y_0) with $y_0 \in F(x_0)$ is said to be a local weak minimum point if there exists a neighbourhood U of x_0 such that $F(x) \subseteq y_0 + (-\operatorname{int} C)^c$ for all $x \in U \cap K$. A point (x_0, y_0) with $y_0 \in F(x_0)$ is said to be a global weak minimum point if $F(x) \subseteq y_0 + (-\operatorname{int} C)^c$ for all $x \in K$.*

Definition 2. *A point (x_0, y_0) with $y_0 \in F(x_0)$ is said to be a local minimum point if there exists a neighbourhood U of x_0 such that $F(x) \subseteq y_0 + (-C\backslash\{0\})^c$ for all $x \in U \cap K$. A point (x_0, y_0) with $y_0 \in F(x_0)$ is said to be a global minimum point if $F(x) \subseteq y_0 + (-C\backslash\{0\})^c$ for all $x \in K$.*

Given a subset K of $\mathbb{R}^n$ we will focus the attention on constrained optimization problems as

$$\min_{x \in K} F(x).$$

We recall that the graph of F (see [1]) is the following subset of $\mathbb{R}^n \times \mathbb{R}^m$

$$\operatorname{graph} F = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m : y \in F(x)\}$$

If $F(x)$ is a closed, compact or convex we say that F is closed, compact or convex value, respectively. A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is a local selection of F at x_0 if there exists a neighbourhood $U(f, x_0)$ such that $f(x) \in F(x)$, $\forall x \in U(x_0)$. We denote by F_{x_0} the set of all local selections of F at x_0. If $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ and $c \in \mathbb{R}^m$ then the level set $\operatorname{lev}_c F$ is defined as

$$\operatorname{lev}_c F = \{x \in \mathbb{R}^n : c \in F(x) + C\}.$$

Several definitions of generalized convexity for multifunctions has been proposed in literature. Here we listed some of these and we show some relations among them.

Definition 3. *Let $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ be multifunction. F is said to be*

- *(i) weak C-convex if*

$$[tF(x_1) + (1-t)F(x_2)] \cap [F(tx_1 + (1-t)x_2) + C] \neq \emptyset$$

  whenever x_1, $x_2 \in \mathbb{R}^n$ and $t \in (0, 1)$;
- *(ii) almost C-convex if*

$$tF(x_1) + (1-t)F(x_2) \subseteq \operatorname{cl}[F(tx_1 + (1-t)x_2) + C]$$

  whenever x_1, $x_2 \in \mathbb{R}^n$ and $t \in (0, 1)$;
- *(iii) C-convex if*

$$tF(x_1) + (1-t)F(x_2) \subseteq F(tx_1 + (1-t)x_2) + C$$

  whenever x_1, $x_2 \in \mathbb{R}^n$ and $t \in (0, 1)$;
- *(iv) convex if*

$$tF(x_1) + (1-t)F(x_2) \subseteq F(tx_1 + (1-t)x_2)$$

  whenever x_1, $x_2 \in \mathbb{R}^n$ and $t \in (0, 1)$.
- *(v) quasi convex if the* $\operatorname{lev}_c F$ *is convex for all $c \in \mathbb{R}^m$.*

Notions (i) and (iii) have been firstly introduced in [4, 9] and the notion (ii) have been studied in [6]. Other works concerning the classical notion of convex multifunctions are [5, 7]. Other definitions of convexity for multifunctions that have been omitted here can be found in [4]. For notion (v) one can see [2]. It is trivial to prove the following relationships among the above definitions

$$\text{convexity} \;\Rightarrow\; C-\text{convexity} \;\Rightarrow \begin{cases} \text{almost C}-\text{convexity} \\ \text{weak C}-\text{convexity.} \end{cases}$$

Furthermore if F is C-convex then it is quasi convex.

2 Arcwise Connected Cone Convex Multifunctions

The aim of this section is to extend the definition of arcwise connected cone convex function (AC functions), introduced by Fu and Wang [3] for vector functions, to multifunctions. The following definition recalls the notion of arcwise connected set (AC set).

Definition 4. *[3] The subset $K \subseteq \mathbb{R}^n$ is to be arcwise connected set (AC set) if for any $x, y \in K$ there exists a continuous function, called arc, $H_{x,y} : [0,1] \to K$ such that $H_{x,y}(0) = x$ and $H_{x,y}(1) = y$.*

It is trivial to show that a convex set is an AC set and a convex multifunction is an AC multifunction. We now recall the definition of AC function.

Definition 5. *[3] Let $K \subset \mathbb{R}^n$ be an AC set and $f : K \to \mathbb{R}^m$. f is called an arcwise connected cone convex function (shortly, AC function) if for any $x, y \in K$ there is and arc $H_{x,y} \subset K$ such that*

$$tf(x) + (1-t)f(y) \subseteq f(H_{x,y}(t)) + C$$

for all $t \in [0,1]$.

Definition 6. *[3] Let $K \subset \mathbb{R}^n$ be an AC set and $f : K \to \mathbb{R}^m$. f is called an arcwise connected cone convex function (shortly, AC function) if for any $x, y \in K$ there is and arc $H_{x,y} \subset K$ such that*

$$tf(x) + (1-t)f(y) \subseteq f(H_{x,y}(t)) + C$$

for all $t \in [0,1]$.

The following definitions recall two notions of generalized derivatives for vector functions. We introduce the notion of Dini AC directional derivative for the first order and Peano AC directional derivative for the second order. These types of derivatives are built taking the set of cluster points of sequences of particular incremental ratios.

Definition 7. *Let $K \subset \mathbb{R}^n$ be an AC set, $f : K \to \mathbb{R}^m$ and $x_0 \in K$. Given $x \in K$ the Dini AC directional derivative of f with respect to $H_{x_0,x}$ is defined as*

$$f'(x_0; H_{x_0,x}) = \left\{ l = \lim_{n \to +\infty} \frac{f(H_{x_0,x}(t_n)) - f(x_0)}{t_n} : t_n \downarrow 0 \right\}.$$

We say that f is AC directional differentiable (see [3]) if

$$\lim_{t \downarrow 0} \frac{f(H_{x_0,x}(t)) - f(x_0)}{t}$$

exists.

The following definition of second order derivative follows the initial idea due to Peano; generalized derivatives can be built by considering generalized Taylor expansions without introducing the second order incremental ratio (that is the incremental ratio of the first order derivative). This idea has been studied by many authors for vector functions, usually under hypotheses of differentiability of the involve data. Here this definition is introduced without any request on the data; we only need that the Dini AC derivative is nonempty.

Definition 8. *Let K be an AC set, $x_0, x \in K$ and $z_0 \in f'(x_0; H_{x_0,x})$. The Peano AC directional derivative of f at (x_0, z_0) with respect to $H_{x_0,x}$ is defined as*

$$f''(x_0, z_0; H_{x_0,x}) = \left\{ l = \lim_{n \to +\infty} 2 \frac{f(H_{x_0,x}(t_n)) - f(x_0) - t_n z_0}{t_n^2} : t_n \downarrow 0 \right\}.$$

The following result characterizes an AC function by directional derivative.

Theorem 1. *[3] Let $K \subset \mathbb{R}^n$ be an AC set and $f : K \to \mathbb{R}^m$ be an AC function. If f is AC directionally differentiable then*

$$f(x) - f(x_0) \in f'(x_0; H_{x_0,x}) + C.$$

2.1 Extensions to Multifunctions

We now introduce the notion of arcwise connected cone convex multifunctions (AC multifunctions).

Definition 9. *Let K be an AC set. A multifunctions $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is called an arcwise connected cone convex multifunctions (shortly, AC multifunction) if for any $x, y \in K$ there is arc $H_{x,y}$ such that*

$$tF(x) + (1-t)F(y) \subseteq F(H_{x,y}(t)) + C$$

for all $t \in [0, 1]$.

It is easy to see when F is single valued this notion reduces to the definition of AC function.

Definition 10. *A set valued map $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is said to be arcwise connected quasi convex multifunction (ACQC multifunction) if the* $\mathrm{lev}_c F$ *is arcwise connected for all $c \in \mathbb{R}^m$.*

Definition 11. *A given multifunction $F : K \rightrightarrows \mathbb{R}^m$ is said to be convexlike (briefly, CL multifunction) if for all $x, y \in K$ and for all $\in [0, 1]$ there exists a $z_t \in K$ such that*

$$tF(x) + (1-t)F(y) \subseteq F(z_t) + C.$$

It is clear that an AC multifunction is also a CL multifunction. As for single-valued functions, for multifunctions it is possible to define several kind of derivatives. Here we extend the previous definitions of AC Dini derivative and AC Peano derivative to multifunctions.

Definition 12. *Let K be an AC set, $x_0 \in K$ and $y_0 \in F(x_0)$. Given $x \in K$ the Dini AC directional derivative of F with respect to $H_{x_0,x}$ is defined as*

$$F'(x_0, y_0; H_{x_0,x}) = \left\{ l = \lim_{n\to+\infty} \frac{f(H_{x_0,x}(t_n)) - y_0}{t_n} : t_n \downarrow 0, f \in \mathcal{F}_{x_0} \right\}.$$

Definition 13. *Let K be an AC set, $x_0, x \in K$, $y_0 \in F(x_0)$ and $z_0 \in F'(x_0, y_0; H_{x_0,x})$. The Peano AC directional derivative of F with respect to $H_{x_0,x}$ is defined as*

$$\begin{aligned} &F''(x_0, y_0, z_0; H_{x_0,x}) \\ &= \left\{ l = \lim_{n\to+\infty} 2\frac{f(H_{x_0,x}(t_n)) - y_0 - t_n z_0}{t_n^2} : t_n \downarrow 0, f \in \mathcal{F}_{x_0} \right\}. \end{aligned}$$

3 Preliminary Properties

Theorem 2. *Let $K \subseteq \mathbb{R}^n$ be an AC set and $F : K \rightrightarrows \mathbb{R}^m$. If F is AC multifunction then it is ACQC multifunction.*

Proof. Let $c \in \mathbb{R}^m$ and $x_1, x_2 \in \mathrm{lev}_c F$. Since F is AC then there exists an arc H_{x_1,x_2} such that

$$tF(x_1) + (1-t)F(x_2) \subseteq F(H_{x_1,x_2}(t)) + C.$$

Then

$$\begin{aligned} c = tc + (1-t)c &\in t(F(x_1) + C) + (1-t)(F(x_2) + C) \\ &= tF(x_1) + (1-t)F(x_2) + C \subseteq F(H_{x_1,x_2}(t)) + C \end{aligned}$$

that is $H_{x_1,x_2}(t) \in \mathrm{lev}_c F$ for all $t \in [0,1]$.

Theorem 3. *Let K be an AC set and $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ be an AC multifunction. If (x_0, y_0) is a local weak minimum then it is a global weak minimum.*

Proof. If (x_0, y_0) is a local weak minimum then there exists a neighbourhood U of x_0 such that

$$F(x) \subseteq y_0 + (-\mathrm{int}\, C)^c$$

for all $x \in K \cap U$. Given $x \in K$, there exists an arc $H_{x_0,x}$ such that

$$tF(x) + (1-t)F(x_0) \subseteq F(H_{x_0,x}(t)) + C$$

for all $t \in [0,1]$. For t sufficiently small we have

$$F(H_{x_0,x}(t)) + C \subseteq y_0 + (-\operatorname{int} C)^c.$$

Then

$$tF(x) + (1-t)y_0 \subseteq y_0 + (-\operatorname{int} C)^c$$

and this implies

$$F(x) \subseteq y_0 + (-\operatorname{int} C)^c$$

that is the thesis.

Theorem 4. *Let K be an AC set and suppose that $F(x) \subseteq f(x) + C$ for all $x \in K$. If f is a selection of F and f is an AC function then F is an AC multifunction.*

Proof. In fact for all $x \in K$ there exists an arc $H_{x_0,x}$ such that

$$tf(x) + (1-t)f(y) - f(H_{x_0,x}(t)) \in C$$

for all $t \in [0,1]$. So we have

$$\begin{aligned} tF(x) + (1-t)F(y) &\subseteq t(f(x) + C) + (1-t)(f(y) + C) \\ &\subseteq tf(x) + (1-t)f(y) + C \\ &\subseteq f(H_{x_0,x}(t)) + C \\ &\subseteq F(H_{x_0,x}(t)) + C \end{aligned}$$

for all $t \in [0,1]$.

Theorem 5. *Let K be an AC set and suppose that $F(x) \subseteq f(x) + C$ for all $x \in K$. If f is a selection of F and F is an AC multifunction then f is an AC function.*

Proof. In fact we have

$$\begin{aligned} tf(x_0) + (1-t)f(x) &\subseteq tF(x_0) + (1-t)F(x) \\ &\subseteq F(H_{x_0,x}(t)) + C \\ &\subseteq f(H_{x_0,x}(t)) + C. \end{aligned}$$

4 Necessary Optimality Conditions

We now prove necessary optimality conditions involving this type of generalized convex multifunctions. These conditions involve the first order and second order generalized directional derivatives we have introduced in section 2.

Theorem 6. *Let K be an AC set, $x_0 \in K$. Suppose that (x_0, y_0) be a weak local minimum point. Then*

$$F'(x_0, y_0, H_{x_0,x}) \cap -\operatorname{int} C = \emptyset$$

for all $x \in K$ and for all arcs $H_{x_0,x} : [0,1] \to K$, $H_{x_0,x}(0) = x_0$ and $H_{x_0,x}(1) = x$.

Proof. Suppose that there exists a $x \in K$ and an arc $H_{x_0,x}$ such that

$$F'(x_0, y_0, H_{x_0,x}) \cap -\operatorname{int} C \neq \emptyset.$$

So there exist a selection $f \in \mathcal{F}_{x_0}$, $t_n \downarrow 0$ and $l \in F'(x_0, y_0, H_{x_0,x}) \cap -\operatorname{int} C$ such that

$$l = \lim_{n \to +\infty} \frac{f(H_{x_0,x}(t_n)) - y_0}{t_n}.$$

Since (x_0, y_0) is a local weak minimum point we have for n large enough

$$\frac{f(H_{x_0,x}(t_n)) - y_0}{t_n} \subseteq (-\operatorname{int} C)^c$$

and this implies $l \in (-\operatorname{int} C)^c$.

Theorem 7. *Let K be an AC set, $x_0 \in K$. Suppose that (x_0, y_0) be a weak local minimum point. If there exists $x \in K$ and an arc $H_{x_0,x} : [0,1] \to K$, $H_{x_0,x}(0) = x_0$ and $H_{x_0,x}(1) = x$ and $F'(x_0, y_0, H_{x_0,x}) \subseteq -C \backslash - \operatorname{int} C$ then*

$$F''(x_0, y_0, z_0; H_{x_0,x}) \cap -\operatorname{int} C = \emptyset$$

for all $z_0 \in F'(x_0, y_0, H_{x_0,x})$.

Proof. If $\exists z_0 \in F'(x_0, y_0, H_{x_0,x})$ such that $F''(x_0, y_0, z_0; H_{x_0,x}) \cap -\operatorname{int} C \neq \emptyset$ then there exists a selection $f \in \mathcal{F}_{x_0}$, $t_n \downarrow 0$ and $l \in F''(x_0, x, z_0; H_{x_0,x}) \cap -\operatorname{int} C$ such that

$$l = \lim_{n \to +\infty} 2\frac{f(H_{x_0,x}(t_n)) - y_0 - t_n z_0}{t_n^2} \in (-\operatorname{int} C)^c + C \backslash \operatorname{int} C = (-\operatorname{int} C)^c.$$

5 Sufficient Optimality Conditions

In this section we prove sufficient conditions for the existence of minimum points for an AC multifunction.

Theorem 8. *Let K be an AC set, $x_0 \in K$ and suppose that F is an AC multifunction at x_0. Furthermore suppose that f is a selection of F and $F(x) \subseteq f(x) + C$ for all $x \in K$. If f is AC directionally differentiable at x_0 for all arcs $H_{x_0,x}$ then*

$$F(x) - y_0 \subseteq F'(x_0, f(x_0); H_{x_0,x}) + C.$$

Proof. From previous theorems we have that f is an AC function, that is for all $x \in K$ there exists an arc $H_{x_0,x}$ (see [3]) such that

$$f(x) - f(x_0) - f'(x_0; H_{x_0,x}) \in C.$$

So

$$\begin{aligned} F(x) - f(x_0) \subseteq f(x) - f(x_0) + C \subseteq f'(x_0; H_{x_0,x}) + C \\ \subseteq F'(x_0, f(x_0); H_{x_0,x}) + C. \end{aligned}$$

In the following we will denote by $B(x, \delta)$ the ball of radius δ centered at x.

Theorem 9. *Let K be an AC set, $x_0 \in K$, $y_0 \in F(x_0)$. Suppose that*

- *for all $x, y \in K$ and for all arcs $H_{x,y}$ we have $H_{x,y}([0,1]) \subset B(x, \|x - y\|) \cup B(y, \|x - y\|)$;*
- *there exist constants $K_{x_0} > 0$ and $\delta > 0$ such that $F(x) \subseteq y_0 + K_{x_0}\|x - x_0\|B(0,1)$ for all $x \in B(x_0, \delta)$;*
- *for all $x \in B(x_0, \delta)$ and for all arcs $H_{x_0,x}$ we have $F'(x_0, y_0; H_{x_0,x}) \cap -C = \emptyset$.*

Then (x_0, y_0) is a local minimum point.

Proof. Ab absurdo, suppose that there exists $x_n \to x_0$ and $y_n \in F(x_n)$, $x_n \in K$, such that $y_n \in y_0 - C \backslash \{0\}$. Eventually by extracting a subsequence, we suppose that $\|x_n - x_0\| \leq \frac{\delta}{n}$. Since K is an AC set between two points x_n and x_{n+1} there is an arc $H_{x_n,x_{n+1}}$. A function H^* from x_1 to x_0 can be built by

$$H^*_{x_1,x_0}(t) = H_{x_n,x_{n+1}}\left(-n\left(n+1\right)\left(t - 1/n\right)\right)$$

with $t \in [1/n, 1/(n+1)]$, $n \in \mathbb{N}$. So $H^*_{x_1,x_0}\left(\frac{1}{n}\right) = x_n$ and $H^*_{x_1,x_0}\left(\frac{1}{(n+1)}\right) = x_{n+1}$. From the first hypothesis easily follows that H_{x_1,x_0} is an arc (that is continuous) starting from $x_1 \in B(x_0, \delta) \cap K$. So we have

$$\frac{y_n - y_0}{1/n} = nK_{x_0}\|x_n - x_0\|b_n$$

and $b_n \to b_0$, $b_0 \in \overline{B(0,1)}$. Then

$$\frac{y_n - y_0}{1/n} \in \frac{F(H^*_{x_1,x_0}(1/n)) - y_0}{1/n}$$

and, eventually by extracting subsequences, $\frac{y_n - y_0}{1/n} \to l \in -C$.

Theorem 10. *Let K be an AC set, $x_0 \in K$ and $y_0 \in F(x_0)$. Suppose that*

- *for all $x \in K$ there exists an arc $H_{x_0,x}$ such that*

$$F(x) \subseteq y_0 + F'(x_0, y_0; H_{x_0,x}) + C;$$

- $F'(x_0, y_0; H_{x_0,x}) \cap -\mathrm{int}\, C = \emptyset$ *for all $x \in K$ and for all arcs $H_{x_0,x}$.*

Then (x_0, y_0) is a weak minimum point.

Proof. Given $x \in K$, there exists an arc $H_{x_0,x}$ such that

$$\begin{aligned} F(x) &\subseteq y_0 + F'(x_0, y_0; H_{x_0,x}) + C \\ &\subseteq y_0 + C + (-\mathrm{int}\, C)^c \subseteq y_0 + (-\mathrm{int}\, C)^c \end{aligned}$$

that is (x_0, y_0) is a weak minimum point.

References

1. Aubin JP, Frankowska H (1990) Set-valued analysis Systems & Control: Foundations & Applications. Birkhuser, Boston, MA
2. Benoist J, Popovici N (2003) Characterizations of convex and quasiconvex set-valued maps Mathematical Methods of Operation Research 57:427–435
3. Fu JY, Wang YH (2003) Arcwise connected cone-convex functions and mathematical programming Journal of Optimization Theory and Application 118:339–352
4. Kuroiwa D (1996) Convexity for Set-Valued Maps Applied Mathematics Letters 9:97–101
5. Robinson SM (1972) Normed convex processes Transaction of the American Mathematical Society 174:127–140
6. Sach PH, Yen ND (1997) Convexity criteria for set-valued maps Set Valued Analysis 5:37–45
7. Ursescu C (1998) Multifunctions with convex closed graph Czechoslovak Mathematical Journal 7:438–441
8. Yang XM, Li D, Wang SY (2001) Near-subconvexlikeness in vector optimization with set valued functions Journal of Optimization Theory and Application 110:413–427
9. Yang XQ (1998) Directional derivatives for set-valued mappings and applications Mathematical Methods of Operations Research 48:273–285

A Sequential Method for a Class of Bicriteria Problems

Laura Martein[1] and Valerio Bertolucci[2]

[1] Department of Statistics and Applied Mathematics, University of Pisa, Italy
lmartein@ec.unipi.it

[2] Department of Statistics and Applied Mathematics, University of Pisa, Italy
bertolucci@ec.unipi.it

Summary. The aim of the paper is to suggest a sequential method for generating the set E of all efficient points of a bicriteria problem P_B where the feasible region is a polytope and whose criteria are a linear function and a concave function which is the sum of a linear and the reciprocal of an affine function. The connectedness of E and some theoretical properties of P_B allow to give a finite simplex-like algorithm based on a suitable post-optimality analysis carried on a scalar parametric problem where the linear criteria plays the role of a parametric constraint.

Key words: Fractional programming, pseudoconcavity, bicriteria problems, parametric optimization.

1 Introduction

The bicriteria problem, that is the constrained problem of maximizing two objective functions, is widely studied and several algorithms have been suggested for several classes of functions [8]. A particular attention has been devoted to a bicriteria linear fractional problem, that is the problem where one criteria f_1 is a linear fractional function and the second one f_2 is linear since this occurs frequently in optimization problems involving criteria that are rates or ratios, such as return on investments, dividend coverage, margin on sales, productivity measures [4, 8, 9, 10, 12, 16]. For such a kind of bicriteria problem a sequential method has been suggested by one of the authors [4].
The aim of the paper is to extend this sequential method when f_1 is the sum of a linear and a linear fractional function (such kind of functions occurs in many applications and it has been studied by several authors [1, 9, 12, 15, 16]). Unfortunately, for this class of problems the set E of all efficient points is in general disconnected since f_1 can have several local maximum points not global. For such a reason we limit ourselves to consider the case where f_1 is also concave.

The concavity of the two criteria ensures the connectedness of E which can be characterized as the union of optimal solutions of a scalar parametric problem $P(\alpha)$ having the concave fractional function f_1 as the objective function while the second linear criteria f_2 plays the role of a parametric constraint.
A sequential method for generating E is obtained performing a suitable post-optimality analysis on $P(\alpha)$ which is based on some theoretical properties of the bicriteria problem. The post-optimality analysis utilizes a finite algorithm (suggested in Section 3) for solving the scalar problem of maximizing f_1 on a polytope.

2 Statement of the Problem and Preliminary Results

In this paper we consider the bicriteria problem where the first objective function is the sum between a linear and a linear fractional function, the second one is linear and the feasible set is a polytope.
It is known [14] that in order to guarantee the connectedness of the set E of all efficient points we must require the strictly quasiconcavity of the nonlinear function.
Recently it has been shown [2] that $f_1(x) = h^T x + \frac{c^T x + c_0}{d^T x + d_0}$ is pseudoconcave (in particular strict quasiconcave) on the halfspace $H = \{x \in \Re^n : d^T x + d_0 > 0\}$ if and only if $f_1(x) = h^T x - \frac{\gamma}{d^T x + d_0}$ with $\gamma > 0$, or $f_1(x) = k d^T x + \frac{c^T x + c_0}{d^T x + d_0}$ with $k < 0$. In particular the first function is concave and the second one is pseudoconcave.
From now on we consider the following concave bicriteria problem

$$P_B : \begin{cases} \max(f_1(x) = h^T x - \frac{\gamma}{d^T x + d_0},\ f_2(x) = q^T x) \\ \qquad\qquad x \in S \end{cases}$$

where $h,\ d \in \Re^n$, $d_0 \in \Re \backslash \{0\}$, $\gamma > 0$, $S = \{x \in \Re^n : Ax = b,\ x \geq 0\}$ is a compact set, A is $m \times n$ matrix with $rank(A) = m < n$, $b \in \Re^m$, $d^T x + d_0 > 0$, $\forall\ x \in S$ and $x \geq 0$ means $x \in \Re^n_+$.
The sequential method that we will suggest in Section 4 can be easily adapted to the case where the nonlinear criteria in P_B is substituted with $f_1(x) = k d^T x + \frac{c^T x + c_0}{d^T x + d_0}$.
In order to state some fundamental properties of P_B, consider the parametric problem

$$P(\alpha) : \begin{cases} \max f_1(x) \\ x \in S \\ q^T x = \alpha \end{cases}$$

with $\alpha \in [\alpha_{\min}, \alpha_{\max}]$ where

$$\alpha_{\max} = \{\max q^T x : x \in S\}$$

$$\alpha_{\min} = \{\max q^T x : x \in S,\ f_1(x) = M\}$$
$$M = \{\max f_1(x) : x \in S\}.$$

Let $S(\alpha)$ be the set of all optimal solutions of the problem $P(\alpha)$.
The following theorem holds.

Theorem 1. *Consider problem P_B. Then:*
i) E is not empty and connected;
ii) $x^0 \in E \cap intS$ if and only if there exists $\beta < 0$ such that

$$h + \frac{\gamma d}{(d^T x + d_0)^2} = \beta q \tag{1}$$

iii) problem $P(\alpha)$ is equivalent to the problem

$$\begin{cases} \max f_1(x) \\ x \in S \\ f_2(x) \geq \alpha \end{cases} \tag{2}$$

iv)

$$E = \bigcup_{\alpha \in [\alpha_{\min}, \alpha_{\max}]} S(\alpha)$$

Proof. i) It follows from the compactness of the feasible set and from the concavity of the objective functions.
ii) It follows applying the Kuhn-Tucker conditions to problem P_B which are necessary and sufficient since the two criteria are concave.
The proof of iii) and iv) can be found in [13].

Remark 1. Condition iii) points out that any optimal solution of problem (2) is binding to the parametric constraint; condition iv) holds in general when one of the objective functions of a bicriteria problem does not have a local maximum point different from the global one.

The algorithm that we will describe in Section 4 requires the knowledge of a sequential method for solving problem $P(\alpha)$; this will be done in the next section.

3 A Sequential Method for Calculating M

In order to suggest a sequential method for generating the set of all efficient points of problem P_B we need of an algorithm for solving the following concave maximization problem:

$$P : \begin{cases} \max f_1(x) = h^T x - \frac{\gamma}{d^T x + d_0} \\ \\ x \in S \end{cases}$$

Set:

$$\xi_{\min} = \{\min d^T x + d_0 : x \in S\},\ \xi_{\max} = \{\max d^T x + d_0 : x \in S\}.$$

Some fundamental properties of problem P are stated in the following theorem.

Theorem 2. *Consider problem P. Then:*
i) f_1 is concave on S;
ii) a local maximum point is also global;
iii) $x^0 \in intS$ is an optimal solution if and only if $h = \beta d$, $\beta < 0$, $-\frac{\gamma}{\beta} \in (\xi_{\min}, \xi_{\max})$;
iv) there exists an optimal solution x^0 of P belonging to an edge of S (in particular x^0 can be a vertex of S).

Proof. *i*) The Hessian matrix of f_1 is $H(x) = -\frac{2\gamma}{(d^T x + d_0)^3} dd^T$ so that $H(x)$ is negative semidefinite on the halfspace $H = \{x \in \Re^n : d^T x + d_0 > 0\}$.
ii) It follows from the concavity of f_1.
iii) The concavity of f_1 implies that $x^0 \in intS$ is a global maximum point if and only if $\nabla f_1(x^0) = h - \frac{\gamma d}{(d^T x + d_0)^2} = 0$. Setting $\beta = -\frac{\gamma}{(d^T x + d_0)^2}$ and taking into account the feasibility of x^0 the thesis is achieved.
iv) Let x^* be an optimal solution for problem P; then the linear program

$$\begin{cases} -\frac{\gamma}{d^T x^* + d_0} + \max h^T x \\ x \in S \cap \{x \in \Re^n : d^T x = d^T x^*\} = S^* \end{cases}$$

has an optimal solution on a vertex x^0 of S^* which belongs to an edge of S and such that $f_1(x^0) = f_1(x^*)$.

Following [7], we solve problem P by means of a suitable post-optimality analysis performed on the parametric problem

$$P(\xi) : \begin{cases} \max(h^T x - \frac{\gamma}{\xi}) \\ x \in S \\ d^T x + d_0 = \xi \end{cases}$$

where $\xi \in [\xi_{\min}, \xi_{\max}]$.
We will refer to every $\xi \in [\xi_{\min}, \xi_{\max}]$ as a feasible level and to any optimal solution of problem $P(\xi)$ as an optimal level solution.
The following theorem holds.

Theorem 3. *i) An optimal solution x^* of P is also an optimal level solution corresponding to the level $\xi^* = d^T x^* + d_0$;*
ii) if an optimal level solution x^* is a local maximum point with respect to an edge of S, then x^* is a global maximum point of P.

Proof. See [7].

The previous theorem suggests a procedure which generates a path of optimal level solutions the last of which is a global maximum point for P.
The basic ideas of the procedure are the following: we start from a vertex x^0 which is an optimal level solution (the first one will be an optimal solution of $P(\xi_{\min})$). If the directional derivatives of f_1 with respect to the edges starting from x^0 are not positive then x^0 is the global maximum point for P, otherwise we choose an edge s_k such that any of its points is an optimal level solution. We consider the restriction of f_1 on the edge s_k; if there exists a feasible maximum point for such a restriction then it is a global maximum point for P, otherwise we move on a suitable adjacent vertex and we repeat the analysis.
In order to describe analitically the procedure we introduce the following notations and we state some theoretical results.
Let x^0 be a vertex of the feasible region S with corresponding basis matrix A_B. We partition the matrix A as $A = [A_B, A_N]$ and the vectors x, h, d as $x^T = (x_B, x_N)$, $h^T = (h_B, h_N)$, $d^T = (d_B, d_N)$. Set:

- $\overline{h}_N^T = h_N^T - h_B^T A_B^{-1} A_N$, $\overline{d}_N^T = d_N^T - d_B^T A_B^{-1} A_N$;
- $\overline{h}_0 = h_B^T A_B^{-1} b$, $\overline{d}_0 = d_B^T A_B^{-1} b + d_0$;
- $\alpha_1 = \frac{x_{B_s}}{a_{sk}} = min\{\frac{x_{B_i}}{a_{ik}} : a_{ik} > 0\}$ where a_{ik} is the $i-$th element of the column $A_B^{-1} A_N^{(k)}$;
- $\Gamma = \overline{d}_0^2 \overline{h}_N + \gamma \overline{d}_N$;
- $J = \{j : \Gamma_j > 0\}$.

It is easy to verify that $\frac{\Gamma}{\overline{d}_0^2}$ is the vector of the directional derivatives of f_1 with respect to the directions associated to the feasible edges starting from x^0.
Let $z(x_{N_k})$ be the restriction of the function f_1 on the feasible edge s_k starting from x^0. We have:

$$z(x_{N_k}) = \overline{h}_{N_k} x_{N_k} + \overline{h}_0 - \frac{\gamma}{\overline{d}_{N_k} x_{N_k} + \overline{d}_0}$$

$$z'(x_{N_k}) = \overline{h}_{N_k} + \frac{\gamma \overline{d}_{N_k}}{(\overline{d}_{N_k} x_{N_k} + \overline{d}_0)^2}$$

The following theorem holds.

Theorem 4. *Let x^0 be a vertex of S which is an optimal level solution. Then*
i) x^0 is an optimal solution of P if and only if $J = \emptyset$;
ii) let k be the index such that

$$\frac{\overline{h}_{N_k}}{\overline{d}_{N_k}} = \max_{j \in J} \frac{\overline{h}_{N_j}}{\overline{d}_{N_j}}$$

Then any point of the feasible edge s_k starting from x^0 associated to x_{N_k} is an optimal level solution;

iii) If $\overline{h}_{N_k} \geq 0$ *then* $z(x_{N_k})$ *is increasing in* $[0, \alpha_2)$ *with* $\alpha_2 = +\infty$;

iv) If $\overline{h}_{N_k} < 0$, *then* $z(x_{N_k})$ *has a maximum point at* $\alpha_2 = -\frac{\overline{d}_0\overline{h}_{N_k}+\sqrt{-\gamma\overline{h}_{N_k}\overline{d}_{N_k}}}{\overline{h}_{N_k}\overline{d}_{N_k}}$.

Now we are able to summarize the algorithm given in [7] with respect to our class of functions.

ALGORITHM 1

STEP 1 Let $x^0 \in S$ be a vertex which is an optimal solution of the problem $\{\min d^T x + d_0,\ x \in S\}$. Go to step 2.

STEP 2 Calculate $\Gamma = \overline{d}_0^2\overline{h}_N + \gamma\overline{d}_N$. If $J = \emptyset$ then x^0 is an optimal solution of P : Stop. Otherwise go to step 3.

STEP 3 Calculate $k,\ \alpha_1,\ \alpha_2$ and set $\alpha = \min(\alpha_1, \alpha_2)$. If $\alpha = \alpha_1$ go to step 4. If $\alpha = \alpha_2$ then $\overline{x} = (x_B(\alpha_2), \alpha_2, 0)$ is an optimal solution of P : Stop.

STEP 4 The non-basic variable x_{N_k} enters the basis by means of a pivot operation on the element a_{sk}. Let x^0 be the new vertex. Go to step 2.

Remark 2. The set of optimal solutions of problem P is a convex set contained in the hyperplane $d^T x + d_0 = d^T x^0 + d_0$ where x^0 is the optimal solution generated by the algorithm.

Example 1. Consider the following problem

$$\begin{cases} \max(2x_1 - 5x_2 - \frac{195}{3x_1+4x_2+2}) \\ \\ 2x_1 - 2x_2 \leq 1 \\ 3x_1 - x_2 \leq 3 \\ x_2 \leq 3 \\ x_1, x_2 \geq 0 \end{cases}$$

Step 1 The denominator $d^T x + d_0 = 3x_1 + 4x_2 + 2$ reaches its minimum at $(0, 0)$ which is a vertex and an optimal level solution. The associated simplex-tableau is the following:

	0	2	−5	0	0	0
	−2	3	4	0	0	0
x_3	1	$\underline{2}$	−2	1	0	0
x_4	3	3	−1	0	1	0
x_5	3	0	1	0	0	1

Go to Step 2.

Step 2 It results

$$\Gamma = 2^2 \begin{pmatrix} 2 \\ -5 \end{pmatrix} + 195 \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 593 \\ 760 \end{pmatrix}.$$

Since $J \neq \emptyset$ go to Step 3.

Step 3 Since $\max\{\frac{2}{3}, -\frac{5}{4}\} = \frac{2}{3}$, we have $k = 1$, $\alpha_1 = \min\{\frac{1}{2}, 1\} = \frac{1}{2}$, $\alpha_2 =$

$+\infty$, so that $\alpha = \alpha_1 = \frac{1}{2}$. Go to Step 4.
Step 4 The nonbasic variable x_1 enters the basis and x_3 leaves the basis.

	-1	0	-3	-1	0	0
	$-\frac{7}{2}$	0	7	$-\frac{3}{2}$	0	0
x_1	$\frac{1}{2}$	1	-1	$\frac{1}{2}$	0	0
x_4	$\frac{3}{2}$	0	2	$-\frac{3}{2}$	1	0
x_5	3	0	1	0	0	1

Go to Step 2.
Step 2 It results

$$\Gamma = (\frac{7}{2})^2 \begin{pmatrix} -3 \\ -1 \end{pmatrix} + 195 \begin{pmatrix} 7 \\ -\frac{3}{2} \end{pmatrix} = \begin{pmatrix} \frac{5313}{4} \\ -\frac{1219}{4} \end{pmatrix}$$

Since $J \neq \emptyset$ go to Step 3.
Step 3 We have $k = 1$, $\alpha_1 = \min\{\frac{3}{4}, 3\} = \frac{3}{4}$, $\alpha_2 = -\frac{(\frac{7}{2})(-3)+\sqrt{-195(-3)7}}{(-3)7} = -\frac{1}{2} + \frac{1}{7}\sqrt{455}$ so that $\alpha = \alpha_1 = \frac{3}{4}$. Go to Step 4.
Step 4 The nonbasic variable x_2 enters the basis and x_4 leaves the basis.

	$\frac{5}{4}$	0	0	$-\frac{13}{4}$	$\frac{3}{2}$	0
	$-\frac{35}{4}$	0	0	$\frac{15}{4}$	$-\frac{7}{2}$	0
x_1	$\frac{5}{4}$	1	0	$-\frac{1}{4}$	$\frac{1}{2}$	0
x_2	$\frac{3}{4}$	0	1	$-\frac{3}{4}$	$\frac{1}{2}$	0
x_5	$\frac{9}{4}$	0	0	$\frac{3}{4}$	$-\frac{1}{2}$	1

Go to Step 2.
Step 2 It results

$$\Gamma = (\frac{35}{4})^2 \begin{pmatrix} -\frac{13}{4} \\ \frac{3}{2} \end{pmatrix} + 195 \begin{pmatrix} \frac{3}{2} \\ -\frac{7}{2} \end{pmatrix} = \begin{pmatrix} \frac{30875}{64} \\ -\frac{18165}{32} \end{pmatrix}$$

We have $k = 1$; $\alpha_1 = 3$, $\alpha_2 = -\frac{(\frac{35}{4})(-\frac{13}{4})+\sqrt{-195(-\frac{13}{4})\frac{15}{4}}}{(-\frac{13}{4})\frac{15}{4}} = \frac{5}{3}$, so that $\alpha = \alpha_2 = \frac{5}{3}$. We have $(x_1, x_2, x_5) = (\frac{5}{4}, \frac{3}{4}, \frac{9}{4}) + \frac{5}{3}(\frac{1}{4}, \frac{3}{4}, -\frac{3}{4}) = (\frac{5}{3}, 2, 1)$, so that the point $(\frac{5}{3}, 2)$ is the optimal solution of the problem.

4 A Sequential Method for a Generalized Fractional Bicriteria Problem

The procedure illustrated in the previous section allows us to propose a sequential method for solving the bicriteria problem P_B.
The theoretical properties established in Section 2 allow us to suggest a simple simplex-like procedure for generating the set E of all efficient points by means of a suitable post-optimality analysis performed on the parametric problem

$P(\alpha)$ starting from $P(\alpha_{\min})$.
In order to calculate M and to determine $\alpha_{\min}$, it is necessary, firstly, to solve problem P and it can be done by means of Algorithm 1.
If there exists a unique optimal solution x^0 for P, then $\alpha_{\min} = q^T x^0$, otherwise we must solve the following problem

$$\{\max q^T x : x \in S,\ h^T x - \frac{\gamma}{d^T x + d_0} = h^T x^0 - \frac{\gamma}{d^T x^0 + d_0}\} \tag{3}$$

which presents a nonlinear constraint.
The following theorem shows that problem (3) is equivalent to a suitable linear problem; more exactly we have the following theorem.

Theorem 5. *Let x^0 be an optimal solution of P. Then problem (3) is equivalent to the following problem*

$$\begin{cases} \max q^T x \\ h^T x = h^T x^0 \\ d^T x + d_0 = d^T x^0 + d_0 \\ x \in S \end{cases}$$

Proof. It is sufficient to note that the constraint $f_1(x) = M$ is verified for every $\overline{x} \in S$ such that $d^T \overline{x} + d_0 = d^T x^0 + d_0$ (see Remark (2)).

Taking into account Theorem 1, we have

$$E = \bigcup_{t \in [0, \alpha_{\max} - \alpha_{\min}]} S(t)$$

where $S(t)$ is the set of optimal solutions of the parametric problem

$$P(t) : \begin{cases} \max h^T x - \frac{\gamma}{d^T x + d_0} \\ Ax = b \\ q^T x = \alpha_{\min} + t \\ x \geq 0 \end{cases}$$

In order to solve problem $P(t)$ for every fixed t we apply the Algorithm 1 where now A is substituted with $\begin{pmatrix} A \\ q^T \end{pmatrix}$ and b is substituted with $\begin{pmatrix} b \\ \alpha_{\min} + t \end{pmatrix}$.
The post-optimality analysis is performed by studying the optimality condition $\Gamma(t) \leq 0$ and the feasibility condition $x_B(t) \geq 0$.
Set $O = \{t : \Gamma(t) \leq 0\}$ and $F = \{t : x_B(t) \geq 0\}$.
For any t which belongs to the intersection of O and F, $(x_B(t), 0)$ is the optimal solution of $P(t)$ and consequently it is an efficient point of P_B.
When t does not belong to O we can restore the optimality applying Theorem 4. When t does not belong to F we can restore the feasibility by means of a suitable pivot operation.
Now we are able to describe a sequential method for generating E.

Let us note that the slack variable associated to the parametric constraint must be a non basic variable in any step of the algorithm so that the column of the simplex tableau associated to such a variable is not involved in the process (it will be deleted in the next examples).

ALGORITHM 2

STEP 0 Determine M and $\alpha_{\min}$ and consider problem $P(t)$; set $\hat{t} = 0$ and go to Step 1.

STEP 1 Determine the sets O and F; if $O \cap F = \emptyset$ go to step 4, otherwise set $\bar{t} = \max\{t : t \in O \cap F\}$; $(x_B(t), 0)$ is an efficient point of P_B $\forall\, t \in [\hat{t}, \bar{t}]$, set $\hat{t} = \bar{t}$ and go to Step 2.

STEP 2 If $\hat{t}$ is an endpoint of F go to step 3; otherwise go to Step 4.

STEP 3 Let i be such that $x_{B_i}(\hat{t}) = 0$; if $a_{ij} \geq 0$, $\forall\ j$: Stop. Otherwise perform a pivot operation on the element a_{ik} such that

$$\frac{\Gamma_k(t)}{a_{ik}} = \min\left\{\frac{\Gamma_j(t)}{a_{ij}},\ a_{ij} < 0\right\}$$

and go to Step 1.

STEP 4 Let k be such that $\frac{\bar{h}_{N_k}}{\bar{d}_{N_k}} = \max_{j \in J} \frac{\bar{h}_{N_j}}{\bar{d}_{N_j}}$; calculate:

$$\alpha_2(t) = -\frac{\bar{d}_0(t)\bar{h}_{N_k} + \sqrt{-\gamma \bar{h}_{N_k}\bar{d}_{N_k}}}{\bar{h}_{N_k}\bar{d}_{N_k}}$$

$$t^* = \max\{t : \bar{x}_B(t) = x_B(t) - \alpha_2(t)A_B^{-1}A_N^{(k)} \geq 0\}.$$

Then $(\bar{x}_B(t), \alpha_2(t), 0)$ is an efficient point for P_B, $\forall\ t \in [\hat{t}, t^*]$. Let s be such that $\bar{x}_{B_s}(t^*) = 0$. Set $\hat{t} = t^*$, perform a pivot operation on a_{sk} and go to Step 1.

Example 2. Consider the following bicriteria problem:

$$\begin{cases} \max(2x_1 - 5x_2 - \frac{195}{3x_1+4x_2+2},\ x_2) \\ \\ 2x_1 - 2x_2 \leq 1 \\ 3x_1 - x_2 \leq 3 \\ x_2 \leq 3 \\ x_1, x_2 \geq 0 \end{cases}$$

Step 0 Referring to Example 1 we have $M = f_1(\frac{5}{3}, 2)$, so that $\alpha_{\min} = 2$. The introduction of the parametric constraint gives the following parametric problem:

$$P(t):\begin{cases}\max(2x_1 - 5x_2 - \frac{195}{3x_1+4x_2+2}) \\ \\ 2x_1 - 2x_2 \le 1 \\ 3x_1 - x_2 \le 3 \\ x_2 \le 3 \\ x_2 = 2 + t \\ x_1, x_2 \ge 0 \end{cases}$$

We insert the parametric constraint in the tableau corresponding to the vertex $(\frac{5}{4}, \frac{3}{4})$ and after a suitable update we obtain:

	$\frac{20}{3} + \frac{13}{3}t$	0	0	0	$-\frac{2}{3}$	0
	$-15 - 5t$	0	0	0	-1	0
x_1	$\frac{5}{3} + \frac{1}{3}t$	1	0	0	$\frac{1}{3}$	0
x_2	$2 + t$	0	1	0	0	0
x_5	$1 - t$	0	0	0	0	1
x_3	$\frac{5}{3} + \frac{4}{3}t$	0	0	1	$-\frac{2}{3}$	0

Go to Step 1.

Step 1 Taking into account that $t \ge 0$, we have $F = [0, 1]$, $\Gamma(t) = (15 + 5t)^2(-\frac{2}{3}) + 195(-1)$, $O = [0, +\infty)$, so that $O \cap F = [0, 1]$ and $\bar{t} = \max\{t : t \in O \cap F\} = 1$. Every point of the segment having $S = (\frac{5}{3}, 2)$, $T = (2, 3)$ as its endpoints is an efficient point. Set $\hat{t} = 1$ and go to Step 2.

Step 2 $\hat{t} = 1$ is an endpoint of F so that we go to Step 3.

Step 3 Any coefficient of the row corresponding to the variable x_5 is non-negative so that the algorithm terminates. It results $E = \overline{TS}$.

Example 3. Consider the following bicriteria problem:

$$\begin{cases}\max(2x_1 - 5x_2 - \frac{195}{3x_1+4x_2+2},\ -3x_1 + x_2) \\ \\ 2x_1 - 2x_2 \le 1 \\ 3x_1 - x_2 \le 3 \\ x_2 \le 3 \\ x_1, x_2 \ge 0 \end{cases}$$

Step 0 Referring to Example 1 we have $M = f_1(\frac{5}{3}, 2)$, so that $\alpha_{\min} = -3$. The introduction of the parametric constraint gives the following parametric problem:

$$P(t):\begin{cases}\max(2x_1 - 5x_2 - \frac{195}{3x_1+4x_2+2}) \\ \\ 2x_1 - 2x_2 \le 1 \\ 3x_1 - x_2 \le 3 \\ x_2 \le 3 \\ -3x_1 + x_2 = -3 + t \\ x_1, x_2 \ge 0 \end{cases}$$

We insert the parametric constraint in the tableau corresponding to the vertex $(\frac{5}{4}, \frac{3}{4})$ and after a suitable update we obtain

$$\begin{array}{c|c|ccccc|}
\cline{2-7}
 & -\frac{3}{2}t+\frac{5}{4} & 0 & 0 & -\frac{13}{4} & 0 & 0 \\
\cline{2-7}
 & \frac{7}{2}t-\frac{35}{4} & 0 & 0 & \frac{15}{4} & 0 & 0 \\
\cline{2-7}
x_1 & -\frac{1}{2}t+\frac{5}{4} & 1 & 0 & -\frac{1}{4} & 0 & 0 \\
x_2 & -\frac{1}{2}t+\frac{3}{4} & 0 & 1 & -\frac{3}{4} & 0 & 0 \\
x_5 & \frac{1}{2}t+\frac{9}{4} & 0 & 0 & \frac{3}{4} & 0 & 1 \\
x_4 & t & 0 & 0 & 0 & 1 & 0 \\
\cline{2-7}
\end{array}$$

and go to Step 1.

Step 1 We have $F = [0, \frac{3}{2}]$, $\Gamma(t) = (-\frac{7}{2}t + \frac{35}{4})^2(-\frac{13}{4}) + 195(\frac{15}{4})$, $O = [\frac{95}{14}, +\infty)$, so that $O \cap F = \emptyset$ and we go to Step 4.

Step 4 We have $k = 1$,

$$\alpha_2(t) = -\frac{(-\frac{7}{2}t + \frac{35}{4})(-\frac{13}{4}) + \sqrt{-195(-\frac{13}{4})(\frac{15}{4})}}{(-\frac{13}{4})(\frac{15}{4})} = \frac{14}{15}t - \frac{5}{3}$$

$$\overline{x}_B(t) = \begin{pmatrix} -\frac{1}{2}t + \frac{5}{4} \\ -\frac{1}{2}t + \frac{3}{4} \\ \frac{1}{2}t + \frac{9}{4} \\ t \end{pmatrix} + (\frac{14}{15}t - \frac{5}{3}) \begin{pmatrix} \frac{1}{4} \\ \frac{3}{4} \\ -\frac{3}{4} \\ 0 \end{pmatrix} = \begin{pmatrix} -\frac{4}{15}t + \frac{5}{3} \\ \frac{1}{5}t + 2 \\ -\frac{1}{5}t + 1 \\ t \end{pmatrix}$$

and $t^* = \max\{t : \bar{x}_B(t) \geq 0\} = 5$ with $\bar{x}_{B_5}(t^*) = 0$. For $t \in [0, 5]$ the point $(-\frac{4}{15}t + \frac{5}{3}, \frac{1}{5}t + 2)$ is an efficient interior point belonging to the segment of endpoints $S_1 = (\frac{5}{3}, 2)$, $T_1 = (\frac{1}{3}, 3)$. We have $x_5 = 0$ when $t = 5$, so that x_5 leaves the basis and x_3 enters the basis. We obtain the following simplex tableau:

$$\begin{array}{c|c|ccccc|}
\cline{2-7}
 & \frac{2}{3}t+11 & 0 & 0 & 0 & 0 & \frac{13}{3} \\
\cline{2-7}
 & t-20 & 0 & 0 & 0 & 0 & -5 \\
\cline{2-7}
x_1 & -\frac{1}{3}t+2 & 1 & 0 & 0 & 0 & \frac{1}{3} \\
x_2 & 3 & 0 & 1 & 0 & 0 & 1 \\
x_3 & \frac{2}{3}t+3 & 0 & 0 & 1 & 0 & \frac{4}{3} \\
x_4 & t & 0 & 0 & 0 & 1 & 0 \\
\cline{2-7}
\end{array}$$

Set $\hat{t} = 5$ and go to Step 1.

Step 1 We have $F = [5, 6]$, $\Gamma(t) = (\frac{1}{3}t - 2)^2(\frac{13}{3}) + 195(-5)$, $O = [5, 35]$, so that $O \cap F = [5, 6]$ and $\overline{t} = \max\{t : t \in O \cap F\} = 6$. Every point of the segment having $S_2 = (\frac{1}{3}, 3), T_2 = (0, 3)$ as endpoints is an efficient point. Set $\hat{t} = 6$ and go to Step 2.

Step 2 $\hat{t} = 6$ is an end point of F so that we go to Step 3.

Step 3 Any coefficient of the row corresponding to the variable x_1 is non-negative so that the algorithm terminates. It results $E = \overline{S_1T_1} \cup \overline{S_2T_2}$.

References

1. Cambini A (1981) An algorithm for a particular class of generalized convex programs. In: Schaible S, Ziemba WT (eds) Generalized concavity in optimization and economics, Academic Press, New York.
2. Cambini A, Crouzeix JP, Martein L (2002) On the pseudoconvexity of a quadratic fractional function. Optimization 51:677-687.
3. Cambini A, Martein L (1986) A modified version of Martos's algorithm for the linear fractional problem. Methods of Operation Research 53:33-44.
4. Cambini A, Martein L (1988) The linear fractional and the bicriteria linear fractional problem. In: Cambini A, Castagnoli E, Martein L, Mazzoleni P, Scaible S (eds) Generalized Convexity and Fractional Programming with Economic Applications, Springer Verlag, Berlin Heidelberg New York.
5. Cambini A, Martein L (1990) Reciprocity in optimization and efficiency in the bicriteria problem: a unified approach. In: Bühler W, Feichtinger G, Hartl RF, Radermacher FJ, Stähly P (eds) Operations Research Proceedings, Springer Verlag, Berlin Heidelberg New York.
6. Cambini A, Martein L (1992) Equivalence in linear fractional programming. Optimization 23:41-51.
7. Cambini A, Martein L, Schaible S (1989) On maximizing a sum of ratios. Journal of Information and Optimization Sciences 10:65-79.
8. Cambini A, Martein L, Stancu-Minasian (1996) Some developments in bicriteria fractional programming. In: Ding-Zhu DU, Xiang-Sun ZHANG, Kan CHENG (eds) Operations Research and its Applications, World Publishing Corporation.
9. Cambini A, Martein L, Stancu-Minasian (1999) A survey of bicriteria fractional programming. Electronic International Journal Advanced Modelling and Optimization 1:9-46.
10. Choo EU, Atkins DR (1982) Bicriteria linear fractional programming. Journal of Optimization Theory and Applications 36:203-226.
11. Ellero A (1996) The optimal level solutions method. Journal of Information and Optimization Sciences 17:355-372.
12. Frenk JBG, Schaible S (2005) Fractional programming. In: Hadjisavvas N, Komlosi S, Schaible S (eds) Handbook of Generalized Convexity and Generalized Monotonicity, Springer.
13. Martein L (1988) On the bicriteria maximization problem. In: Cambini A, Castagnoli E, Martein L, Mazzoleni P, Schaible S (eds) Generalized Convexity and Fractional Programming with Economic Applications, Springer Verlag, Berlin Heidelberg New York.
14. Schaible S (1983) Bicriteria quasiconcave programs. Cahiers du C.E.R.O. 25:93-101.
15. Schaible S (1981) A survey of fractional programming. In: Schaible S, Ziemba WT (eds) Generalized concavity in optimization and economics, Academic Press New York.
16. Schaible S (1995) Fractional programming. In: Host R, Pardalos PM (eds) Handbook of Global Optimization, Kluwer Acadamic Publishers.

Decomposition of the Measure in the Integral Representation of Piecewise Convex Curves

Mariana Nedelcheva

Technical University Varna, Department of Mathematics
9010 Varna, Bulgaria, mariana@mnet.bg

Summary. The notions of a convex arc and piecewise convex curve in the plane generalize the notion of a convex curve, the latter is usually defined as the boundary of a planar compact convex set with nonempty interior. The integral representation of a piecewise convex curve through a Riemann-Stieltjes integral with a corresponding one-dimensional measure is studied. It is shown that the Minkowski operations known from the convex sets can be generalized to piecewise convex curves. It is shown that the decomposition of the measure in the integral representation of the piecewise convex curve leads to a decomposition of the piecewise convex curve into a sum of corresponding piecewise convex curves. On this base, applying the natural decomposition of the one-dimensional measure into an absolutely continuous function, a jump function, and a singular function, the structure of a piecewise convex curve is investigated. As some curious consequences, the existence of polygons with infinitely many sides and no vertices, and polygons with infinitely many vertices and no sides is shown.

Key words: Convex arcs, convex curves, piecewise convex curves.

1 Introduction

The convex sets and convex curves in two dimensions are important topics in convex set theory. The convex curves usually are defined as boundaries of planar compact convex sets with nonempty interiors (we call such sets convex figures). By mean of this definition a convex curve is a simple closed curve in the plane. However the closedness of the curve gives restrictions in applications of the tool of integral representation introduced initially in Vitale [11]. Vitale [11] associates to a compact convex set a measure with the purpose to characterize its support function. Even though Vitale's paper appeared only as a preprint, it gained some popularity and its results have been used by other authors, say in [7], [9], [6] and [5]. The tool of the integral representations is developed and clarified later in [10]. For this purpose the notion of a convex curve is generalized in [10], where convex arcs and piecewise convex curves are defined. We recall these definitions in Section 2. The present paper continues the study of the integral representation of piecewise convex curves. In Section 3 it is shown that Minkowski operations known from the convex sets can be generalized to piecewise convex curves. In Section 4 it is shown that the decomposition of the measure in the integral representation of the piecewise convex curve leads to a decomposition of the piecewise convex curve into a sum of corresponding piecewise convex curves. On this base, applying in particular the natural decomposition of the one-dimensional measure into a jump function, an absolutely continuous function and a singular function, in Section 5 the structure of piecewise convex curves and convex curves in particular is investigated. Some curious consequences are the existence of convex curves being either polygonal curves with infinitely many sides and no vertices, or polygonal curves with infinitely many vertices and no sides.

In the present paper the author continues after [10] to develop the tool of the integral representation of piecewise convex curves, with the intention to apply it later to explain phenomena, which occur in the approximation of convex curves by polygonal curves, like the ones described in [2], [3] and [9]. Other applications are also in sight. For instance the developed tool could give another point of view to some of the problems concerning convex curves, say the ones exposed in [12]. As possible interdisciplinary application, let us mention that in economics the indifference curves and the level lines of the utility functions possess convexity properties, hence their analysis could be based on the developed here tool.

2 Convex Arcs and Piecewise Convex Curves

All considerations in this paper concern the Euclidean plane $\mathbb{R}^2$. The points in $\mathbb{R}^2$ and their radius-vectors are identified with pairs of reals. We make use of the transformations $T^+ : \mathbb{R}^2 \to \mathbb{R}^2$, $a = (a_1, a_2) \mapsto T^+a = (-a_2, a_1)$ and $T^- : \mathbb{R}^2 \to \mathbb{R}^2$, $a = (a_1, a_2) \mapsto T^-a = (a_2, -a_1)$ being in fact rotations on a right angle respectively in counter-clockwise and clockwise directions. For any two points c^1, $c^2 \in \mathbb{R}^2$ we denote by $\overline{c^1c^2}$ the segment with an initial point c^1 and a final point c^2. We denote also by $e_\theta = (\cos\theta, \sin\theta)$ the unit vector constituting with the x-axis an angle with measure θ.

We call an arc each set $\Gamma \subset \mathbb{R}^2$ homeomorphic to a compact interval $[\alpha, \beta] \subset \mathbb{R}$. In this convention the points in $\mathbb{R}^2$ are also arcs, since each point is homeomorphic to a degenerate interval. The homeomorphism $h : [\alpha, \beta] \to \Gamma$ introduces an ordering $\leq$ on Γ with the agreement $h(t_1) \leq h(t_2)$ if $t_1 \leq t_2$. We write also $h(t_1) < h(t_2)$ if $t_1 < t_2$. The arc with the introduced ordering is called an oriented arc. The points $h(\alpha)$ and $h(\beta)$ are called initial and final points of the oriented arc Γ.

We call the oriented arc Γ a convex arc if for any three points $c^i = (c^i_1, c^i_2) \in \Gamma$, $i = 1, 2, 3$, such that $c^1 < c^2 < c^3$, it holds

$$(c^1, c^2, c^3) := \begin{vmatrix} 1 & c_1^1 & c_2^1 \\ 1 & c_1^2 & c_2^2 \\ 1 & c_1^3 & c_2^3 \end{vmatrix} \geq 0 \,. \tag{1}$$

The above determinant gives the doubled value of the oriented area of the oriented triangle $c^1c^2c^3$, hence it is nonnegative if this triangle is counter-clockwise oriented. Roughly speaking, we call the oriented arc convex if it is counter-clockwise curved.

Here there are some examples of convex arcs. Each point in $\mathbb{R}^2$ is a convex arc. The segment $\overline{ab}$ is a convex arc with an initial point a and a final point b. The graph of a continuous convex function of one variable defined on a compact interval is a convex arc.

Denote by Γ a convex arc, and by a and b its initial and final points. Given any two points $c^1 < c^2$ in Γ, then we put $\Gamma_{c^1c^2} = \{c \in \Gamma \mid c^1 \leq c \leq c^2\}$. Since $\Gamma_{c^1c^2}$ is the image of the restriction of the homeomorphism determining Γ on a compact interval, we see that $\Gamma_{c^1c^2}$ is a convex arc.

Following [12] we call a convex figure any convex compact set in the plane with nonempty interior. The next theorem determines the structure of the convex arcs.

Theorem 1 ([10]). *Let Γ be a convex arc with a and b being its initial and final points. Then the following cases may occur:*
a) If $a = b$, then Γ degenerates to a point.
b) If $a \neq b$ and $(c^1, c^2, c^3) = 0$ for any three points $c^1 < c^2 < c^3$ of Γ, then Γ is the segment $\overline{ab}$.
c) If $a \neq b$ and $(c^1, c^2, c^3) > 0$ for at least one triple of points $c^1 < c^2 < c^3$ of Γ, then $\Gamma \cup \overline{ba}$ is the boundary of a convex figure.

For a convex arc Γ we introduced the set $\Phi_\Gamma = \operatorname{co}\Gamma$. After Theorem 1 we see that Φ_Γ is a point in case a), a segment in case b), and a convex figure in case c). In each case Φ_Γ is a compact convex set in the plane. Recall that the boundary of a convex figure is usually called a convex curve [12]. Therefore, each convex arc is either a point, or a segment, or a connected and closed proper subset of a convex curve. In the last case the convex curve can be taken as the boundary of $\operatorname{co}\Gamma$.

It is shown in [1] that each convex figure possesses a perimeter, that is each convex curve is rectifyable. Consequently, each convex arc Γ is rectifyable and therefore it admits an equation in natural parameter

$$\Gamma: \;\; r = f(s)\,, \quad 0 \leq s \leq L\,, \tag{2}$$

where the natural parameter s is the length of the arc from the initial point a to the current point. Here L is the length of Γ. The function f is continuous. In fact, as in differential geometry, it can be shown easily, that f is Lipschitz with constant 1. Since Γ has no multiple points, the function f is injective. It is well known that each continuous and injective mapping with domain a compact set is a homeomorphism. Therefore (2) is a homeomorphic representation of the arc Γ. Further, it can be shown that the passing from a parameter t determining the convex arc Γ to the natural parameter s is realized by a monotonely increasing function $s = s(t)$. This shows that the natural parameter s determines the same ordering on Γ as the parameter t, that is property (1) holds with respect to the ordering determined by the parameter s. Therefore (2) is a representation of Γ as a convex arc, which can be referred to as representation in natural parameter.

In Theorem 2 below we describe the function f in (2) in terms of a parameter θ being connected with the support function of Φ_Γ in direction e_θ. The support functions are an important tool when treating problems concerning convex figures. The representation obtained in Theorem 2 could play similar role when studying convex arcs. Further we will introduce the notion of a piecewise convex curve as a generalization of both the notion of a convex arc and a convex curve, and will extend the representation from Theorem 2 to piecewise convex curves. We need first the following notations.

Let K be a convex set in $\mathbb{R}^2$. We call a support function of K the function

$$\Lambda : \mathbb{R} \to \mathbb{R}\,, \quad \Lambda(\theta) = \sup\{r \cdot e_\theta \mid r \in K\}\,.$$

Here $r \cdot e_\theta$ denotes the scalar product of the radius-vector r and the vector e_θ.

The straight line $p_\theta : r \cdot e_\theta = \Lambda(\theta)$ is said to be a support line of K in direction e_θ. We will consider p_θ as an axis with orientation determined by the vector $T^+ e_\theta$ being colinear to p_θ.

Suppose that Γ is a convex arc with initial point a, final point b and parametric representation in natural parameter given by (2). When $a \neq b$ we denote by γ a real number, for which $e_\gamma = T^-(a-b)/\|a-b\|$. Here $\|\cdot\|$ denotes the Euclidean norm. When $a = b$, which according to Theorem 1 has place only if Γ degenerates to a point, we denote by γ any real number.

Let $\theta \in [\gamma, \gamma + 2\pi]$ and p_θ is the support line of Φ_Γ in direction e_θ. Let $\Phi_\Gamma \cap p_\theta$ be the segment with end points $r^-(\theta)$ and $r^+(\theta)$ where the direction from $r^-(\theta)$ to $r^+(\theta)$ coincides with the orientation on p_θ.

We put

$$c^-(\theta) = \begin{cases} a\,, & \theta = \gamma, \\ r^-(\theta)\,, & \gamma < \theta \le \gamma + 2\pi, \end{cases} \qquad c^+(\theta) = \begin{cases} r^+(\theta)\,, & \gamma \le \theta < \gamma + 2\pi, \\ b\,, & \theta = \gamma + 2\pi. \end{cases}$$

We determine the functions $s^-, s^+ : [\gamma, \gamma + 2\pi] \longrightarrow \mathbb{R}$ by

$$f(s^-(\theta)) = c^-(\theta), \quad f(s^+(\theta)) = c^+(\theta),$$

where f is the function from the representation (2) of Γ in natural parameter. In fact $s^-(\theta)$ gives the length of $\Gamma_{ac^-(\theta)}$ and $s^+(\theta)$ gives the length of $\Gamma_{ac^+(\theta)}$.

Theorem 2 ([10]). *Let Γ be a convex arc with initial point a and final point b. We put $\theta_a = \gamma$ and $\theta_b = \gamma + 2\pi$. Now $a = c^-(\theta_a)$ and $s^-(\theta_a) = 0$.*
a) The following integral representation has place:

$$\begin{gathered} c^+(\theta) = c^-(\theta_0) + T^+ e_{\theta_0} \left(s^+(\theta_0) - s^-(\theta_0) \right) + \int_{\theta_0}^{\theta} T^+ e_\lambda \, ds^+(\lambda) \,, \\ c^-(\theta) = c^-(\theta_0) + \int_{\theta_0}^{\theta} T^+ e_\lambda \, ds^-(\lambda) \,, \end{gathered} \tag{3}$$

for all $\theta_0 \in [\theta_a, \theta_b)$ and $\theta \in [\theta_0, \theta_b]$ (the integrals are in the sense of Riemann-Stieltjes).

b) The function f from the representation (2) in natural parameter is given by

$$f(s) = \begin{cases} c^-(\theta) \,, & s = s^-(\theta), \\ c^+(\theta) \,, & s = s^+(\theta), \\ c^-(\theta) \dfrac{s^+(\theta) - s}{s^+(\theta) - s^-(\theta)} + c^+(\theta) \dfrac{s - s^-(\theta)}{s^+(\theta) - s^-(\theta)} \,, & s^-(\theta) < s < s^+(\theta). \end{cases} \tag{4}$$

c) The support function Λ of Φ_Γ satisfies

$$\Lambda(\theta) = e_\theta \cdot c^-(\theta) = e_\theta \cdot c^+(\theta) \,, \quad \theta_a \le \theta \le \theta_b \,.$$

The obtained result is illustrated by the following example.

Example 1. Let K be the triangle $K = \{(x, y) \mid -1 \le x \le 0, \ -1 - x \le y \le 1 + x\}$. Define the convex arc Γ to be the counter-clockwise oriented part of the boundary of K from the point (0, 1) to the point (0, −1). We have $a = (0,\, 1)$, $\gamma = 0$, and $s^-,\ s^+ : [0,\, 2\pi] \to \mathbb{R}$ are given by

$$s^-(\theta) = \begin{cases} 0 \,, & 0 \le \theta \le 3\pi/4 \,, \\ \sqrt{2} \,, & 3\pi/4 < \theta \le 5\pi/4 \,, \\ 2\sqrt{2} \,, & 5\pi/4 < \theta \le 2\pi \,. \end{cases} \qquad s^+(\theta) = \begin{cases} 0 \,, & 0 \le \theta < 3\pi/4 \,, \\ \sqrt{2} \,, & 3\pi/4 \le \theta < 5\pi/4 \,, \\ 2\sqrt{2} \,, & 5\pi/4 \le \theta \le 2\pi \,. \end{cases}$$

Formula (3) gives

$$c^-(\theta) = \begin{cases} (0,\, 1), & 0 \le \theta \le 3\pi/4 \,, \\ (-1,\, 0), & 3\pi/4 < \theta \le 5\pi/4 \,, \\ (0,\, -1), & 5\pi/4 < \theta \le 2\pi \,, \end{cases} \quad c^+(\theta) = \begin{cases} (0,\, 1), & 0 \le \theta < 3\pi/4 \,, \\ (-1,\, 0), & 3\pi/4 \le \theta < 5\pi/4 \,, \\ (0,\, -1), & 5\pi/4 \le \theta \le 2\pi \,. \end{cases}$$

For the function f we get

$$f(s) = \begin{cases} (-s/\sqrt{2},\, 1 - s/\sqrt{2}) \,, & 0 \le s \le \sqrt{2} \,, \\ (-2 + s/\sqrt{2},\, 1 - s/\sqrt{2}) \,, & \sqrt{2} \le s \le 2\sqrt{2} \,. \end{cases}$$

Concerning Example 1 we can make the following remark. The function s^- and s^+ are defined on the interval $[0, 2\pi]$. The function s^- starts as a constant on the interval $[0, 3\pi/4]$. Similarly s^+ finishes as a constant on the interval $[5\pi/4, 2\pi]$. In such a situation the function f can be obtained if the calculations in Theorem 2 are done with any $\theta_a \in [0, 3\pi/4]$ and any $\theta_b \in [5\pi/4, 2\pi]$, in particular with $\theta_a = 3\pi/4$ and $\theta_b = 5\pi/4$. While Theorem 2 was formulated with interval $[\theta_a, \theta_b]$ with length 2π, Example 1 shows that sometimes the same calculations can be done applying a smaller interval.

Formula (3) with $\theta_0 = \theta_a$ transforms into

$$\begin{gathered} c^+(\theta) = a + T^+ e_{\theta_a}\, s^+(\theta_a) + \int_{\theta_a}^{\theta} T^+ e_\lambda \, ds^+(\lambda)\,, \\ c^-(\theta) = a + \int_{\theta_a}^{\theta} T^+ e_\lambda \, ds^-(\lambda)\,, \end{gathered} \tag{5}$$

true for all $\theta \in [\theta_a, \theta_b]$. This can be considered as an integral representation of the convex arc Γ, since in virtue of (4), once we have got the functions c^- and c^+, we can restore Γ. Let us underline, that the essential information in (5) is the knowledge of the initial point a, the interval $[\theta_a, \theta_b]$ and the function $s^+ : [\theta_a, \theta_b] \to \mathbb{R}$, which is monotonely increasing, nonnegative, and continuous from the right. The latter is seen from the next Theorem 3, where it is shown that the function s^- can be expressed by s^+. Turn attention there, that the knowledge of only s^- is not enough to restore s^+, for the value $s^+(\theta_b)$ cannot be obtained by s^-.

Let us underline that the Riemann-Stieltjes integral from a continuous function with respect to an increasing function exists always [4]. The function $\lambda \to T^+ e_\lambda = (-\sin\lambda, cos\lambda)$ is continuous. Therefore, the integrals in (5) exist always.

Theorem 3 ([10]). *Let $s+ : [\theta_a, \theta_b] \to \mathbb{R}$ be monotonely increasing, nonnegative, and continuous from the right function and $a \in \mathbb{R}^2$.*
Determine the function $s^- : [\theta_a, \theta_b] \to \mathbb{R}$ from the condition

$$s^-(\theta) = \begin{cases} 0\,, & \theta = \theta_a\,, \\ \lim_{\theta_1 \to \theta - 0} s^+(\theta_1) = s^+(\theta - 0)\,, & \theta_a < \theta \le \theta_b\,. \end{cases} \tag{6}$$

Determine $c^+(\theta)$ and $c^-(\theta)$ from (5) for all $\theta \in [\theta_a, \theta_b]$.
Under these conditions it holds:
The limit in (6) exists. The function s^- is monotonely increasing, nonnegative, and continuous from the left, moreover

$$\begin{gathered} s^+(\theta) = s^-(\theta + 0)\,,\ \theta_a \le \theta < \theta_b\,, \\ s^-(\theta) \le s^+(\theta)\,, \quad \theta_a \le \theta \le \theta_b\,. \end{gathered} \tag{7}$$

The function c^+ is continuous from the right, and c^- is continuous from the left, moreover

$$\begin{aligned} &c^+(\theta) = c^-(\theta + 0)\,,\; \theta_a \le \theta < \theta_b\,, \\ &c^-(\theta) = c^+(\theta - 0)\,,\; \theta_a < \theta \le \theta_b\,. \end{aligned} \tag{8}$$

In Example 1 we determined initially the convex curve Γ and on this base we obtained the point a, the interval $[\theta_a, \theta_b]$ and the function s^+. Now we pose the reverse question:

Given the point a, the interval $[\theta_a, \theta_b]$ and the function $s^+ : [\theta_a, \theta_b] \to \mathbb{R}$ being monotonely increasing, nonnegative, and continuous from the right. Define the function $s^- : [\theta_a, \theta_b] \to \mathbb{R}$ as $s^-(\theta_a) = 0$ and $s^-(\theta) = s^+(\theta - 0)$ for $\theta_a < \theta \le \theta_b$. The question is: What can be said for the curve Γ having equation (2) with function f determined by the integral representation (5) and formula (4)? In particular:

A. Is Γ a convex arc, at least when $\theta_b - \theta_a \le 2\pi$?
B. Is s a natural parameter for the obtained curve, i. e. is s the length of the arc from the initial point a to the current point?

In connection with Question A, let us say that in Theorem 2 it was $\theta_b - \theta_a = 2\pi$, but the possibility to take $\theta_b - \theta_a < 2\pi$ was noticed as a remark after Example 1. Still, let us say that the answer of Question A is negative as the following example shows.

Example 2. Let θ_a be any real and let $a = (0, 0)$. Put $\theta_b = \theta_a + \pi$ and define

$$s^+ : [\theta_a, \theta_b] \to \mathbb{R}\,, \quad s^+(\theta) = \begin{cases} 1\,, & \theta_a \le \theta < \theta_b\,, \\ 2\,, & \theta = \theta_b\,. \end{cases}$$

Then (5) does not represent a convex arc, since we get multiple points. In fact, the curve Γ corresponding to the representation (5) is the segment $\overline{ab}$ walked twice, once from a to b, and once from b to a. Here b is the point $b = (-\sin\theta_a, \cos\theta_a) = T^+ e_{\theta_a}$.

Example 2 shows that the answer of Question A is negative when $\theta_b - \theta_a \ge \pi$. The next theorem shows however, that the answer is still positive when $\theta_b - \theta_a < \pi$.

Theorem 4 ([10]). *Suppose that $a \in \mathbb{R}^2$ is a given point, $[\theta_a, \theta_b]$ is a given interval with length $\theta_b - \theta_a < \pi$, and $s^+ : [\theta_a, \theta_b] \to \mathbb{R}$ is a given monotonely increasing function, which is nonnegative and continuous from the right. Define the function $s^- : [\theta_a, \theta_b] \to \mathbb{R}$ as $s^-(\theta_a) = 0$ and $s^-(\theta) = s^+(\theta - 0)$ for $\theta_a < \theta \le \theta_b$. Then the curve Γ given by equation (2) with function f determined by the integral representation (5) and formula (4) is a convex arc and the parameter s in equation (2) is the natural parameter of Γ.*

Theorem 4 clarifies the answer of the posed problem. The curve Γ determined by $a \in \mathbb{R}^2$ and $s^+ : [\theta_a, \theta_b] \to \mathbb{R}$ is for sure a convex arc only if $\theta_b - \theta_a < \pi$. In the case, when this inequality is not satisfied, we can take a partition $\theta_a = \theta_0 < \theta_1 < \ldots < \theta_n = \theta_b$, of the interval $[\theta_a, \theta_b]$ with $\theta_i - \theta_{i-1} <$

π, $i = 1, \ldots, n$. These inequalities show that the restriction of the integral representation (5) to the interval $[\theta_{i-1}, \theta_i]$ gives a convex arc. More precisely, the convex arc that we have in mind when speaking for a restriction of (5) to the interval $[\theta_{i-1}, \theta_i]$ is obtained from (5) by obvious change of the initial point and diminishing s^+ by a constant, the latter in order for s^+ to play the role of a natural parameter. In fact this representation is the one given by (3), where θ_0 must be replaced by θ_{i-1} and the measures in the integrals must be diminished by $s^-(\theta_{i-1})$, see further (9). The variable θ then ranges in the interval $[\theta_{i-1}, \theta_i]$.

The made observation makes natural the following definition.

We call a piecewise convex curve any curve, which admits a representation given by (5) in the explained above sense, with some point $a \in \mathbb{R}^2$, interval $[\theta_a, \theta_b]$, and a monotonely increasing, nonnegative and continuous from the right function $s^+ : [\theta_a, \theta_b] \to \mathbb{R}$.

3 Operations with Piecewise Convex Curves

According to the definition given in the previous section a piecewise convex curve Γ can be identified with a triple $\Gamma = (a,\, I,\, s^+)$, where $a \in \mathbb{R}^2$ is a given point, $I = [\theta_a, \theta_b]$ is a given interval, and $s^+ : I \to \mathbb{R}$ is a nonnegative, continuous from the right, and monotonely increasing function. Since all the components of Γ appear in formulae (5), where $\theta \in [\theta_a, \theta_b]$, we may consider (5) as an integral representation of Γ. We say then that the corresponding parametric curve (2) is generated by the piecewise convex curve Γ. The reasoning made at the end of the previous section and based on Theorem 4 form the following result.

Theorem 5. *Suppose that $\Gamma = (a,\, [\theta_a, \theta_b],\, s^+)$ is a piecewise convex curve, which generates the parametric curve (2). Then the parameter s in (2) is the natural parameter.*

To interpret geometrically a piecewise convex curve $\Gamma = (a,\, I,\, s^+)$, where $I = [\theta_a, \theta_b]$, we take a partition $\theta_a = \theta_0 < \theta_1 < \cdots < \theta_n = \theta_b$ of the interval I, for which $\theta_i - \theta_{i-1} < \pi$, $i = 1, \ldots, n$. For $i = 1, \ldots, n$, we consider the piecewise convex curves $\Gamma^i = (c^-(\theta_{i-1}), [\theta_{i-1}, \theta_i], s_i^+)$, where

$$s_i^+ : [\theta_{i-1}, \theta_i] \to \mathbb{R}\,, \quad s_i^+ = \begin{cases} s^+(\theta) - s^-(\theta_{i-1})\,, & \theta_{i-1} \le \theta < \theta_i\,, \\ s^-(\theta_i) - s^-(\theta_{i-1})\,, & \theta = \theta_i,\ i = 1, \ldots, n-1, \\ s^+(\theta_i) - s^-(\theta_{i-1})\,, & \theta = \theta_i,\ i = n. \end{cases} \tag{9}$$

Let $\gamma^i : r = f^i(s)$, $0 \le s \le L_i$, be the parametric curve generated by Γ^i. Then Γ generates the parametric curve γ being the sum of $\gamma^1 + \ldots + \gamma^n$, the latter means that the initial point of each succeeding curve γ^i is the final point of the preceding curve γ^{i-1}. More precisely, γ is the parametric curve given by

$\gamma : r = f(s)$, $0 \le s \le L$, where $L = L_1 + \cdots + L_n$, and $f(s) = f^i(s - \sum_{k=1}^{i-1} L_i)$ for $L_{i-1} \le s \le L_i$, $i = 1, \dots, n$ (here we accept $L_0 = 0$ and $\sum_{i=1}^{0} L_i = 0$). Since further we define a sum of curves in other sense, the described here sum will be denoted $\gamma = \gamma^1 \oplus \cdots \oplus \gamma^n$ and we will call γ an oriented sum of $\gamma^1, \dots, \gamma^n$ (since the sign $\oplus$ is used here only in this meaning, there will be no confusion with other accepted practices of usage of this sign). Similarly, we write $\Gamma = \Gamma^1 \oplus \cdots \oplus \Gamma^n$ (and call it oriented sum) for the piecewise convex curves (even when not necessarily $\theta_i - \theta_{i-1} < \pi$, $i = 1, \dots, n$). This notation is justified, since as we show below, when $\theta_i - \theta_{i-1} < \pi$, the piecewise convex curve Γ^i can be identified with γ^i.

Theorem 6. *Let the piecewise convex curve $\Gamma = (a, I, s^+)$, with interval $I = [\theta_a, \theta_b]$ having length $\theta_b - \theta_a < \pi$ generate the parametric curve $\gamma : r = f(s)$, $0 \le s \le L$. Then, having fixed in advance the interval I, we can identify Γ with γ, in other words γ determines uniquely a and s^+.*

Proof. The equality $a = f(0)$ shows that γ determines uniquely the initial point a. Now we show that also s^+ is determined uniquely. According to Theorem 4 γ is a convex arc. According to Theorem 1 the following cases may occur:

a) $\gamma = \{a\}$, that is γ is a point. Then necessarily $s^+(\theta) = 0$, $\theta \in I$.

b) $\gamma = \overline{ab}$, $b \ne a$, that is γ is a nondegenerate segment.

Let θ_* be any real, such that $e_{\theta_*} = T^-(b-a)/\|b-a\|$. We claim that then there exists a uniquely determined $\theta_0 \in I$, such that

$$s^+(\theta) = \begin{cases} 0\,, \ \theta \in [\theta_a, \theta_0)\,, \\ \|b-a\|\,, \ \theta \in [\theta_0, \theta_b]\,. \end{cases}$$

It holds $\theta_0 = \theta_* + 2k_0\pi$ with some $k_0 \in \mathbb{Z}$ (with $\mathbb{Z}$ we denote the set of the integer reals).

To show this we observe that s^+ is a constant on each interval $[\theta', \theta''] \subset I$, such that $[\theta', \theta''] \cap \{\theta_* + m\pi \mid m \in \mathbb{Z}\} = \emptyset$. Assume in the contrary, that this is not true. Since γ is the segment $\overline{ab}$, we would have

$$\begin{aligned} 0 = e_{\theta_*} \cdot (c^+(\theta'') - c^+(\theta')) &= \int_{\theta'}^{\theta''} e_{\theta_*} \cdot T^+ e_\lambda \, ds^+(\lambda) \\ &= \int_{\theta'}^{\theta''} \sin(\lambda - \theta_*) \, ds^+(\lambda) \ne 0\,, \end{aligned} \tag{10}$$

a contradiction (the integral has the sign of $\sin(\lambda - \theta_*)$, which does not change for $\lambda \in [\theta', \theta'']$). Therefore s^+ is a jump function, whose jumps are only on the set $I \cap \{\theta_* + m\pi \mid m \in \mathbb{Z}\}$. This set is a unique point $\theta_0 = \theta_* + m_0\pi$, because the length of I is less than π. Now a direct calculation shows that Γ generates the segment $\overline{ab}$ only if $m_0 = 2k_0$ for some $k_0 \in \mathbb{Z}$ and the jump of s^+ at θ_0 is $\|b-a\|$.

c) $\gamma \cup \overline{ba}$ is the boundary of a convex figure Φ_γ, where $b \ne a$ is the final point of γ.

Let $[\bar{\theta}_a, \bar{\theta}_b]$ be an interval with less possible length, such that $e_{\bar{\theta}_a}$ and $e_{\bar{\theta}_b}$ are support directions for Φ_γ at the points a and b respectively. For any $\theta \in [\bar{\theta}_a, \bar{\theta}_b]$ we denote by p_θ the support line of Φ_γ in direction e_θ. Let $\Phi_\gamma \cap p_\theta = \overline{r^-(\theta)\, r^+(\theta)}$, where the direction from $r^-(\theta)$ to $r^+(\theta)$ coincides with the orientation on p_θ. Denote by $\sigma(\theta)$ the length of the arc $\gamma_{a\, r^+(\theta)}$. We claim that then there exists a uniquely determined $k_0 \in \mathbb{Z}$, such that $[\bar{\theta}_a + 2k_0\pi,\, \bar{\theta}_b + 2k_0\pi] \subset I$ and

$$s^+(\theta) = \begin{cases} 0\,, & \theta \in [\theta_a,\, \bar{\theta}_a + 2k_0\pi)\,, \\ \sigma(\theta - 2k_0\pi)\,, & \theta \in [\bar{\theta}_a + 2k_0\pi,\, \bar{\theta}_b + 2k_0\pi]\,, \\ \sigma(\bar{\theta}_b)\,, & \theta \in (\bar{\theta}_b + 2k_0\pi,\, \theta_b]\,. \end{cases} \tag{11}$$

To prove the claim we denote by k_0 the smallest $k \in \mathbb{Z}$, such that $\theta_a \le \bar{\theta}_a + 2k\pi < \bar{\theta}_b + 2k\pi$ (the strict inequality $\bar{\theta}_a < \bar{\theta}_b$ holds, since γ is not a segment). We prove that $\bar{\theta}_b + 2k_0\pi \le \theta_b$. If this is not the case, we would have $\theta_b - 2k_0\pi < \bar{\theta}_b$. Since b is the final point of γ, it holds $\sigma(\theta_b - 2k_0\pi) = \sigma(\bar{\theta}_b)$. This equality shows, that the interval $[\bar{\theta}_a, \bar{\theta}_b]$ can be diminished to $[\bar{\theta}_a, \theta_b - 2k_0\pi]$, and still $e_{\theta_b - 2k_0\pi}$ is the support direction of Φ_γ at b, which contradicts the minimality of $[\bar{\theta}_a, \bar{\theta}_b]$.
Thus $[\bar{\theta}_a + 2k_0\pi,\, \bar{\theta}_b + 2k_0\pi] \subset [\theta_a,\, \theta_b]$, whence in particular $0 < \bar{\theta}_b - \bar{\theta}_a \le \theta_b - \theta_a < \pi$. Like in case b) we show that s^+ is a constant on any interval $[\theta', \theta'']$ contained in $[\theta_a,\, \bar{\theta}_a + 2k_0\pi)$ or $(\bar{\theta}_b + 2k_0\pi,\, \theta_b]$ (the left-hand side in (10) is 0, since $c^+(\theta') = c^+(\theta'')$ and θ_* is replaced by $\theta_0 = (\theta_a + \theta_b)/2$). Since s^+ is continuous from the right, it is also a constant on the closed interval $[\bar{\theta}_b + 2k_0\pi,\, \theta_b]$. Let $\theta \in [\bar{\theta}_a + 2k_0\pi, \bar{\theta}_b + 2k_0\pi]$. Now $\theta - 2k_0\pi \in [\bar{\theta}_a, \bar{\theta}_b]$ and $e_\theta = e_{\theta - 2k_0\pi}$. From the definitions of c^+ (in 5) and r^+ we have $c^+(\theta) = r^+(\theta - 2k_0\pi)$. According to Theorem 4 the length of $\gamma_{a\, c^+(\theta)} = \gamma_{a\, r^+(\theta - 2k_0\pi)}$ is given by $s^+(\theta) = \sigma(\theta - 2k_0\pi)$, which proves the claim. □

We consider two piecewise convex curves $\Gamma^i = (a^i,\, I_i,\, s_i^+)$ as different, if they differ in at least one component. The following example shows that the conclusion of Theorem 6 does not hold in general for larger intervals. Namely, we show that two different piecewise convex curve can generate the same parametric curve.

Example 3. Consider the piecewise convex curves $\Gamma^i = (a^i,\, I_i,\, s_i^+)$, $i = 1, 2$, determined by the point $a^1 = a^2 = (1, 0)$, the interval $I_1 = I_2 = [0, 4\pi]$ and the functions $s_i^+ : I_i \to \mathbb{R}$, $i = 1, 2$, given by

$$s_1^+(\theta) = \begin{cases} \theta\,, & 0 \le \theta \le 2\pi\,, \\ 2\pi\,, & 2\pi \le \theta \le 4\pi\,, \end{cases} \qquad s_2^+(\theta) = \begin{cases} 0\,, & 0 \le \theta \le 2\pi\,, \\ \theta - 2\pi\,, & 2\pi \le \theta \le 4\pi\,. \end{cases}$$

Then these piecewise convex curves are different, but they generate the same parametric curve, namely the circle $r = (\cos s,\, \sin s)$, $0 \le s \le 2\pi$.

Still, the result of Theorem 6 admits some improvement.

Theorem 7. *Let the piecewise convex curve* $\Gamma = (a, I, s^+)$ *with interval* $I = [\theta_a, \theta_b]$ *having length* $\theta_b - \theta_a \le 2\pi$ *generate the parametric curve* $\gamma : r = f(s)$, $0 \le s \le L$. *If* $\theta_b - \theta_a = 2\pi$ *suppose also that* $s^+(\theta_b - 0) = s^+(\theta_b)$. *Then, having fixed in advance the interval* I, *we can identify* Γ *with* γ, *in other words* γ *determines uniquely* a *and* s^+.

Proof. We suppose first $\theta_b - \theta_a < 2\pi$. Put $\theta_0 = (\theta_a + \theta_b)/2$. Consider the oriented sum $\Gamma = \Gamma^1 \oplus \Gamma^2$ corresponding to the partition $\theta_a < \theta_0 < \theta_b$. For the respective generated parametric curves we have $\gamma = \gamma^1 \oplus \gamma^2$, where γ^1 and γ^2 are convex arcs because of the intervals $[\theta_a, \theta_0]$ and $[\theta_0, \theta_b]$ have lengths smaller than π. We denote by $[\bar{\theta}_a, \bar{\theta}_b]$ an interval with less possible length, such that $e_{\bar{\theta}_a}$ and $e_{\bar{\theta}_b}$ are support directions for Φ_{γ^1} and Φ_{γ^2} at the points a and b respectively. Obviously e_{θ_0} is a support direction for Φ_{γ^1} and Φ_{γ^2} at the point $c = c^-(\theta_0)$, which is the final point for γ^1 and the initial point for γ^2. Let p_θ be the support line of Φ_{γ^1} for $\theta \in [\bar{\theta}_a, \theta_0)$ or the support line of Φ_{γ^2} for $\theta \in [\theta_0, \bar{\theta}_b]$. Denote $\Phi_{\gamma^i} \cap p_\theta = \overline{r^-(\theta) r^+(\theta)}$ with an orientation of p_θ coinciding with the orientation from $r^-(\theta)$ to $r^+(\theta)$. Let $\sigma(\theta)$ is the length of the curve $\gamma_{ar^+(\theta)}$. In a similar way like in the proof of Theorem 6 case c) we show that there exists a unique $k_0 \in \mathbb{Z}$, for which $[\bar{\theta}_a + 2k_0\pi, \bar{\theta}_b + 2k_0\pi] \subset I$ and (11) holds. The uniqueness of $s^+(\theta)$ is proved.

Let now $\theta_b - \theta_a = 2\pi$ and $s^+(\theta_b - 0) = s^+(\theta_b)$. For a sequence $\theta_n \to \theta_b$ with $\theta_n < \theta_b$, $n = 1, 2, \dots$, we consider the piecewise convex curves $\Gamma_n = (a, I, s_n^+)$, $n = 1, 2, \dots$, where

$$s_n^+(\theta) = \begin{cases} s^+(\theta)\,, \ \theta_a \le \theta < \theta_n\,, \\ s^+(\theta_n)\,, \ \theta_n \le \theta \le \theta_b\,. \end{cases}$$

Let γ_n be the generated by Γ_n parametric curve. The relation $\theta_n - \theta_a < 2\pi$ implies that s_n^+ is uniquely determined by γ_n. Obviously $s_n^+(\theta) \to s^+(\theta)$ according to the continuity from the left of s^+ at the point θ_b, whence s^+ is uniquely determined. □

The proved theorem shows that a piecewise convex curve $\Gamma = (a, I, s^+)$, for which the length of the interval I is not greater than 2π (and in the case of 2π length the function s^+ is continuous at the right end of I) can be identified with the generated by Γ parametric curve. The following simple example shows that in the case of a length 2π the assumption that s^+ is continuous at the right end point of I is essential.

Example 4. Consider the piecewise convex curves $\Gamma^i = (a^i, I_i, s_i^+)$, $i = 1, 2$, determined by the point $a^1 = a^2 = (0, 0)$, the interval $I_1 = I_2 = [0, 2\pi]$, and the functions $s_i^+ : I_i \to \mathbb{R}$, $i = 1, 2$, given by $s_1^+(\theta) = 1$ for $0 \le \theta \le 2\pi$, and $s_2^+(\theta) = 0$ for $0 \le \theta < 2\pi$ and $s_2^+(2\pi) = 1$. Then both Γ^1 and Γ^2 generate the same parametric curve $\gamma : \ r = (0, s)$, $0 \le s \le 1$.

In geometry by *curve* we mean usually the image of an interval by a continuous mapping. Different mapping can define the same curve. Each particular of these mappings is called a parametric curve, or a parameterization of the given curve. Since often the properties of a curve are derived by its parameterization, the notion of a parametric curve plays an important role to the extent that we identify curves and parametric curves. This comment underlines the importance of Theorems 6 and 7, which describe when a piecewise convex curve could be identified with a parametric curve. Example 4 shows however, that in general such an identification does not hold. A piecewise convex curve Γ is a more complicated object than a parametric curve. Namely, it is a parametric curve γ and a given normal vector $\theta \to e_\theta$, $\theta \in [\theta_a, \theta_b]$ at the points on $\overline{c^-(\theta)\, c^+(\theta)} \subset \gamma$. The normal vector at the initial point a is e_{θ_a} and the normal vector at the final point b is e_{θ_b}. The existence of a normal vector to a piecewise convex curve $\Gamma = (a, I, s^+)$ allows to introduce a support function $\Lambda : I \to \mathbb{R}$ putting $\Lambda(\theta) = e_\theta \cdot c^+(\theta)$. We write also Λ_Γ instead of Λ to underline the dependence on Γ. Now the tool of the support functions can be applied to piecewise convex curves in the way, in which support functions are applied to investigate convex sets and convex curves.

Now we define the following operations with piecewise convex curves.

Multiplication of a piecewise convex curve with a nonnegative scalar. Given the piecewise convex curve $\Gamma = (a, I, s^+)$ and a nonnegative real λ, we put $\lambda\, \Gamma = (\lambda a, I, \lambda s^+)$.

Sum of piecewise convex curves. Given the piecewise convex curves $\Gamma^i = (a^i, I, s_i^+)$, $i = 1, 2$, defined on the same interval I, we put $\Gamma^1 + \Gamma^2 = (a^1 + a^2, I, s_1^+ + s_2^+)$.

In the above definitions both $\lambda\, \Gamma$ and $\Gamma^1 + \Gamma^2$ are piecewise convex curves. In fact, if s^+ is a nonnegative, continuous from the right, monotonely increasing function on I, the same property obeys $\lambda\, s^+$. Similarly, if both s_1^+ and s_2^+ are nonnegative, continuous from the right, monotonely increasing functions on I, the same is true for $s_1^+ + s_2^+$.

More generally, we define a linear combination of piecewise convex curves with nonnegative coefficients as follows. Given the piecewise convex curves $\Gamma^i = (a^i, I, s_i^+)$, $i = 1, \ldots, n$, defined on the same interval I, and the nonnegative reals λ_i, we put $\sum_{i=1}^n \lambda_i\, \Gamma^i = \left(\sum_{i=1}^n \lambda_i\, a^i, I, \sum_{i=1}^n \lambda_i s_i^+\right)$. According to Theorem 5 in the above linear combination the sum $\sum_{i=1}^n \lambda_i s_i^+$ is in an obvious manner related to the natural parameter of the generated curve. In particular, if $L(\Gamma)$ denotes the length of a piecewise convex curve, then $L(\sum_{i=1}^n \lambda_i\, \Gamma^i) = \sum_{i=1}^n \lambda_i\, L(\Gamma^i)$.

According to Theorem 7, in the case when the interval I has a length not greater than 2π (and in the case of 2π length the function s^+ is continuous at the right end of I), the introduced operations between piecewise convex curves can be considered also as operations between the generated by them parametric curves and vice versa. Assuming that the operations are defined in some direct manner on the parametric curves, we carry them immediately over the piecewise convex curves.

Let us underline, that in the case of an interval I with a length greater than 2π, the operations on the piecewise convex curves cannot be determined by suitable operations on the generated parametric curves. To demonstrate this, observe that both the two piecewise convex curves Γ^i, $i = 1, 2$, in Example 3 generate the same parametric curve, namely the unit circle $r = (\cos s, \sin s)$, $0 \le s \le 2\pi$. The piecewise convex curve $\Gamma^1 + \Gamma^2$ generates the parametric curve $r = (\cos s, \sin s)$, $0 \le s \le 4\pi$, that is the unit circle circumscribed twice. At the same time the piecewise convex curves $\Gamma^1 + \Gamma^1 = 2\,\Gamma^1$ and $\Gamma^2 + \Gamma^2 = 2\,\Gamma^2$ generate the parametric curve $r = 2\,(\cos(s/2), \sin(s/2))$, $0 \le s \le 4\pi$, that is a circle with radius 2 circumscribed once.

Now we discuss closed piecewise convex curves. Such a curve $\Gamma = (a, I, s^+)$, where $I = [\theta_a, \theta_b]$, is called closed if its initial and final points coincide, that is if $a = c^+(\theta_b)$. According to (5) this condition can be written as

$$T^+ e_{\theta_a} s^+(\theta_a) + \int_{\theta_a}^{\theta_b} T^+ e_\lambda \, ds^+(\lambda) = 0 \,. \tag{12}$$

In general a closed piecewise convex curve Γ can have multiple points, that is it can generate a self-intersecting parametric curve. The next theorem shows, that this is not the case if the curve is determined by an interval, whose length is not greater than 2π, and it does not degenerate to a segment passed twice.

Theorem 8. *Let $\Gamma = (a, I, s^+)$ be a closed piecewise convex curve with interval $I = [\theta_a, \theta_b]$ having length $\theta_b - \theta_a \le 2\pi$. Then Γ generates and can be identified to a parametric curve being either a point, or a segment passed twice, or a convex curve.*

Proof. Using the notation from Theorem 7, we take the decomposition $\Gamma = \Gamma^1 \oplus \Gamma^2$, where $\theta_0 = \frac{1}{2}(\theta_a + \theta_b)$. Then for the respective parametric curves we have $\gamma = \gamma^1 \oplus \gamma^2$. Then Γ^1 can be identified with γ^1, whose initial and final points are a and $c^-(\theta_0)$. Similarly Γ^2 can be identified with γ^2, whose initial and final points are $c^-(\theta_0)$ and $c^+(\theta_b) = a$. In the case when $\theta_b - \theta_a < 2\pi$ according to Theorem 7 also Γ can be identified with γ. Then $\theta_b - \theta_0 = \theta_0 - \theta_a = \frac{1}{2}(\theta_b - \theta_a) < \pi$, and according to Theorem 4 γ^1 and γ^2 are convex arcs. If at least one of them is a point, the other is a point too, and γ degenerates to a point. If both $\gamma_a(\theta_0)$ and $\gamma_b(\theta_0)$ are segments, then obviously γ is a segment passed twice. Assume that at least one of γ^1 or γ^2 does not degenerate to a point and is not a segment. Then we can show that the set $\Phi = \Phi_{\gamma^1} \cup \Phi_{\gamma^2}$ is a convex figure. This can be demonstrated (similarly to the proof of Theorem 1 case c) given in [10]) by showing that γ is a simple closed curve, Φ contains interior points, and each straight line passing through an interior point of Φ intersects γ in exactly two points. This property implies that Φ is a convex figure (see Problem 5, page 17 in [12]).

The obtained result can be extended to the case $\theta_b - \theta_a = 2\pi$ in a routine way (though not ad hoc), so the proof is omitted. Let us underline, that according to Theorem 7, when $\theta_b - \theta_a = 2\pi$ and $\Gamma = (a, I, s^+)$ is such that

$s^+(\theta_b - 0) = s^+(\theta_b)$, then $\mathit{\Gamma}$ can be identified with γ. To show that also without this assumption $\mathit{\Gamma}$ can be identified with γ, we make the following reasoning. Suppose that γ is a simple closed curve (the cases of γ being a point, or a segment passed twice are considered similarly). Consider the piecewise convex curves $\mathit{\Gamma} = (a, I, s^+)$ and $\bar{\mathit{\Gamma}} = (c^-(\theta_b), I, \bar{s}^+)$, where $\bar{s}^+ : I \to \mathbb{R}$ is defined by

$$\bar{s}^+(\theta) = \begin{cases} s^+(\theta) + s^+(\theta_b) - s^+(\theta_b - 0)\,, & \theta_a \le \theta < \theta_b\,, \\ s^+(\theta_b)\,, & \theta = \theta_b\,. \end{cases}$$

Since $\bar{s}^+(\theta_b - 0) = \bar{s}^+(\theta_b)$, according to Theorem 7 the curve $\bar{\mathit{\Gamma}}$ can be identified with the generated by it parametric curve $\bar{\gamma}$. The curve $\bar{\gamma}$ coincides with γ as a set of points in the plane. The only difference is that $\mathit{\Gamma}$ and $\bar{\mathit{\Gamma}}$ determine two different initial points a and $c^-(\theta_b)$. Since we usually identify the coinciding as point sets simple closed curves regardless of the chosen initial points, we may identify γ and $\bar{\gamma}$. We will identify on this base also the curves $\mathit{\Gamma}$ and $\bar{\mathit{\Gamma}}$, and any two curves to which corresponds the same parametric curve. With this agreement, giving in fact a relation of equivalence on the piecewise convex curves defined on I, we get, that each closed piecewise convex curve defined on an interval with length 2π can be identified with the generated by it parametric curve. □

Let $\mathit{\Gamma} = (a, I, s^+)$ be a closed piecewise convex curve with interval $I = [\theta_a, \theta_b]$ having length $\theta_b - \theta_a \le 2\pi$. Denote by γ the parametric curve generated by $\mathit{\Gamma}$. We will put $\mathit{\Phi}_{\mathit{\Gamma}} = \operatorname{co}\gamma$ (in this formula we consider γ as a set of points). Then $\mathit{\Phi}_{\mathit{\Gamma}}$ is a compact convex set, which is either a convex figure having γ as its boundary, or a segment (then γ is the segment passed twice), or a point. Now $\mathit{\Gamma}$ can be identified with γ, which in turn is into one-to-one correspondence with $\mathit{\Phi}_{\mathit{\Gamma}}$. Therefore, in this case the introduced operations between piecewise convex curves can be interpreted in terms of operations between convex sets. For this purpose it is useful to establish what is the relation between the support functions of the given piecewise convex curves and the support function of their linear combination with nonnegative coefficients.

Theorem 9. *Let $\mathit{\Gamma}^i = (a^i, I, s_i^+)$, $i = 1, \dots, n$, be closed piecewise convex curves with the same interval $I = [\theta_a, \theta_b]$, and let $\lambda_i \ge 0$ be nonnegative reals. Then $\mathit{\Lambda}_{\sum_{i=1}^n \lambda_i \mathit{\Gamma}^i} = \sum_{i=1}^n \lambda_i \mathit{\Lambda}_{\mathit{\Gamma}^i}$. In consequence, if the interval I has length $\theta_b - \theta_a \le 2\pi$, then $\sum_{i=1}^n \lambda_i \mathit{\Gamma}^i$ is a closed piecewise convex curve and it holds $\mathit{\Phi}_{\sum_{i=1}^n \lambda_i \mathit{\Gamma}^i} = \sum_{i=1}^n \lambda_i \mathit{\Phi}_{\mathit{\Gamma}^i}$, where the right-hand side stands for the respective Minkowski operation between convex sets.*

Proof. We have $\sum_{i=1}^n \lambda_i \mathit{\Gamma}^i = (\sum_{i=1}^n \lambda_i a^i, I, \sum_{i=1}^n \lambda_i s_i^+)$. Applying the representation (5) now for $\theta \in [\theta_a, \theta_b]$ we get

$$\Lambda_{\sum_{i=1}^n \lambda_i \Gamma^i}(\theta) = e_\theta \cdot c^+_{\sum_{i=1}^n \lambda_i \Gamma^i}(\theta)$$
$$= e_\theta \cdot \left(\sum_{i=1}^n \lambda_i a^i + T^+ e_{\theta_a} \sum_{i=1}^n \lambda_i s_i^+(\theta_a) + \int_{\theta_a}^{\theta} T^+ e_\lambda \, d \sum_{i=1}^n \lambda_i s_i^+(\lambda) \right)$$
$$= \sum_{i=1}^n \lambda_i \; e_\theta \cdot c_i^+(\theta) = \sum_{i=1}^n \lambda_i \, \Lambda_{\Gamma^i}(\theta) \, .$$

If Γ^i are closed piecewise convex curves we have according to (12)

$$T^+ e_{\theta_a} \sum_{i=1}^n \lambda_i s_i^+(\theta_a) + \int_{\theta_a}^{\theta_b} T^+ e_\lambda \, d \sum_{i=1}^n \lambda_i s_i^+(\lambda)$$
$$= \sum_{i=1}^n \lambda_i \left(T^+ e_{\theta_a} \, s_i^+(\theta_a) + \int_{\theta_a}^{\theta_b} T^+ e_\lambda \, d s_i^+(\lambda) \right) = 0 \, ,$$

whence $\sum_{i=1}^n \lambda_i \Gamma^i$ is also a closed piecewise convex curve. When I has length $\theta_b - \theta_a \le 2\pi$ according to Theorem 8 the sum $\sum_{i=1}^n \lambda_i \Gamma^i$ can be identified with a simple closed curve being the boundary of the set $\Phi_{\sum_{i=1}^n \lambda_i \Gamma^i} = \operatorname{co} \sum_{i=1}^n \lambda_i \Gamma^i$. Let $\Gamma = (a, I, s^+)$ be a closed piecewise convex curve with interval $I = [\theta_a, \theta_b]$ having length $\theta_b - \theta_a \le 2\pi$. Then according to the definitions the support functions Λ_Γ of the piecewise convex curve Γ and Λ_{Φ_Γ} of the convex set $\Phi_\Gamma = \operatorname{co} \Gamma$ are equal. Now the equality $\Lambda_{\sum_{i=1}^n \lambda_i \Gamma^i} = \sum_{i=1}^n \lambda_i \Lambda_{\Gamma^i}$ implies

$$\Lambda_{\Phi_{\sum_{i=1}^n \lambda_i \Gamma^i}} = \sum_{i=1}^n \lambda_i \, \Lambda_{\Phi_{\Gamma^i}} = \Lambda_{\sum_{i=1}^n \lambda_i \, \Phi_{\Gamma^i}} \, ,$$

where the right-hand side equality represents a known relation between Minkowski operations of convex sets and their support functions. Since the support function determines uniquely the compact convex set, we have $\Phi_{\sum_{i=1}^n \lambda_i \Gamma^i} = \sum_{i=1}^n \lambda_i \Phi_{\Gamma^i}$. □

The Minkowski operations between sets define Minkowski operations between convex curves. Theorem 9 shows that the introduced here linear combination for piecewise convex curves $\sum_{i=1}^n \lambda_i \Gamma^i$, where $\lambda_i \ge 0$, is a generalization of the Minkowski operations from convex curves to piecewise convex curves.

The connection of the introduced operations with the Minkowski operations of sets is illustrated on the following example.

Example 5. Let $K_0 = \operatorname{co}\{(0, 1), (0, -1)\}$, $K_1 = \{r \in \mathbb{R}^2 \mid \|r\| \le 1\}$ and $K = K_0 + K_1$. Define the convex arc Γ to be the counter-clockwise oriented part of the boundary of K from the point $(1, 1)$ to the point $(1, -1)$. We have $a = (1, 1)$, $\gamma = 0$, and $s^-, s^+ : [0, 2\pi] \to \mathbb{R}$ are given by

$$s^-(\theta) = \begin{cases} \theta \, , & 0 \le \theta \le \pi \, , \\ 2 + \theta \, , & \pi < \theta \le 2\pi \, , \end{cases} \qquad s^+(\theta) = \begin{cases} \theta \, , & 0 \le \theta < \pi \, , \\ 2 + \theta \, , & \pi \le \theta \le 2\pi \, . \end{cases}$$

Formula (3) with $\theta_0 = 0$ gives

$$c^-(\theta) = \begin{cases} (\cos\theta,\ 1+\sin\theta), & 0 \le \theta \le \pi\,, \\ (\cos\theta,\ -1+\sin\theta), & \pi < \theta \le 2\pi\,, \end{cases}$$

$$c^+(\theta) = \begin{cases} (\cos\theta,\ 1+\sin\theta), & 0 \le \theta < \pi\,, \\ (\cos\theta,\ -1+\sin\theta), & \pi \le \theta \le 2\pi\,. \end{cases}$$

The convex arc Γ can be identified with the piecewise convex curve $\Gamma = \Gamma^1 + \Gamma^2$, where $\Gamma^i = (a^i,\ I,\ s_i^+)$, $i = 1,\ 2$, are given by $I = [0,\ 2\pi]$, $a^1 = (1,\ 0)$, $s_1^+(\theta) = \theta$ for $0 \le \theta \le 2\pi$, and $a^2 = (0,\ 1)$, $s_2^+(\theta) = 0$ for $0 \le \theta < \pi$, and $s_2^+(\theta) = 2$ for $\pi \le \theta \le 2\pi$.

4 Decomposition of the Measure

For two piecewise convex curves $\Gamma^i = (a^i, I, s^+)$, $i = 1,\ 2$, which differ only with respect to the initial points, formulae (5) give $c_2^+(\theta) - c_1^+(\theta) = a^2 - a^1$ and $c_2^-(\theta) - c_1^-(\theta) = a^2 - a^1$. Consequently, for the generated by Γ^i parametric curves $\gamma^i : r = f^i(s)$, $0 \le s \le L_i$, we have $L_1 = L_2$ and $f^2(s) = f^1(s) + (a^2 - a^1)$. Hence γ^2 is obtained translating γ^1 with the vector $a^2 - a^1$. On this base we can speak also that Γ^2 is obtained translating Γ^1 on the vector $a^2 - a^1$. All the piecewise convex curves obtained one from the other by a translation form a class of equivalence in the set of piecewise convex curves. Each class of equivalence is determined by a piecewise convex curve $\Gamma = (0,\ I,\ s^+)$ with initial point the origin. Wishing to study the properties of the piecewise convex curves determined only by s^+, we may confine to piecewise convex curves with initial points at the origin, and to state that these properties are assigned to each piecewise convex curve from the class of equivalence.

The function s^+ for the piecewise convex curve $\Gamma = (0,\ I,\ s^+)$ with $I = [\theta_a, \theta_b]$ is called sometimes measure. Such a saying is used once that s^+ measures the length in the generated parametric curve, but mainly since s^+ determines the measure in the Riemann-Stieltjes integrals (5). The measure s^+ is monotonely increasing function. Each such function admits a decomposition

$$s^+(\theta) = s_j^+(\theta) + s_a^+(\theta) + s_\sigma^+(\theta)\,, \quad \theta \in I = [\theta_a, \theta_b]\,, \tag{13}$$

as a sum of a jump function s_j^+, an absolutely continuous function s_a^+, and a singular function s_σ^+. For the role of this decomposition in integration theory see for instance [8] or [4]. The decomposition (13) is unique, if we agree that at the beginning of the interval it holds $s_a^+(\theta_a) = 0$ and $s_\sigma^+(\theta_a) = 0$. If s^+ is nonnegative, monotonely increasing and continuous from the right, the same properties obey the functions s_j^+, s_a^+ and s_σ^+. Therefore, together with the piecewise convex curve $\Gamma = (0,\ I,\ s^+)$, we determine uniquely the piecewise convex curves $\Gamma_j = (0,\ I,\ s_j^+)$, $\Gamma_a = (0,\ I,\ s_a^+)$ and $\Gamma_\sigma = (0,\ I,\ s_\sigma^+)$. Further, the equality (13) implies $\Gamma = \Gamma_j + \Gamma_a + \Gamma_\sigma$. We call this representation a decomposition of Γ into a sum of a jump, an absolutely continuous and a singular components.

The properties of Γ can be investigated on the base of its decomposition. But first of all let us underline as the next example shows, that even if Γ is a convex curve, its components need not be convex curves. Therefore, the notion of a piecewise convex curve is important also when we wish to study convex curves through their decomposition.

Example 6. Consider the piecewise convex curve $\Gamma = (0, I, s^+)$, $I = [0, 2\pi]$, given by

$$s^+(\theta) = \begin{cases} 1+\theta & 0 \le \theta < 3\pi/2\,, \\ 2+3\pi/2\,, & 3\pi/2 \le \theta \le 2\pi\,. \end{cases}$$

Then Γ is a closed piecewise convex curve, hence it is also a convex curve. It is decomposed into $\Gamma = \Gamma_j + \Gamma_a$ with non-closed components given by

$$s_j^+(\theta) = \begin{cases} 1\,, & 0 \le \theta < 3\pi/2\,, \\ 2\,, & 3\pi/2 \le \theta \le 2\pi\,, \end{cases} \qquad s_a^+(\theta) = \begin{cases} \theta\,, & 0 \le \theta < 3\pi/2\,, \\ 3\pi/2\,, & 3\pi/2 \le \theta \le 2\pi\,. \end{cases}$$

5 The Structure of Piecewise Convex Curves

Now we give some applications of the obtained in the previous section decomposition.

Consider the piecewise convex curve $\Gamma = (0, I, s^+)$ with $I = [\theta_a, \theta_b]$.

We say that Γ has a side for $\theta \in I$ if $\ell(\theta) = s^+(\theta) - s^-(\theta) > 0$. The number $\ell(\theta)$ is called then length of the side. Recall that $s^-(\theta) = 0$ for $\theta = \theta_a$ and $s^-(\theta) = s^+(\theta - 0)$ for $\theta_a < \theta \le \theta_b$. In fact Γ has a side for some θ if $c^-(\theta) \ne c^+(\theta)$. Then we accept that the segment $\overline{c^-(\theta)\, c^+(\theta)}$ is this side. The length of this segment is in fact $\ell(\theta)$. The direction e_θ is considered then as a normal for this side. Since s_a^+ and s_σ^+ are zero for $\theta = \theta_a$ and they are continuous for $\theta_a \le \theta \le \theta_b$, we see that Γ has a side for θ if and only if the jump component Γ_j has a side for θ and the lengths of these sides of Γ and Γ_j coincide.

Given any $\theta \in I$, we denote by $I(\theta) = [\alpha(\theta), \beta(\theta)]$ the maximal interval $I(\theta) \subset I$, such that $\theta \in I(\theta)$ and s^+ is a constant on $I(\theta) \setminus \{\beta(\theta)\} = [\alpha(\theta), \beta(\theta))$. We say that Γ has a vertex for $\theta \in I$ if the interval $I(\theta)$ is non-degenerate. Then the point $c^+(\lambda)$ is the same for all $\lambda \in I(\theta) \setminus \{\beta(\theta)\}$ and is called the vertex for θ. The number $m(\theta) = \beta(\theta) - \alpha(\theta)$ is called the measure of this angle. To underline the dependence of $I(\theta)$ on Γ we write also $I_\Gamma(\theta) = [\alpha_\Gamma(\theta), \beta_\Gamma(\theta))$. Since the functions s_j^+, s_a^+ and s_σ^+ are monotonely increasing, we see that $I(\theta) = I_{\Gamma_j}(\theta) \cap I_{\Gamma_a}(\theta) \cap I_{\Gamma_\sigma}(\theta)$. Therefore Γ has a vertex for θ if this intersection is a non-degenerate interval.

We call the piecewise convex curve Γ a broken line, if $s^+ = s_j^+$ and s_j^+ is a scale function. In other words Γ is a broken line if the interval I is a union of finitely many intervals on which s^+ is a constant. Sometimes the broken lines are called polygonal curves, but we use here the notion of a polygonal curve in a more general sense.

Recall that s^+ gives a measure on the interval I. In the case when Γ is a broken line the support of this measure is a finite set (we say also that the measure is concentrated on a finite set). This observation leads to the following generalization of the notion of a broken line. We say that Γ is a piecewise convex polygonal curve if the measure s^+ is concentrated on a set with Lebesgue measure zero. The following remarks concern polygonal curves. The definition gives immediately that a curve is polygonal if and only if the decomposition (13) does not contain absolutely continuous part, that is if $s_a^+ = 0$. Consequently $\Gamma = \Gamma_j + \Gamma_\sigma$ is decomposed in only jump curve and singular curve. The existence of an absolutely continuous component Γ_a of Γ is connected with the notion of a curvature of the piecewise convex curve and will be investigated in a separate paper. Now we confine to some notes concerning polygonal curves.
If Γ is a broken line and s^+ is not a constant, then Γ has both sides and vertices. Asking, whether the same property is true for polygonal curves, we discover, that there exist polygonal curves with infinitely many sides and no vertices or with infinitely many vertices and no sides.

Example 7. Let $I = [0, \pi/2]$ and $\{\theta_n\}_{n=1}^\infty$ be a dense sequence of different numbers of the open interval $(0, \pi/2)$ (say $\{\theta_n\}_{n=1}^\infty$ could be the set of all the rational points of this interval). Let $\{\varepsilon_n\}_{n=1}^\infty$ be a sequence of positive numbers for which the series $\sum_{n=1}^\infty \varepsilon_n$ converges. Define the function $s^+ : I \to \mathbb{R}$, $s^+(\theta) = \sum\{\varepsilon_n \mid \theta_n \le \theta\}$. Then s^+ is nonnegative, monotonely increasing and continuous from the right jump function, which is not a constant on any interval. Then the piecewise convex curve $\Gamma = (0, I, s^+)$ has only a jump component $\Gamma = \Gamma_j$ having a side at each point θ_n, $n = 1, 2, \ldots$, and having no vertices at all.

Example 8. Let s^+ be any singular function on the interval $I = [0, \pi/2]$ and let $\{I_n\}_{n=1}^\infty$ be the subintervals of I on which s^+ is constant. Recall that a singular function has no points of discontinuity. Then the piecewise convex curve $\Gamma = (0, I, s^+)$ has only a singular component $\Gamma = \Gamma_\sigma$ having a vertex at each $\theta \in I_i$ and no sides at all.

A natural question is whether there exist convex curves being polygonal curves with the properties described in the previous two examples. The following example gives an affirmative answer of this question.

Example 9. Let $\hat{s} : [0, \pi/2] \to \mathbb{R}$ be any nonnegative, monotonely increasing and continuous from the left function, such that $\hat{s}^+(0) = 0$ and $\hat{s}^+(\pi/2 - 0) = \hat{s}^+(\pi/2)$. Put as usual $\hat{s}^-(0) = 0$ and $\hat{s}^-(\theta) = \hat{s}^+(\theta - 0)$ for $0 < \theta \le \pi/2$. Let $I = [0, 2\pi]$ and define the function $s^+ : I \to \mathbb{R}$ by

$$s^+(\theta) = \begin{cases} \hat{s}^+(\theta)\,, & 0 \le \theta < \pi/2\,, \\ 2\,\hat{s}^+(\pi/2) - \hat{s}^-(\pi - \theta)\,, & \pi/2 \le \theta < \pi\,, \\ 2\,\hat{s}^+(\pi/2) + \hat{s}^+(\theta - \pi)\,, & \pi \le \theta < 3\pi/2\,, \\ 4\,\hat{s}^+(\pi/2) - \hat{s}^-(2\pi - \theta)\,, & 3\pi/2 \le \theta \le 2\pi\,. \end{cases}$$

The function s^+ is nonnegative, monotonely increasing and continuous from the right. Consider the piecewise convex curve $\Gamma = (0,\, I,\, s^+)$. Then condition (12) is satisfied, whence Γ is closed and according to Theorem 8 it is a convex curve. When for $\hat{s}^+$ we take the function from Example 7 we get a convex curve being a polygonal curve with infinitely many sides and no vertices, and when we take for $\hat{s}^+$ the function from Example 8 we get a convex curve being a polygonal curve with infinitely many vertices and no sides.

To check condition (12) in the last example, we apply substitution in the integrals getting

$$T^+e_0\, s^+(0) + \int_0^{2\pi} T^+e_\lambda\, ds^+(\lambda) = \int_0^{\pi/2} T^+e_\lambda\, d\hat{s}^+(\lambda)$$

$$- \int_{\pi/2}^{\pi} T^+e_\lambda\, d\hat{s}^-(\pi-\lambda) + \int_{\pi}^{3\pi/2} T^+e_\lambda\, d\hat{s}^+(\lambda-\pi) - \int_{3\pi/2}^{2\pi} T^+e_\lambda\, d\hat{s}^-(2\pi-\lambda)$$

$$= \int_0^{\pi/2} \left(T^+e_\lambda + T^+e_{\pi-\lambda} + T^+e_{\lambda+\pi} + T^+e_{2\pi-\lambda}\right)\, d\hat{s}^+(\lambda) = 0\,.$$

The last integral is zero, since $T^+e_\lambda + T^+e_{\pi-\lambda} + T^+e_{\lambda+\pi} + T^+e_{2\pi-\lambda} = 0$.

References

1. Blaschke W (1956) Kreis und Kugel. Walter de Gruyter & Co, Berlin
2. Kenderov PS (1983) Polygonal approximation of plane convex compacts. J. Approx. Theory 38:221–239
3. Kenderov PS, Kirov N (1993) A dynamical systems approach to the polygonal approximation of plane convex compacts. J. Approx. Theory 74:1–15
4. Kolmogorov AN, Fomin SV (1972) Elements of the theory of functions and of functional analysis (In Russian). Nauka, Moscow
5. Ligun AA, Shumeĭko AA (2000) Description of convex curves. Ukrainian Math. Zh. 52:1040–1051
6. Ludwig M (1998) Asymptotic approximation of convex curves: the Hausdorff metric case. Arch. Math. (Basel) 70:331–336
7. McClure DE, Vitale RA (1975) Polygonal approximation of plane convex bodies. J. Math. Anal. Appl. 51:326–358
8. Natanson IP (1950) Theory of the functions of a real variable (In Russian). Moscow–Leningrad
9. Nedelcheva MD (1985) Characterization of convex subsets of the plane through their local approximation properties. Serdica Math. J. 11:165–170
10. Nedelcheva MD (2006) Piecewise convex curves and their integral representation. Serdica Math. J. 32:7–30
11. Vitale RA (1979) Support functions of plane convex sets. Report, Claremont Graduate School, Claremont
12. Yaglom IM, Boltyanski VG (1981) Convex figures (In Russian). Gos. Izdatel'stvo Tehnicheskoy Literatury, Moskva Leningrad

Rambling Through Local Versions of Generalized Convex Functions and Generalized Monotone Operators

Huynh Van Ngai and Jean-Paul Penot

Université de Pau, Faculté des Sciences, Laboratoire de Mathématiques Appliquées, CNRS UMR 5142 Av. de l'Université 64000 PAU, France.

Summary. Two classes of functions encompassing the cone of convex functions and the space of strictly differentiable functions are presented and compared. Related properties for sets and multimappings are dealt with.

Key words: Approximate convexity, nonsmooth analysis, paraconvexity, subdifferentials.

1 Introduction

In the present paper we survey some generalized convexity properties of functions and sets and we characterize them in terms of generalized monotonicity properties. They are of a different nature than the generalized convexity properties considered up to now in the series of meetings devoted to the subject. They are local (or rather infinitesimal) rather than global. Thus, one cannot expect from them global optimality conditions or duality properties. However, they can be combined with the usual generalized convexity and generalized monotonicity properties and they are important from the point of view of nonsmooth analysis. For the proofs, and a more detailed analysis, we refer to [41], [42], [43], [44] and their references. We introduce here some new concepts related to the two classes we study, in particular a notion of approximately affine map and a variant of it and a notion of approximately multimapping. We also raise some open problems.

Motivated by various problems, several authors have introduced some favorable classes of functions on normed vector spaces (in short, n.v.s.). Let us mention a few, referring to the papers in the bibliography for precise definitions.

- p-paraconvex functions introduced by Rolewicz [55]-[59] and studied by Bougeard [9], Bougeard-Penot-Pommellet [10], Canino [11], Castellani-

Pappalardo [13], Jourani [30], Ngai-Penot [44], Penot [47], Penot-Volle [49]...

- in the case $p = 2$ these functions are also called semiconvex (Lasry-Lions [32], Attouch-Azé [1], Cannarsa-Sinestrari [12]...) or subsmooth (Aussel-Daniilidis-Thibault [3]), property (ω) (Colombo-Goncharov [16]), weakly convex (Vial [63]), lower-C^2 (Rockafellar [53], Spingarn [62], Penot [47]...);
- (p,q)-convex functions introduced by De Giorgi-Marino-Tosques [20] and studied by Canino [11], Degiovanni [21], Marino [34], [35] and their co-authors;
- Lower-C^1 functions introduced by Spingarn [62] and studied by Rockafellar [53], Penot [47], Daniilidis-Georgiev-Penot [19];
- Lower-C^k functions and lower-T^k functions ($k \in \mathbb{N}\backslash\{0,1\}$) studied by Rockafellar [53] and Penot [47];
- approximately convex functions introduced by Ngai-Luc-Théra [40] and studied by Aussel-Daniilidis-Thibault [3], Colombo-Goncharov [16], Daniilidis - Georgiev [18], Ngai-Penot [42];
- approximately starshaped functions introduced by Penot [48]
- semismooth functions introduced by Mifflin [38] in the locally Lipschitz case and Ngai-Penot in the ower semicontinuous (l.s.c.) case [41].
- prox-regular functions considered in [6], [7], [8], [50], [51].

These concepts have directional versions which will not be considered here. In the present survey, we will focus attention on the main streams of these classes, referring to the quoted papers for more specialized properties. We will also relate these classes of functions to some classes of sets which have some regularity properties. It is one of the most remarkable achievements of nonsmooth analysis to enable easy (or at least natural) passages from functions to sets and from sets to functions. Therefore, we hope that the present tentative of synthesis will prove useful to the reader.

2 Some Concepts of Nonsmooth Analysis

Our study requires some knowledge of nonsmooth analysis. We gather these elements in the present section for the reader's convenience. Of course, the limited setting of the present contribution imposes conciseness. The notation we use ($\partial^{(\cdot)}, N^{(\cdot)}$, $f^{(\cdot)}$, $T^{(\cdot)}$) stresses the fact that several devices exist, each of which being affected by some letter or symbol and the fact that for a given device $(\cdot)$ one disposes of related constructions for subdifferentials, normal cones and, sometimes, generalized derivatives of functions and tangent cones to sets. Letters usually refer either to some author or to some characteristic feature of the construction. Both denominations are debattable. For instance both Severi and Bouligand worked on tangent cones, while Dini, Fréchet, Hadamard have not touched subdifferentials. Here we distinguish directional notions from firm notions which require stronger, uniform estimates; but other

names exist with some evocative character (adjacent or incident cones, contingent cones...). Symbols are less subject to such wanderings.
One of the key features of the studies concerning the two classes we deal with consists in the observation that the different devices we mentioned yield the same objects within these two classes. This fact explains why we do not look for completeness. This remarkable fact also enables to combine the advantages of the various devices.

2.1 Subdifferentials

Given a subset $\mathcal{F}(X)$ of the set $\mathcal{S}(X)$ of lower semicontinuous (lsc for short) functions $f : X \to \mathbb{R} \cup \{+\infty\}$ on a nvs X with topological dual X^*, we define here a *subdifferential* as a correspondence $\partial^{(\cdot)} : \mathcal{F}(X) \times X \rightrightarrows X^*$ satisfying:

- $\partial^{(\cdot)} f(\overline{x}) = \varnothing$ when $\overline{x} \notin \operatorname{dom} f$ (i.e. when $f(\overline{x}) = +\infty$)
- $0 \in \partial^{(\cdot)} f(\overline{x})$ when $\overline{x}$ is a minimizer of a Lipschitzian function $f \in \mathcal{F}(X)$.

Such conditions are versatile, but loose requirements; they are often supplemented by other conditions:

- if $\partial^{(\cdot)} f(x) \neq \varnothing$ for x in a dense subset of X for any $f \in \mathcal{L}(X)$ (the set of Lipschitzian functions) X is said to be a $\partial^{(\cdot)}$-*subdifferentiability space.*
- (*Exact mean value theorem*) $\partial^{(\cdot)}$ is said to be *Lipschitz-valuable* on X if for any $\overline{x}, \overline{y} \in X$, any $f \in \mathcal{L}(X)$ one can find $w \in [\overline{x}, \overline{y}]$ and $w^* \in \partial^{(\cdot)} f(w)$ such that $f(y) - f(x) = \langle w^*, y - x \rangle$.
- (*Fuzzy mean value theorem*) $\partial^{(\cdot)}$ is said to be *valuable* on X if for any $\overline{x} \in X$, $\overline{y} \in X \backslash \{\overline{x}\}$, $f \in \mathcal{F}(X)$ finite at $\overline{x} \in X$ and for any $r \in \mathbb{R}$ such that $f(\overline{y}) \geq r$, there exist $u \in [\overline{x}, \overline{y})$ and sequences $(u_n) \to u$, (u_n^*) such that $u_n^* \in \partial^{(\cdot)} f(u_n)$, $(f(u_n)) \to f(u)$,

$$\lim_n \|u_n^*\| \, d(u_n, [\overline{x}, \overline{y}]) = 0,$$
$$\liminf_n \langle u_n^*, \overline{y} - \overline{x} \rangle \geq r - f(\overline{x}),$$
$$\liminf_n \langle u_n^*, (x - u_n) \,/\, \|x - u\| \rangle \geq (r - f(\overline{x})) \,/\, \|\overline{y} - \overline{x}\|$$

  for all $x \in (\overline{x} + \mathbb{R}_+(\overline{y} - \overline{x})) \backslash [\overline{x}, u)$.

The terms “valuable”, “Lipschitz-valuable” evoke the Mean Value theorem; but, in view of the numerous applications of this theorem, it also qualifies a subdifferential which may be useful for several purposes. In several cases, such a property for a specific subdifferential is valid only in a restrictive class of spaces (for instance Asplund spaces for subdifferentials larger then the Fréchet subdifferential); the Clarke subdifferential is valuable in any Banach space, but it is not as accurate as the Fréchet subdifferential. These two examples are defined in the next subsection with few other ones.

2.2 Some Subdifferentials

In this section we describe some examples of subdifferentials and some basic constructions. Several subdifferentials are derived from a directional derivative $f^{(\cdot)}$ of some sort of the function f via the following relation

$$\partial^{(\cdot)} f(x) := \{x^* \in X^* : x^*(\cdot) \leq f^{(\cdot)}(x, \cdot)\}.$$

- The (lower) directional derivative (or lower Hadamard derivative) of f

$$f'(x, v) := \liminf_{(t,w)\to(0_+,v)} \frac{f(x+tw) - f(x)}{t}.$$

- The Clarke–Rockafellar derivative [52] or circa-derivative:

$$f^{\uparrow}(x, v) := \inf_{r>0} \limsup_{\substack{(t,y)\to(0_+,x) \\ f(y)\to f(x)}} \inf_{w\in B(v,r)} \frac{f(y+tw) - f(y)}{t}.$$

- The dag derivative which majorizes both these two derivatives is given by

$$f^{\dagger}(x, v) := \limsup_{(t,y)\to(0_+,x), f(y)\to f(x)} \frac{f(y + t(v + x - y)) - f(y)}{t}.$$

  Its interest seems to be limited to its role of upper bound; however, this role could be played by $f^{\uparrow}$ which is already extremely large.

Other generalized derivatives exist, such as the adjacent (or incident or intermediate) derivative ([2]), the moderate (or Michel-Penot) derivative ([36], [37]), but they will not be used here. The subdifferentials associated with f', $f^{\uparrow}$, $f^{\diamond}$, $f^{\dagger}$ are denoted by ∂, $\partial^{\uparrow}$, $\partial^{\diamond}$, $\partial^{\dagger}$ respectively.
Several subdifferentials are not derived from directional derivatives:

- The *firm (or Fréchet) subdifferential*:

  $$x^* \in \partial^- f(x) \Leftrightarrow \forall\varepsilon > 0\, \exists\delta > 0, \forall u \in \delta\overline{B}_X : f(x+u) \geq f(x)+\langle x^*, u\rangle-\varepsilon \|u\| .$$

- The p-proximal subdifferential, with $p \in (1, 2]$:

  $$x^* \in \partial^p f(x) \Leftrightarrow \exists c > 0, \rho > 0, \forall u \in \rho\overline{B}_X : f(x+u) \geq f(x)+\langle x^*, u\rangle-c \|u\|^p .$$

- The *approximate subdifferential* or Ioffe subdifferential ([25], [26], [27], [28]).
- The *limiting subdifferential* associated with a subdifferential $\partial^{(\cdot)}$:

$$\overline{\partial^{(\cdot)}} f(x) := w^* - \limsup_{(u,f(u))\to(x,f(x))} \partial^{(\cdot)} f(u),$$

  where the w^*-limsup is the set of cluster points of bounded nets $(u_i^*)_{i\in I}$ with $u_i^* \in \partial^{(\cdot)} f(u_i)$, $(u_i)_{i\in I} \to x$, $(f(u_i))_{i\in I} \to f(x)$.

2.3 Normal Cones

Since we will consider some generalized convexity properties of sets, we need to introduce some geometric concepts. With any subdifferential $\partial^{(\cdot)}$ is associated a notion of normal cone to $E \subset X$ at $e \in E$:

$$N^{(\cdot)}(E, e) := \mathbb{R}_+ \partial^{(\cdot)} \iota_E(e),$$

where ι_E is the indicator function of E ($\iota_E(x) = 0$ if $x \in E$, $+\infty$ otherwise). Conversely, a normal cone notion $N^{(\cdot)}$ yields a notion of subdifferential $\partial^{(\cdot)}$:

$$\partial^{(\cdot)} f(x) := \{x^* \in X^* : (x^*, -1) \in N^{(\cdot)}(E_f, x_f)\}$$

where $E_f := \{(x, r) \in X \times \mathbb{R} : r \geq f(x)\}$ is the epigraph of f and $x_f := (x, f(x))$.
Given a normal cone notion $N^{(\cdot)}$ one defines the associated coderivative of a multimapping $F : X \rightrightarrows Y$ between two normed vector spaces by

$$D^{(\cdot)} F(x, y)(y^*) := \{x^* \in X^* : (x^*, -y^*) \in N^{(\cdot)}(F, (x, y))\},$$

F being identified with its graph. In particular, for a function f on X, one has $\partial^{(\cdot)} f(x) = D^{(\cdot)} f(x, f(x))(1)$.
Examples:

- The (usual or contingent or directional) *normal cone* $N(E, x)$ to a subset E of X at $x \in \mathrm{cl}(E)$ is the polar cone of the tangent cone $T(E, x)$ to E at x which is the set of vectors $v \in X$ such that there exist sequences $(t_n) \to 0_+$, $(x_n) \xrightarrow{E} x$ for which $(t_n^{-1}(x_n - x)) \to v$.
- The *firm normal cone* (or Fréchet normal cone) to E at x is given by

  $$x^* \in N^-(E, x) \Leftrightarrow \forall \varepsilon > 0 \exists \delta > 0, \forall u \in E \cap B(x, \delta) : \langle x^*, u - x \rangle \leq \varepsilon \|u - x\| .$$

- The *Clarke normal cone* $N^\uparrow(E, x)$ to a subset E of X at $x \in \mathrm{cl}(E)$ is the polar cone of the Clarke tangent cone $T^\uparrow(E, x)$ to E at x which is the set of vectors $v \in X$ such that for any sequence $(x_n) \xrightarrow{E} x$ there exist sequences $(t_n) \to 0_+$ and (y_n) in E for which $(t_n^{-1}(y_n - x_n)) \to v$.
- The *limiting normal cone* associated with a normal cone $N^{(\cdot)}$:

  $$\overline{N^{(\cdot)}}(E, x) := w^* - \limsup_{(u, f(u)) \to (x, f(x))} N^{(\cdot)}(E, u),$$

  where the w^*-limsup is the set of cluster points of bounded nets $(u_i^*)_{i \in I}$ with $u_i^* \in \partial^{(\cdot)} N(E, u_i)$, $(u_i)_{i \in I} \xrightarrow{E} x$, i.e. $(u_i)_{i \in I} \to x$ with $u_i \in E$ for each $i \in I$.

3 Approximately Convex Functions and Paraconvex Functions

The following class of functions has been introduced by Ngai-Luc-Théra [40].

Definition 1. *A function $f : X \to \mathbb{R} \cup \{+\infty\}$ is said to be approximately convex around $\overline{x} \in X$ if for any $\varepsilon > 0$ there exists $\delta > 0$ such that for any $x, x' \in B(\overline{x}, \delta)$ and any $t \in [0,1]$ one has*

$$f(tx + (1-t)x') \leq tf(x) + (1-t)f(x') + \varepsilon t(1-t) \|x - x'\| .$$

It is a rather general class: obviously, any convex function and any function which is strictly differentiable at $\overline{x}$ is approximately convex around $\overline{x} \in X$. In view of Proposition 1 below, various combinations of such functions are approximately convex around $\overline{x}$. It can be shown that approximately convex functions retain some of the nice properties of convex functions. In particular they are continuous on segments contained in their domains and have radial derivatives (c.f. Ngai-Luc-Théra [40]). We show elsewhere ([45]) that approximately convex functions on Asplund spaces are densely differentiable as are convex functions.
An important subclass of the class of approximately convex functions is the class of p-paraconvex functions when $p > 1$.

Definition 2. *Given some $p > 1$, a function $f : X \to \mathbb{R}\cup\{+\infty\}$ on a n.v.s. X is said to be p-paraconvex around $\overline{x} \in \operatorname{dom} f := f^{-1}(\mathbb{R})$ if there exist $c, \delta > 0$ such that for any $x, x' \in B(\overline{x}, \delta)$ and any $t \in [0,1]$ one has*

$$f(tx + (1-t)x') \leq tf(x) + (1-t)f(x') + ct(1-t) \|x - x'\|^p . \qquad (1)$$

These two classes have interesting stability properties, as shown in Ngai-Luc-Théra [40] in the case of approximately convex functions.

Proposition 1. *The set of approximately convex (resp. p-paraconvex) functions around $\overline{x} \in X$ is stable (i.e. invariant) under addition, multiplication by positive numbers and finite suprema.*

The proofs of these assertions are simple. As an example, we give the proof for the supremum f of a finite family $(f_i)_{i \in I}$ of p-paraconvex around $\overline{x}$ functions. For $i \in I$, let $c_i, \delta_i > 0$ be such that for any $x, x' \in B(\overline{x}, \delta)$ and any $t \in [0,1]$ one has

$$f_i(tx + (1-t)x') \leq tf_i(x) + (1-t)f_i(x') + c_i t(1-t) \|x - x'\|^p .$$

Set $c := \max_{i \in I} c_i$, $\delta := \min_{i \in I} \delta_i$. Given $x, x' \in B(\overline{x}, \delta)$, $t \in [0,1]$, we pick $i \in I$ such that $f_i(tx + (1-t)x') = f(tx + (1-t)x')$. Then, replacing $f_i(x)$ and $f_i(x')$ by their respective majorants $f(x)$ and $f(x')$ in the preceding inequality, we get relation (1).
A stability result by composition with strictly differentiable mappings in [40] can be extended by using the following new concept.

Definition 3. *Given normed vector spaces X, Y, a mapping $g : X \to Y$ is said to be approximately affine around $\overline{x} \in X$ if for any $\varepsilon > 0$ there exists $\delta > 0$ such that for any $x, x' \in B(\overline{x}, \delta)$ and any $t \in [0,1]$ one has*

$$\|g(tx + (1-t)x') - tg(x) - (1-t)g(x')\| \leq \varepsilon t(1-t) \|x - x'\| .$$

It is p-para-affine around $\overline{x}$ if there exist $c, \rho > 0$ such that for any $x, x' \in B(\overline{x}, \delta)$ and any $t \in [0,1]$ one has

$$\|g(tx + (1-t)x') - tg(x) - (1-t)g(x')\| \leq ct(1-t) \|x - x'\|^p .$$

Clearly, a function $f : X \to \mathbb{R}$ is approximately affine around $\overline{x} \in X$ if and only if it is both approximately convex and approximately concave (i.e. $-f$ is approximately convex) around $\overline{x}$. A similar assertion holds for p-para-affine functions. Moreover, an easy use of the Hahn-Banach theorem yields a chacterization of approximately affine maps in terms of approximately affine functions.

Proposition 2. *For a mapping $g : X \to Y$ between two normed vector spaces X, Y, and $\overline{x} \in X$, the following assertions are equivalent:*
(a) g is approximately affine around $\overline{x}$;
(b) for any continuous linear form f on Y, $f \circ g$ is approximately affine around $\overline{x}$;
(c) for any continuous linear form f on Y, $f \circ g$ is approximately convex around $\overline{x}$.

A similar statement holds for a p-para-affine around $\overline{x}$ map g.

Proposition 3. *(a) A mapping $g : X \to Y$ which is strictly differentiable at $\overline{x}$ is approximately affine around $\overline{x}$.*
(b) A mapping $g : X \to Y$ of class C^2 around $\overline{x}$ is 2-para-affine around $\overline{x}$.

Proof. (a) Given $\varepsilon > 0$, let $\delta > 0$ be such that for every $x, x' \in B(\overline{x}, \delta)$ one has

$$\|g(x) - g(x') - Dg(\overline{x})(x - x')\| \leq \varepsilon \|x - x'\| ;$$

taking $x, x' \in B(\overline{x}, \delta)$ and $t \in [0,1]$ one has $x_t := tx + (1-t)x' \in B(\overline{x}, \delta)$ and, since $x_t - x = (1-t)(x' - x)$, $x_t - x' = t(x - x')$,

$$\begin{aligned}
&\|g(x_t) - tg(x) - (1-t)g(x')\| \leq t \|g(x_t) - g(x) - (1-t)Dg(\overline{x})(x' - x)\| \\
&+(1-t) \|g(x_t) - g(x') - tDg(\overline{x})(x - x')\| \\
&\leq 2\varepsilon t(1-t) \|x - x'\| .
\end{aligned}$$

(b) Let $g : X \to Y$ be of class C^2 around $\overline{x}$ and let $c > \left\|D^2 g(\overline{x})\right\|$. For $x, x' \in B(\overline{x}, \delta)$ with $\delta > 0$ small enough, one has a $\left\|D^2 g(x)\right\| \leq c$ for all $x \in B(\overline{x}, \delta)$ and a Taylor's expansion yields

$$\begin{aligned}
&\|tg(x) + (1-t)g(x') - g(x_t)\| \leq t\,\|g(x) - g(x_t) - (1-t)Dg(x_t)(x - x')\| \\
&+(1-t)\,\|g(x') - g(x_t) - tDg(x_t)(x' - x)\| \\
&\leq (1/2)t(1-t)^2 c\,\|x - x'\|^2 + (1/2)(1-t)t^2 c\,\|x' - x\|^2 \\
&\leq (1/2)t(1-t)c\,\|x - x'\|^2 .
\end{aligned}$$

Proposition 4. *Let $f = h \circ g$, where g and h are Lipschitzian around $\overline{x}$ and $\overline{y} := g(\overline{x})$ respectively.*
(a) If $g : X \to Y$ is approximately affine around $\overline{x} \in X$ and $h : Y \to \mathbb{R}$ is approximately convex around $\overline{y}$ then f is approximately convex around $\overline{x}$.
(b) If g and h are p-para-affine and p-paraconvex respectively around $\overline{x}$ and $\overline{y}$ respectively, then f is p-paraconvex.

Proof. (a) Let κ, $\rho > 0$ be such that g is Lipschitzian with rate κ on $B(\overline{x}, \rho)$ and let λ, $\sigma > 0$ be such that h is Lipschitzian with rate λ on $B(\overline{y}, \sigma)$. Given $\varepsilon > 0$, let $\delta \in (0, \sigma)$ be such that for any $y, y' \in B(\overline{y}, \delta)$ and any $t \in [0, 1]$ one has

$$h(ty + (1-t)y') \leq th(y) + (1-t)h(y') + (\varepsilon/2\kappa)t(1-t)\,\|y - y'\| .$$

Let $\gamma \in (0, \rho)$ be such that $g(B(\overline{x}, \gamma)) \subset B(\overline{y}, \delta)$ and such that for every $x, x' \in B(\overline{x}, \gamma)$, $t \in [0, 1]$

$$\|g(tx + (1-t)x') - tg(x) - (1-t)g(x')\| \leq (\varepsilon/2\lambda)t(1-t)\,\|x - x'\| .$$

Then, for $x, x' \in B(\overline{x}, \gamma)$, $t \in [0, 1]$, setting $y := g(x)$, $y' := g(x')$, one has

$$\begin{aligned}
f(tx + (1-t)x') &\leq h(ty + (1-t)y') \\
&\quad + \lambda\,\|g(tx + (1-t)x') - tg(x) - (1-t)g(x')\| \\
&\leq th(y) + (1-t)h(y') + (\varepsilon/2\kappa)t(1-t)\,\|y - y'\| \\
&\quad + (\varepsilon/2)t(1-t)\,\|x - x'\| \\
&\leq tf(x) + (1-t)f(x') + \varepsilon t(1-t)\,\|x - x'\| .
\end{aligned}$$

(b) The proof of the second assertion is analogous and left to the reader.

4 Generalized Convexity Versus Generalized Monotonicity

We devote this section to characterizations of approximate convexity and paraconvexity. Previous results of this kind have been obtained in Aussel-Daniilidis-Thibault [3], [16], Daniilidis-Georgiev [18] under some restrictive assumptions on the space or on the functions (a local Lipschitz property, for instance). They use concepts introduced by Spingarn (under the names of strict submonotonicity and hypomonotonicity respectively).

Definition 4. *A multimapping* $M : X \rightrightarrows X^*$ *is approximately monotone around* $\overline{x} \in \mathrm{dom}(M)$ *provided that for each* $\varepsilon > 0$ *there exists* $\rho > 0$ *such that*

$$\forall x_i \in B(\overline{x}, \rho),\ x_i^* \in M(x_i),\ i = 1,2 \qquad \langle x_1^* - x_2^*, x_1 - x_2 \rangle \geq -\varepsilon \|x_1 - x_2\|$$

Definition 5. *A multimapping* $M : X \rightrightarrows X^*$ *is said to be p-paramonotone around* $\overline{x}$ *on a subset* E *of* X *if there exist some* $m, \delta > 0$ *such that for any* $x_1, x_2 \in E \cap B(\overline{x}, \delta)$, $x_1^* \in M(x_1)$, $x_2^* \in M(x_2)$ *one has*

$$\langle x_1^* - x_2^*, x_1 - x_2 \rangle \geq -m \|x_1 - x_2\|^p .$$

For $E = X$ *one simply says that* M *is p-paramonotone around* $\overline{x}$.

In the characterization which follows we denote by S_X the unit sphere of X and by $B(\overline{x}, \rho)$ the open ball with center $\overline{x}$ and radius ρ.

Theorem 1. *Let* $\overline{x} \in \mathrm{dom}\, f$, f *l.s.c. and let* $\partial^{(\cdot)}$ *be a subdifferential. Suppose* $\partial^{(\cdot)} f \subset \partial^{\dagger} f$. *Then, among the following assertions, one has the implications (a)⇒(b)⇒(c)⇔(c')⇒(d).*
If moreover $\partial^{(\cdot)}$ *is valuable on* X, *all these assertions are equivalent.*
(a) f *is approximately convex around* $\overline{x}$;
(b) $\forall \varepsilon > 0\ \exists \rho > 0$ *such that* $\forall x \in B(\overline{x}, \rho)$, $\forall v \in B(0, \rho)$ *one has*

$$f^{\dagger}(x, v) \leq f(x + v) - f(x) + \varepsilon \|v\| ;$$

(c) $\forall \varepsilon > 0$, $\exists \rho > 0$ *such that* $\forall x \in B(\overline{x}, \rho), x^* \in \partial^{(\cdot)} f(x), (u, t) \in S_X \times (0, \rho)$ *one has*

$$\langle x^*, u \rangle \leq \frac{f(x + tu) - f(x)}{t} + \varepsilon;$$

(c') $\forall \varepsilon > 0\ \exists \rho > 0$ *such that* $\forall x \in B(\overline{x}, \rho)$, $\forall x^* \in \partial^{(\cdot)} f(x)$, $\forall v \in \overline{B}(0, \rho)$ *one has*

$$\langle x^*, v \rangle \leq f(x + v) - f(x) + \varepsilon \|v\| ;$$

(d) $\partial^{(\cdot)} f$ *is approximately monotone around* $\overline{x}$.

The implications are easy consequences of the definitions, except the implication (d)⇒(a); for f locally Lipschitzian it suffices to use the exact mean value theorem (see [3]). In the case of a l.s.c. function, the fuzzy mean value theorem is required ([42]).

Corollary 1. *The preceding assertions (a), (b), (c), (d) are equivalent when*
(i) X *is an arbitrary Banach space and* ∂ *is the Clarke or the Ioffe subdifferential;*
(ii) X *is an Asplund space and* ∂ *is the Fréchet subdifferential or the Hadamard subdifferential.*
Moreover, they are equivalent to the variant of assertion (b) obtained by replacing $f^{\dagger}$ *by* $f^{\uparrow}$ *(and, if* X *is an Asplund space, by* $f^{!}$*).*

Now let us turn to characterizations of paraconvexity.

Theorem 2. *Let $p \in [1, \infty)$, $f \in \mathcal{F}(X)$, $\overline{x} \in \operatorname{dom} f$. Suppose $\partial^{(\cdot)} f \subset \partial^{\dagger} f$. Then, among the following assertions, one has the implications (a)⇒(b)⇒(c)⇒(d).*
If moreover $\partial^{(\cdot)}$ is valuable, all these assertions are equivalent.
(a) f is p-paraconvex around $\overline{x}$;
(b) $\exists \rho > 0, c \in \mathbb{R}$, s. t. $\forall x \in B(\overline{x}, \rho) \cap \operatorname{dom} f$, $\forall v \in B(0, \rho)$ one has

$$f^{\dagger}(x, v) \leq f(x + v) - f(x) + c \|v\|^p ;$$

(c) $\exists \rho > 0, c \in \mathbb{R}$, such that $\forall x \in B(\overline{x}, \rho)$, $x^* \in \partial^{(\cdot)} f(x)$, $v \in B(0, \rho)$ one has

$$\langle x^*, v \rangle \leq f(x + v) - f(x) + c \|v\|^p ;$$

(d) $\partial^{(\cdot)} f$ is p-paramonotone around $\overline{x}$.

Corollary 2. *Suppose $f \in \mathcal{F}(X)$ is finite at $\overline{x} \in X$ and p-paraconvex around $\overline{x}$. Then for any subdifferential $\partial^{(\cdot)}$ such that $\partial^p f \subset \partial^{(\cdot)} f \subset \partial^{\dagger} f$ one has $\partial^{(\cdot)} f(x) = \partial^p f(x) = \partial^{\dagger} f(x)$. In particular, $\partial^- f(x) = \partial^{\uparrow} f(x)$ for x near $\overline{x}$.*

For the following supplement, let us recall that X is said to be *superreflexive* if it admits an equivalent uniformly convex norm.

Corollary 3. *Suppose X is uniformly smooth and $\partial^{(\cdot)}$ is a valuable subdifferential on X contained in $\partial^{\dagger}$. Then, for $p > 1$, the assertions (a)-(d) of the preceding theorem are consequences of the following one:*
(e) $\exists \rho, \sigma > 0$ such that $f + \sigma \|\cdot\|^p$ is convex on $B(\overline{x}, \rho)$.
If the norm is uniformly convex, assertions (a)-(d) imply (e).
In particular, if X is superreflexive, conditions (a) and (e) are equivalent.

Corollary 4. *Suppose $f : U \to \mathbb{R}$ is a differentiable function on some open subset U of X with a locally $(p-1)$-Hölderian derivative, with $p \in (1, 2]$. Then f is p-paraconvex on U.*

Other results about p-paraconvex functions and approximate convex functions can be found in [3], [42] and [44]. In particular, representations of such functions as marginal functions are presented there.

5 Approximate Convexity and Paraconvexity of Sets

We observe that using the notions of approximate convexity and p- paraconvexity for the indicator function ι_E of a subset E of X would lead to convexity of E and not to a relaxed form of convexity. Therefore, we rather use the distance function $d_E(\cdot) := \inf_{e \in E} d(\cdot, e)$. In the sequel $\overline{x}$ is a point of E and $p \in (1, +\infty)$.

Definition 6. *A subset E of X is said to be approximately convex (respectively p-paraconvex) around $\overline{x}$ if its associated distance function d_E is approximately convex (respectively p-paraconvex) around $\overline{x}$.*

Example. $E := \{(r,s) \in \mathbb{R}^2 : s \geq |r| - r^2\}$ is p-paraconvex but nonconvex. The following result is an easy consequence of Theorem 1; when $\partial^{(\cdot)} = \partial^{\uparrow}$ it can be deduced from [18].

Theorem 3. *Let $\partial^{(\cdot)}$ be a subdifferential on the family $\mathcal{L}(X)$ of Lipschitz functions on X such that $\partial^{(\cdot)} f \subset \partial^{\uparrow} f$ for any $f \in \mathcal{L}(X)$ and let $\overline{x}$ be an element of a subset E of X. Then, among the following assertions, one has the implications (a)⇒(b)⇒(c)⇒(d). If moreover $\partial^{(\cdot)}$ is Lipschitz-valuable on X, in particular if $\partial^{(\cdot)} := \partial^{\uparrow}$, $\partial^{\diamond}$, all these assertions are equivalent.*
(a) E is approximately convex around $\overline{x}$ in the sense that d_E is approximately convex around $\overline{x}$;
(b) for any $\varepsilon > 0$ there exists $\rho > 0$ such that for any $x \in B(\overline{x}, \rho)$ and any $v \in B(0, \rho)$ one has

$$d_E^{\uparrow}(x,v) \leq d_E(x+v) - d_E(x) + \varepsilon \|v\| ; \tag{2}$$

(c) for any $\varepsilon > 0$ there exists $\rho > 0$ such that for any $x \in B(\overline{x}, \rho)$, any $x^* \in \partial^{(\cdot)} d_E(x)$ and any $(u,t) \in S_X \times (0, \rho)$ one has

$$\langle x^*, u \rangle \leq \frac{d_E(x+tu) - d_E(x)}{t} + \varepsilon; \tag{3}$$

(c') for any $\varepsilon > 0$ there exists $\rho > 0$ such that for any $x \in B(\overline{x}, \rho)$, any $x^* \in \partial^{(\cdot)} d_E(x)$ and any $v \in \rho \overline{B}_X$ one has

$$\langle x^*, v \rangle \leq d_E(x+v) - d_E(x) + \varepsilon \|v\| ; \tag{4}$$

(d) $\partial^{(\cdot)} d_E$ is approximately monotone around $\overline{x}$;

Corollary 5. *If E is approximately convex around $\overline{x}$ then, for any subdifferential $\partial^{(\cdot)}$ such that $\partial^{-} \subset \partial^{(\cdot)} \subset \partial^{\uparrow}$ one has $\partial^{-} d_E(\overline{x}) = \partial^{(\cdot)} d_E(\overline{x}) = \partial^{\uparrow} d_E(\overline{x})$.*

Theorem 4. *Let $\partial^{(\cdot)}$ be a subdifferential on the family $\mathcal{L}(X)$ of Lipschitz functions on X such that $\partial^{(\cdot)} f \subset \partial^{\uparrow} f$ for any $f \in \mathcal{L}(X)$. Then, among the following assertions, one has the implications (a)⇒(b)⇒(c)⇒(d).*
If moreover $\partial^{(\cdot)}$ is Lipschitz-valuable on X all these assertions are equivalent.
(a) E is p-paraconvex around $\overline{x}$;
(b) $\exists c, \rho > 0$ such that $\forall x \in B(\overline{x}, \rho)$, $\forall v \in B(0, \rho)$ one has

$$d_E^{\uparrow}(x,v) \leq d_E(x+v) - d_E(x) + c \|v\|^p ; \tag{5}$$

(c) $\exists c, \rho > 0$ such that $\forall x \in B(\overline{x}, \rho)$, $\forall x^* \in \partial^{(\cdot)} d_E(x)$, $\forall (u,t) \in S_X \times (0, \rho)$ one has

$$\langle x^*, u \rangle \leq \frac{d_E(x+tu) - d_E(x)}{t} + ct^{p-1}; \tag{6}$$

(c') $\exists c, \rho > 0$ such that $\forall x \in B(\overline{x}, \rho)$, $\forall x^* \in \partial^{(\cdot)} d_E(x)$, $\forall v \in \rho \overline{B}_X$ one has

$$\langle x^*, v \rangle \leq d_E(x+v) - d_E(x) + c \|v\|^p; \tag{7}$$

(d) $\partial^{(\cdot)} d_E$ is p-paramonotone around $\overline{x}$.

The following immediate consequence is stronger than the conclusion of Corollary 5 since it holds for points around $\overline{x}$ and not just for $\overline{x}$.

Corollary 6. *If E is p-paraconvex around $\overline{x}$ for some $p > 1$, then, for any subdifferential $\partial^{(\cdot)}$ such that $\partial^p f \subset \partial^{(\cdot)} f \subset \partial^{\uparrow} f$ for any Lipschitz function f, one has $\partial^p d_E(x) = \partial^- d_E(x) = \partial^{(\cdot)} d_E(x) = \partial^{\uparrow} d_E(x)$ for x close to $\overline{x}$.*

Proof. Relation (5) shows that for any $x \in B(\overline{x}, \rho)$, $v \in B(0, \rho)$ and any $x^* \in \partial^{\uparrow} f(x)$ one has

$$\langle x^*, v \rangle \leq d_E^{\uparrow}(x, v) \leq d_E(x+v) - d_E(x) + c \|v\|^p,$$

hence $x^* \in \partial^- f(x)$ since $p > 1$.

6 Intrinsic Approximate Convexity and *p*-Paraconvexity

In the present section we deal with two variants of the classes of sets studied in the preceding section. We raise the problem: are these classes different from the preceding ones?

Definition 7. *A subset E of X is said to be intrinsically approximately convex around $\overline{x} \in E$ if for any $\varepsilon > 0$ there exists $\rho > 0$ such that for any $x_1, x_2 \in E \cap B(\overline{x}, \rho)$, $t \in [0, 1]$, one has*

$$d_E((1-t)x_1 + tx_2) \leq \varepsilon t(1-t) \|x_1 - x_2\|. \tag{8}$$

It is intrinsically approximately convex if it is intrinsically approximately convex around each of its points.

Definition 8. *Given $p \in (1, +\infty)$, a subset E of X is said to be intrinsically p-paraconvex around $\overline{x} \in E$ if there exist $c, \rho > 0$ such that for any $x_1, x_2 \in E \cap B(\overline{x}, \rho)$, $t \in [0, 1]$, one has*

$$d_E((1-t)x_1 + tx_2) \leq ct(1-t) \|x_1 - x_2\|^p. \tag{9}$$

It is intrinsically p-paraconvex if it is intrinsically p-paraconvex around each of its points.

Characterizations can be given as follows. When one of the assertions (b)-(d) holds, we say that E is $\partial^{(\cdot)}$-intrinsically p-paraconvex around $\overline{x}$.

Theorem 5. *Suppose $\partial^{(\cdot)} f \subset \partial^{\uparrow} f$ for any Lipschitz function f on X. Then, among the following assertions one has the implications (a)⇒(b)⇒(c)⇔(d)⇐(e).*
For $\partial^{(\cdot)} = \partial^{\uparrow}$ one has (b)⇔(c).
When X is a $\partial^{(\cdot)}$-subdifferentiability space one has (e)⇒(a). If X is an Asplund space and $\partial^{-} \subset \partial^{(\cdot)} \subset \partial^{\uparrow}$, then assertions (a)-(e) are equivalent.
(a) E is intrinsically p-paraconvex around $\overline{x}$;
(b) $\exists c, \delta > 0$ such that $\forall x, x' \in E \cap B(\overline{x}, \delta)$, one has

$$d_E^{\uparrow}(x, x' - x) \leq c \|x - x'\|^p ; \tag{10}$$

(c) $\exists c, \delta > 0$ such that $\forall x, x' \in E \cap B(\overline{x}, \delta)$, $x^* \in \partial^{(\cdot)} d_E(x)$, one has

$$\langle x^*, x' - x \rangle \leq c \|x - x'\|^p ; \tag{11}$$

(d) $\partial^{(\cdot)} d_E(\cdot)$ is p-paramonotone around $\overline{x}$ on E : there exist $c, \delta > 0$ such that for any $x_1, x_2 \in E \cap B(\overline{x}, \delta)$, $x_1^* \in \partial^{(\cdot)} d_E(x_1)$, $x_2^* \in \partial^{(\cdot)} d_E(x_2)$ one has

$$\langle x_1^* - x_2^*, x_1 - x_2 \rangle \geq -c \|x_1 - x_2\|^p ; \tag{12}$$

(e) there exist $c, \rho > 0$ such that $\forall w \in B(\overline{x}, \rho)$, $x \in E \cap B(\overline{x}, \rho)$, $w^* \in \partial^{(\cdot)} d_E(w)$ one has

$$d_E(w) + \langle w^*, x - w \rangle \leq c \|x - w\|^p . \tag{13}$$

Now let us give some specializations to some specific subdifferentials and normal cones.

Corollary 7. *If E is intrinsically p-paraconvex around $\overline{x}$ then there exists $\delta > 0$ such that for $x \in E \cap B(\overline{x}, \delta)$ one has $N^{\uparrow}(E, x) = N^{-}(E, x)$ and*
(f) $\exists c, \delta > 0$ such that for any $\forall x, x' \in E \cap B(\overline{x}, \delta)$, $x^* \in N^{-}(E, x)$ one has

$$\langle x^*, x' - x \rangle \leq c \|x^*\| \|x - x'\|^p . \tag{14}$$

The preceding property can be related to a global one as in the works of Canino [11], Colombo-Goncharov [16], De Giorgi-Marino-Tosques [20], Degiovani-Marino-Tosques [21], Federer [23] for X a Hilbert space.

Definition 9. *Given a subset E of X, $p \in (1, +\infty)$ and a continuous function $\varphi : E \times E \to \mathbb{R}_+$, the subset E of X is said to be φ−p-convex if for any $x, y \in E$ and $x^* \in N^{-}(E, x)$ one has*

$$\langle x^*, y - x \rangle \leq \varphi(x, y) \|x^*\| \|x - y\|^p . \tag{15}$$

The choice of the firm normal cone is natural since if this definition holds for some other normal cone $N^{(\cdot)}(E, x)$ one has $N^{(\cdot)}(E, x) \subset N^{-}(E, x)$.
The following result clarifies the links between $\varphi - p$-convexity and p - paraconvexity.

Proposition 5. *Let $p \in (1, +\infty)$, a subset E of a Banach space X, a continuous function $\varphi : E \times E \to \mathbb{R}_+$ be such that E is $\varphi - p$-convex. Then E is intrinsically firmly p-paraconvex around each point of E in the sense that assertion (c) of Theorem 5 is satisfied with $\partial^{(\cdot)} = \partial^-$.*
Conversely, if E is intrinsically firmly p-paraconvex around each point of E, then E is $\varphi - p$-convex for some continuous function $\varphi : E \times E \to \mathbb{R}_+$.

Let us introduce the following concept.

Definition 10. *A multimapping $F : X \rightrightarrows Y$ between two n.v.s. is said to be approximately convex (resp. intrinsically approximately convex) around $(\overline{x}, \overline{y})$ if its graph is an approximately convex (resp. intrinsically approximately convex) subset of $X \times Y$ around $(\overline{x}, \overline{y})$.*

Let us give some properties. We start with intersection of subsets.

Proposition 6. *Let $E_1, E_2, \cdots, E_n$ be intrinsically $p-$paraconvex sets around $\overline{x} \in E := E_1 \cap \cdots \cap E_n$. Suppose that the following standard qualification condition is satisfied: there exist $b, \delta > 0$ such that*

$$\|x_1^* + \cdots + x_n^*\| \geq b \quad \forall x \in B(\overline{x}, \delta) \setminus E,\ \forall x_i^* \in \partial^{\uparrow} d_{E_i}(x).$$

Then E is intrinsically $p-$paraconvex around $\overline{x}$.

Proposition 7. *Let $E \subset X$ be an intrinsically $p-$paraconvex set around $\overline{x} \in E$ and let $f : X \to Y$ be a nonexpansive mapping onto another n.v.s. Y. Then $F := f(E)$ is intrinsically $p-$paraconvex around $\overline{y} := f(\overline{x})$.*
If f is open around $\overline{x}$ and if E is intrinsically approximately convex around $\overline{x}$, then $F := f(E)$ is intrinsically approximately convex around $\overline{y} := f(\overline{x})$.

Corollary 8. *If $F : X \rightrightarrows Y$ is an intrinsically approximately convex multimapping around $(\overline{x}, \overline{y})$, then its domain and its image are intrinsically approximately convex around $\overline{x}$ and $\overline{y}$ respectively.*

7 Links Between *p*-Paraconvex Sets and Functions

In the sequel, we endeavour to relate paraconvexity of sets and paraconvexity of functions. We start with epigraphs. We endow the product space $X := W \times \mathbb{R}$ of a n.v.s. W with $\mathbb{R}$ with a product norm, i.e. a norm such that the projections and the insertions $w \mapsto (w, 0)$ and $r \mapsto (0, r)$ are nonexpansive. Then, for each $(w, r) \in W \times \mathbb{R}$ we have

$$\max\left(\|w\|, |r|\right) \leq \|(w, r)\| \leq \|w\| + |r|.$$

Proposition 8. *Let W be a normed vector space and let $f : W \to \mathbb{R} \cup \{+\infty\}$ be a l.s.c. function which is approximately convex (resp. p-paraconvex) around $\overline{w} \in W$. Then, for any $\overline{r} \geq f(\overline{w})$, the epigraph E of f is intrinsically approximately convex (resp. p-paraconvex) around $\overline{x} := (\overline{w}, \overline{r})$.*

Let us complete the preceding result with the following one.

Proposition 9. *Let $f : W \to \mathbb{R}$ be a function which is Lipschitzian with rate $\ell > 0$ on some ball $B(\overline{w}, \rho)$. Suppose $X := W \times \mathbb{R}$ is endowed with the norm given by $\|(w, r)\| = \ell \|w\| + |r|$. If f is p-paraconvex around $\overline{w}$, then, for any $\overline{r} \geq f(\overline{w})$, the epigraph E of f is p-paraconvex around $\overline{x} := (\overline{w}, \overline{r})$.*

Let us give a kind of converse to the preceding propositions.

Theorem 6. *Let W be a Banach space and let $f : W \to \mathbb{R}$ be a function which is locally Lipschitzian around $\overline{w} \in W$ and such that the epigraph E of f is an intrinsically approximately convex (resp. intrinsically p-paraconvex) subset of $X := W \times \mathbb{R}$ around $\overline{x} := (\overline{w}, f(\overline{w}))$. Then f is an approximately convex (resp. p-paraconvex) function around $\overline{w}$.*

Let us say that a function $f : X \to \mathbb{R} \cup \{+\infty\}$ is an *approximately quasi-convex function* if its sublevel sets are approximately convex. The following result shows that under some qualification condition an approximately convex function is approximately quasi-convex.

Proposition 10. *Let X be a Banach space with a norm which is Fréchet differentiable off 0 and let $f : X \to \mathbb{R}$ be a continuous function. Suppose f is approximately convex around $\overline{x} \in S := \{x \in X : f(x) \leq 0\}$ and there exist $c > 0$, $r > 0$ such that $\|x^*\| \geq c$ for each $x \in (X \backslash S) \cap B(\overline{x}, r)$ and each $x^* \in \partial^- f(x)$. Then S is intrinsically approximately convex around $\overline{x}$.*

8 Paraconvex Sets and Projections

The following result is reminiscent of [15, Thm 4.1] which takes place in a Hilbert space. However, here U is not a uniform entourage of E; it may be small (or large) and far from E.

Theorem 7. *Suppose that the norm of X is Fréchet differentiable on $X \backslash \{0\}$. Let E be a closed subset of X and let U be an open subset of X. Consider the following assertions*
(a) Each $w \in U$ has a unique metric projection $P_E(w)$ in E and the mapping $P_E(\cdot)$ is continuous on $U \setminus E$.
(b) $d_E(\cdot)$ is continuously differentiable on $U \setminus E$.
(c) $d_E(\cdot)$ is approximately convex on $U \setminus E$.
Then, one has (a) $\Rightarrow$ (b) $\Rightarrow$ (c). If X is uniformly Fréchet smooth, then (a) $\Rightarrow$ (b) $\Longleftrightarrow$ (c).
If, in addition, X is strictly convex and the norm of X has the Kadec-Klee property, then (a) $\Longleftrightarrow$ (b) $\Longleftrightarrow$ (c).

Theorem 8. *Let X be a super-reflexive Banach space, let E be a closed subset of X and let U be an open subset of X. The following assertions relative to some choices of $p \in (1,2]$ and of an equivalent norm on X are equivalent:*
(a) Each $w \in U$ has a unique metric projection $P_E(w)$ in E and the mapping $w \mapsto P_E(w)$ is locally Hölderian on $U \setminus E$.
(b) $d_E(\cdot)$ is differentiable with a locally Hölderian derivative on $U \setminus E$.
(c) For each $\overline{w} \in U \setminus E$, $d_E(\cdot)$ is p-paraconvex around $\overline{w}$ for some $p > 1$.

References

1. Attouch, H. and Azé, D., Approximation and regularization of arbitrary functions in Hilbert spaces by the Lasry-Lions method, Ann. Inst. H. Poincaré, Anal. Non Linéaire 10, No.3 (1993), 289-312.
2. Aubin, J.-P., Frankowska, H., *Set-Valued Analysis*, Birkhaüser, Boston (1990).
3. Aussel, D., Daniilidis, A. and Thibault, L., Subsmooth sets: functional characterizations and related concepts, Trans. Amer. Math. Soc. 357, No.4 (2005), 1275-1301.
4. Azé, D. and Penot, J.-P., Uniformly convex and uniformly smooth convex functions, Ann. Fac. Sci. Toulouse IV (4) (1995), 705-730.
5. Benyamini, Y, Lindenstrauss, J., *Geometric Nonlinear Functional Analysis*, Amer. Math. Soc. Colloquium Publications 48, Providence (2000).
6. Bernard, F., Thibault, L., Prox-regularity of functions and sets in Banach spaces, Set-Valued Anal. 12, No.1-2, 25-47 (2004).
7. Bernard, F., Thibault, L., Uniform prox-regularity of functions and epigraphs in Hilbert spaces, Nonlinear Anal., Theory Methods Appl. 60A, No.2, 187-207 (2005).
8. Bernard, F., Thibault, L., Prox-regular functions in Hilbert spaces, J. Math. Anal. Appl. 303, No.1, 1-14 (2005)
9. Bougeard, M. L., Morse theory for some lower-C^2 functions in finite dimension, Math. Program., Ser. A 41, No.2 (1988), 141-159.
10. Bougeard, M., Penot, J.-P. and Pommellet, A., Towards minimal assumptions for the infimal convolution regularization, J. Approximation Theory 64, No.3 (1991), 245-270.
11. Canino, A., On p-convex sets and geodesics, J. Differ. Equations 75, No.1 (1988), 118-157.
12. Cannarsa, P. and Sinestrari, C., *Semiconcave Functions, Hamilton-Jacobi Equations and Optimal Control*, Birkhäuser, Basel (2004).
13. Castellani, M. and Pappalardo, M., Characterization of ρ-convex functions, in *Generalized Convexity, Generalized Monotonicity: Recent Results*, Crouzeix, J.-P. et al., eds., Kluwer, Dordrecht ((1998), 219-233.
14. Clarke, F.H., *Optimization and Nonsmooth Analysis*, Wiley Interscience, New York, New York (1983).
15. Clarke, F.H., Stern, R.J. and Wolenski, P.R., Proximal smoothness and the lower-C^2 property, J. Convex Anal. 2 (1,2) (1995), 117-144.
16. Colombo, G., and Goncharov, V., Variational inequalities and regularity properties of closed sets in Hilbert spaces, J. Convex Anal. 8 (2001), 197-221.
17. Correa, R., Jofre, A., and Thibault, L., Subdifferential monotonicity as a characterization of convex functions, Numer. Funct. Anal. Optim. 15 (1994), 531-535.

18. Daniilidis, A. , Georgiev, P., Approximate convexity and submonotonicity, J. Math. Anal. Appl. 291 (2004), 292-301.
19. Daniilidis, A., Georgiev, P. and Penot, J.-P., Integration of multivalued operators and cyclic submonotonicity, Trans. Amer. Math. Soc. 355 (2003), 177-195.
20. De Giorgi, E., Marino, A., and Tosques, M., (p, q)-convex functions, Atti Accad. Naz. Lincei, VIII. Ser., Rend., Cl. Sci. Fis. Mat. Nat. 73, 6-14 (1982).
21. Degiovanni, M., Marino, A., and Tosques, M., General properties of (p, q)-convex functions and (p, q)-monotone operators, Ric. Mat. 32, 285-319 (1983).
22. Deville, R., Godefroy, G., Zizler, V., *Smoothness and Renormings in Banach Spaces*, Pitman Monographs 64, Longman, Harlow (1993).
23. Federer, H., Curvature measures, Trans. Amer. Math. Soc. 93 (1959), 418-491.
24. Georgiev, P., Submonotone Mappings in Banach Spaces and Applications, Set-Valued Analysis 5 (1997), 1-35.
25. Ioffe, A.D., Approximate subdifferentials and applications. III: The metric theory, Mathematika 36, No.1, 1-38 (1989).
26. Ioffe, A.D., Proximal analysis and approximate subdifferentials, J. London Math. Soc. 41 (1990), 175-192.
27. Ioffe, A., Codirectional compactness, metric regularity and subdifferential calculus, Théra, M. (ed.), *Constructive, experimental, and nonlinear analysis. Selected papers of a workshop, Limoges, France, September 22-23, 1999*, American Mathematical Society (AMS), publ. for the Canadian Mathematical Society. CMS Conf. Proc. 27, Providence, RI (2000), 123-163.
28. Ioffe, A.D., Metric regularity and subdifferential calculus, Russ. Math. Surv. 55, No.3, 501-558 (2000); translation from Usp. Mat. Nauk 55, No.3, 103-162 (2000).
29. Ivanov, M. and N. Zlateva, Abstract subdifferential calculus and semi-convex functions, Serdica Math. J. 23 (1997), 35-58.
30. Jourani, A., Subdifferentiability and subdifferential monotonicity of γ - paraconvex functions, Control Cybern. 25, No.4 (1996), 721-737.
31. Kruger, A.Y., and Mordukhovich, B. S., Extremal points and the Euler equation in nonsmooth optimization problems, Dokl. Akad. Nauk BSSR 24 (1980), 684-687.
32. Lasry, J.-M. and Lions P.-L., A remark on regularization in Hilbert spaces, Isr. J. Math. 55 (1986), 257-266.
33. Luc, D.T., Ngai, H.V., and Théra, M., On ε-convexity and ε-monotonicity, in *Calculus of Variations and Differential Equations*, A. Ioffe, S. Reich and I. Shafrir (eds.), Research Notes in Maths. Chapman & Hall, (1999), 82-100.
34. Marino, A. and Saccon, C., Subdifferential analysis and plates subjected to unilateral constraints, Nonsmooth optimization and related topics, Proc. 4th Course Int. Sch. Math., Erice/Italy 1988, Ettore Majorana Int. Sci. Ser., Phys. Sci. 43 (1989), 265-276.
35. Marino, A., and Tosques M., Some variational problems with lack of convexity and some partial differential inequalities, *Methods of Nonconvex Analysis*, Lect. 1st Sess. CIME, Varenna/Italy 1989, Lect. Notes Math. 1446 (1990), 58-83.
36. Michel, Ph. and Penot, J.-P., Calcul sous-différentiel pour des fonctions lipschitziennes et non lipschitziennes, C. R. Acad. Sci., Paris, Sér. I 298 (1984), 269-272.
37. Michel, Ph. and Penot, J.-P., A generalized derivative for calm and stable functions, Differ. Integral Equ. 5, No.2 (1992), 433-454.

38. Mifflin, R., Semismooth and semiconvex functions in constrained optimization, SIAM J. Control Optim. 15 (1977), 959-972.
39. Mordukhovich, B. S., and Shao Y., Nonsmooth sequential analysis in Asplund spaces, Trans. Amer. Math. Soc. 348 (4) (1996), 1235-1280.
40. Ngai, H.V., Luc, D.T., and Théra, M., Approximate convex functions, J. Nonlinear and Convex Anal. 1 (2) (2000), 155-176.
41. Ngai, H.V., and Penot, J.-P., Semismoothness and directional subconvexity of functions, preprint, May 2003, revised May 2005.
42. Ngai, H.V., and Penot, J.-P., Approximately convex functions and approximately monotone operators, preprint, October 2004.
43. Ngai, H.V., and Penot, J.-P., Approximately convex sets, preprint, October 2004.
44. Ngai, H.V., and Penot, J.-P., Paraconvex functions and paraconvex sets, preprint, October 2004.
45. Ngai, H.V., and Penot, J.-P., Generic differentiability of approximately convex functions, preprint, December 2005.
46. Ngai, H.V., and Théra, M., Metric inequality, subdifferential calculus and applications, Set-Valued Analysis, 9, 187-216 (2001).
47. Penot, J.-P., Favorable classes of mappings and multimappings in nonlinear analysis and optimization, J. Convex Analysis 3 (1996), 97-116.
48. Penot, J.-P., Softness, sleekness and regularity properties, preprint, 2005.
49. Penot, J.-P. and Volle, M., On strongly convex and paraconvex dualities, in *Generalized convexity and fractional programming with economic applications, Proc. Workshop, Pisa/Italy 1988*, Lect. Notes Econ. Math. Syst. 345 (1990), 198-218.
50. Poliquin, R.A., Rockafellar, R.T., Prox-regular functions in variational analysis, Trans. Am. Math. Soc. 348, No.5, 1805-1838 (1996).
51. Poliquin, R.A., Rockafellar, R.T. and Thibault, L., Local differentiability of distance functions, Trans. Am. Math. Soc. 352, No.11, 5231-5249 (2000).
52. Rockafellar, R.T., Generalized directional derivatives and subgradients of nonconvex functions, Can. J. Math. 32 (2) (1980), 257-280.
53. Rockafellar, R.T., Favorable classes of Lipschitz continuous functions in subgradient optimization, in *Nondifferentiable Optimization* (1982), Nurminski E. (eds), Pergamon Press, New York.
54. Rockafellar, R.T. and Wets, R. J-B, *Variational Analysis*, Springer, New York (1998).
55. Rolewicz, S., On γ-paraconvex multifunctions, Math. Japon. 24 (1979), 31-47.
56. Rolewicz, S., On $\alpha(\cdot)$-monotone multifunctions and differentiability of γ-paraconvex functions, Studia Math. 133 (1) (1999), 29-37.
57. Rolewicz, S., On $\alpha(\cdot)$-paraconvex and strongly $\alpha(\cdot)$-paraconvex functions, Control and Cybernetics 29 (1) (2000), 367-377.
58. Rolewicz, S., On uniformly approximate convex and strongly $\alpha(\cdot)$-paraconvex functions, Control and Cybernetics 30 (3) (2001), 323-330.
59. Rolewicz, S., On the coincidence of some subdifferentials in the class of $\alpha(\cdot)$-paraconvex functions, Optimization 50 (2001), 353-360.
60. Rolewicz, S., On $\alpha(\cdot)$-monotone multifunctions and differentiability of strongly $\alpha(\cdot)$-paraconvex functions, Control and Cybernetics 31 (3) (2002), 601-619.
61. Rolewicz, S., Paraconvex analysis, preprint, Polish Acad. Sci. (2004).
62. Spingarn, J.E., Submonotone subdifferentials of Lipschitz functions, Trans. Amer. Math. Soc. 264 (1981), 77-89.

63. Vial, J.-Ph., Strong and weak convexity of sets and functions, *Math. Oper. Research* 8 (2) (1983), 231-259.

Monotonicity and Dualities

Jean-Paul Penot

Université de Pau, Faculté des Sciences, Laboratoire de Mathématiques Appliquées, CNRS UMR 5142 Av. de l'Université 64000 PAU, France.

Summary. There is a recent surge of interest for the representation of monotone operators by convex functions. It can explained by the success of convex analysis in obtaining the fundamental results about maximal monotone operators. Convex analysis can also be combined with variational analysis to get new convergence results. Here we take another direction and connect such a stream with the concept of duality in a general framework, heavily using order methods.

Key words: Conjugacy, convexity, duality, generalized convexity, generalized monotonicity, monotonicity, polarity.

1 Introduction

It is one of the purposes of mathematics to clarify a question by putting it in general, abstract terms which avoid the complexity of real-world problems and thus make them tractable. Of course, the gain lies in the balance between simplication (i.e. abandon of contingent peculiarities) and abstraction.

In [45] we proposed a topological approach to the question of extension of usual operations for monotone operators. Another vein is the use of order, another fundamental tool in mathematics (see for instance [5], [33]). What follows is an attempt to introduce explicit order methods and abstract convexity and duality in the representation of monotone operators; see also [41].

Representations of monotone operators have been known since the pioneering works of Krauss ([22]-[24]), and Fitzpatrick ([15]). Since it has been shown that such representations are not just aesthetic, but have some usefulness ([40], [44], [46]- [7], [55]...), the subject is experiencing a great expansion ([4], [6], [8], [9]-[11], [29], [30], [47], [56], [62]...).

The concept of duality is a convenient framework ([2], [5], [14], [19], [21], [25]-[28], [31]-[43], [49]-[53], [57], [48]); it has shown to be effective in a number of situations. Recall that a *duality* between two ordered spaces $\mathcal{L}$ and $\mathcal{M}$ is an antitone (i.e. order reversing) mapping $D : \mathcal{L} \to \mathcal{M}$ such that

$$D(\bigwedge_{i\in I} \ell_i) = \bigvee_{i\in I} D(\ell_i)$$

for any family $(\ell_i)_{i\in I}$ in $\mathcal{L}$ for which $\bigwedge_{i\in I} \ell_i := \inf_{i\in I} \ell_i$ exists. Dualities are often used in the case $\mathcal{L}$ and $\mathcal{M}$ are lattices or complete lattices. Conjugacies are much studied examples of duality ([31]): given two sets W, Z and a map $b : W \times Z \to \overline{\mathbb{R}}$, taking $\mathcal{L} := \overline{\mathbb{R}}^W$, $\mathcal{M} := \overline{\mathbb{R}}^Z$ the conjugacy associated with the coupling function b is the map $D : \mathcal{L} \to \mathcal{M}$ given by

$$f^b(z) := -\inf_{w\in W} (f(w) - b(w, z)) \qquad\qquad f \in \mathcal{L},\ z \in Z.$$

Conjugacies have been characterized among dualities as the dualities D for which the relation $D(f + r) = -(r - D(f))$ for every $f \in \overline{\mathbb{R}}^W$, $r \in \overline{\mathbb{R}}$ holds, where the addition of $\mathbb{R}$ is extended to $\overline{\mathbb{R}}$ by setting $r + (+\infty) = +\infty$ for each $r \in \overline{\mathbb{R}}$ and $(-\infty) + (-\infty) = -\infty$ and where $r - s := r + (-s)$. Here, for the sake of simplicity we avoid the use of the opposite convention; the price to be paid is a certain awkwardness in some formulae such as the preceding one. One also has to be careful enough in cancelling terms or reorganizing inequalities or equalities; see [31] and Lemma 1 below.
When the ordered sets $\mathcal{L}$ and $\mathcal{M}$ are subsets of the power sets 2^W, 2^Z of some spaces W, Z ,the orders in $\mathcal{L}$ and $\mathcal{M}$ being the opposite of the inclusion, and when $P : \mathcal{L} \to \mathcal{M}$ is a duality, i.e. satisfies

$$P(\bigcup_{i\in I} A_i) = \bigcap_{i\in I} P(A_i),$$

one says that P is a *polarity*. Here we will deal with the mixed case in which one of the spaces is a power set and the other one is a function space, with its pointwise order. The passage from sets to functions and the reverse passage have proved to be fruitful in various areas of mathematics. Among many instances, let us mention the passage from a closed subset A of a metric space to its *distance function* $d_A := \inf_{a\in A} d(a, \cdot)$, the passage from a subset A of a set W to its *indicator* function ι_A given by $\iota_A(w) = 0$ if $w \in A$, $\iota_A(w) = +\infty$ if $w \in W\backslash A$ and, in the reverse direction, the passage from a function to its graph or its epigraph. In particular, given a polarity $P : \mathcal{L} \to \mathcal{M}$, let us observe that considering the injections $A \mapsto \iota_A$ and $B \mapsto \iota_B$ of $\mathcal{L}$ and $\mathcal{M}$ into $\overline{\mathbb{R}}^W$ and $\overline{\mathbb{R}}^Z$ respectively as identifications, any polarity can be considered as a special duality.
The aim of the present note consists in trying to take advantage of the numerous dualities which have been defined in function spaces (see examples and references in [34], [36], [43], [51], [57], [60]) in order to construct new polarities. As a byproduct, we obtain a means to represent certain generalized monotone operators by functions. Much more remains to be done in this second direction.

2 A General Framework

In the sequel, it will be convenient to say that a map $M : \mathcal{L} \to \mathcal{M}$ between two ordered sets is a *coduality* if it is a duality for the reverse orders in $\mathcal{L}$ and $\mathcal{M}$, i.e. if it satisfies

$$M(\sup_{i\in I} \ell_i) = \inf_{i\in I} M(\ell_i)$$

for any family $(\ell_i)_{i\in I}$ in $\mathcal{L}$ for which $\sup_{i\in I} \ell_i$ exists. Taking the reverse order in $\mathcal{M}$, one gets a familiar notion. In particular, if $\mathcal{L}$ and $\mathcal{M}$ are *sup-lattices* (in the sense that for any family of elements the supremum of the family exists) and if

$$M(\sup_{i\in I} \ell_i) = \sup_{i\in I} M(\ell_i)$$

for any family $(\ell_i)_{i\in I}$ in $\mathcal{L}$, we say that M is a *morphism of sup-lattices.* A similar definition can be given for morphisms of inf-lattices; note that here we use the expression inf-lattice or sup-lattice for complete inf-sublattice or sup-sublattice respectively.
The following result is probably the transcription to codualities of a classical fact for dualities; but we are not aware of a precise statement under such general assumptions. The notation we choose takes into account the fact that in the sequel $\mathcal{P}$ will be a power set.

Proposition 1. *Given a sup-lattice $\mathcal{P}$, an inf-lattice $\mathcal{F}$ and an antitone map $J : \mathcal{P} \to \mathcal{F}$, there is a smallest antitone map $J^\dagger : \mathcal{F} \to \mathcal{P}$ such that $J^\dagger(J(S)) \geq S$ for every $S \in \mathcal{P}$. It is given by*

$$J^\dagger(f) = \bigvee\{S \in \mathcal{P} : J(S) \geq f\}. \tag{1}$$

Moreover, if J is a coduality, one has $J(J^\dagger(f)) \geq f$ for every $f \in \mathcal{F}$ and $J^\dagger$ is a coduality.

When J is a coduality, $J^\dagger$ will be called the *reverse coduality* of J.
Proof. Let $J^\dagger$ be defined by (1). Clearly $J^\dagger$ is antitone and $J^\dagger$ is such that $J^\dagger(J(S)) \geq S$ for every $S \in \mathcal{P}$. Let $N : \mathcal{F} \to \mathcal{P}$ be an antitone map such that $N(J(S)) \geq S$ for every $S \in \mathcal{P}$. Given $f \in \mathcal{F}$, for every $S \in \mathcal{P}$ such that $J(S) \geq f$, we have $N(f) \geq N(J(S)) \geq S$, hence $N(f) \geq J^\dagger(f)$.
If J is a coduality, for every $f \in \mathcal{F}$, one has $J(J^\dagger(f)) = \inf\{J(S) : S \in \mathcal{P},\ J(S) \geq f\} \geq f$. Let us show that $J^\dagger$ is a coduality. Let $(f_i)_{i\in I}$ be any family in $\mathcal{F}$ for which $f := \sup_{i\in I} f_i$ exists. Since $J^\dagger$ is antitone, we have $J^\dagger(f) \leq J^\dagger(f_i)$ for every $i \in I$. On the other hand, if $S \in \mathcal{P}$ is such that $S \leq J^\dagger(f_i)$ for every $i \in I$, then we have $J(S) \geq J(J^\dagger(f_i)) \geq f_i$ for every $i \in I$, hence $J(S) \geq f$ and $J^\dagger(f) \geq S$ by (1); that shows that $J^\dagger(f)$ is the greatest lower bound of $(J^\dagger(f_i))_{i\in I}$ and that $J^\dagger$ is a coduality. □
Given a set Z we consider a sub-sup-lattice (in the sense of complete sublattices, i.e. a subset stable by arbitrary suprema) $\mathcal{L}$ of the set $\mathcal{F} := \overline{\mathbb{R}}^Z$

of functions from Z to $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$ for the pointwise order and a sub-inf-lattice (i.e. a subset stable by arbitrary infima) $\mathcal{M}$ of the power set $\mathcal{P} := \mathcal{P}(Z)$ of Z (the set of subsets of Z, ordered by inclusion). The following easy observation is the basis of our study.

Proposition 2. *Given a coduality $J : \mathcal{P} \to \mathcal{F}$, a duality $C : \mathcal{F} \to \mathcal{F}$ and a coduality $M : \mathcal{F} \to \mathcal{P}$, the map $P := M \circ C \circ J : \mathcal{P} \to \mathcal{P}$ is a polarity. If $\mathcal{M}$ is a sub-inf-lattice of $\mathcal{P}$ such that $M(C(J(S))) \in \mathcal{M}$ for every $S \in \mathcal{M}$, then the restriction of P to $\mathcal{M}$ is a polarity from $\mathcal{M}$ to $\mathcal{M}$.*

Proof. Given a family $(S_i)_{i \in I}$ of subsets of Z one has

$$P(\bigcup_{i \in I} S_i) = M\left(C\left(\inf_{i \in I} J(S_i)\right)\right) = M\left(\sup_{i \in I} C\left(J(S_i)\right)\right) = \bigcap_{i \in I} M\left(C\left(J(S_i)\right)\right)$$
$$= \bigcap_{i \in I} P(S_i).$$

The second assertion is obvious. □

Under the assumptions of the preceding proposition, the map $L := C \circ J : \mathcal{P} \to \mathcal{F}$ is a morphism of sup-lattices. We will study it in more detail in the following sections, assuming a more structured framework.

3 A Useful Construction

Now we suppose that a canonical function $c : Z \to \mathbb{R}_{-\infty} := \mathbb{R} \cup \{-\infty\}$ is given on Z. In the sequel c will be a coupling function; here we take into account that many coupling functions take the value $-\infty$ but not the value $+\infty$ (see [36, section 4]). As mentioned above, we use the familiar extension to $\overline{\mathbb{R}}$ of the addition given by $+\infty + s = +\infty$ for every $s \in \overline{\mathbb{R}}$, $(-\infty) + (-\infty) = -\infty$ and we set $r - s := r + (-s)$ for $r, s \in \overline{\mathbb{R}}$. The following equivalence will be useful.

Lemma 1. *For any $r, s, t \in \mathbb{R}_{-\infty}$ one has the implications*

$$-s \leq r - t \Rightarrow t \leq r + s, \tag{2}$$
$$t \leq r + s \Rightarrow -s \leq r - t. \tag{3}$$

Proof. In (2) the second inequality is obvious for $t = -\infty$ and satisfied when $t \in \mathbb{R}$ since then r, s, t are finite. The same argument applies for the proof of (3). □

We consider the map $J : \mathcal{P} := \mathcal{P}(Z) \to \mathcal{F} := \overline{\mathbb{R}}^Z$ given by

$$J(S) := c_S := \iota_S + c,$$

where ι_S is the indicator function of S. This mapping satisfies

$$J(\bigcap_{i\in I} S_i) = \sup_{i\in I} J(S_i), \qquad J(\bigcup_{i\in I} S_i) = \inf_{i\in I} J(S_i) \tag{4}$$

for any family $(S_i)_{i\in I}$ in $\mathcal{P}$; thus it is an injective duality and coduality for the usual orders in $\mathcal{P}$ (the inclusion) and $\mathcal{F}$ (the pointwise order). In the reverse direction, for $f \in \mathcal{F} := \overline{\mathbb{R}}^Z$, we set

$$M(f) := \{z \in Z : f(z) \le c(z)\}. \tag{5}$$

The following lemma is obvious.

Lemma 2. *The relation M defines a coduality from $\mathcal{F} := \overline{\mathbb{R}}^Z$ into the power set $\mathcal{P} = 2^Z$ of Z, i.e. an antitone mapping which satisfies the following relation for any family $(f_i)_{i\in I}$ in $\mathcal{F}$:*

$$M(\sup_{i\in I} f_i) = \bigcap_{i\in I} M(f_i).$$

The following lemma gives the hint that M might be a more special coduality.

Lemma 3. *One has $M(J(S)) = S$ for any $S \in \mathcal{P}$ and $J(M(f)) \ge f$ for any $f \in \mathcal{F}$.*

Proof. The first assertion is immediate: for $z \in S$ one has $J(S)(z) = c(z)$ and for $z \in Z\backslash S$ one has $J(S)(z) = +\infty > c(z)$. Given $f \in \mathcal{F}$, let $S := M(f)$. For $z \in S$ one has $J(S)(z) = c(z) \ge f(z)$ by definition of $M(f)$; for $z \in Z\backslash S$ one has $J(S)(z) = +\infty \ge f(z)$. Thus $J(S) \ge f$. □

The following observation confirms the preceding hint and provides an interpretation of M showing its status in terms of codualities.

Proposition 3. *The reverse coduality $J^\dagger$ of J is M.*

Proof. This follows from the fact that for any $f \in \mathcal{F}$ one has

$$J^\dagger(f) = \bigcup\{S \in \mathcal{P} : J(S) \ge f\} = \{z \in Z : \iota_{\{z\}} + c \ge f\} = M(f).$$

4 A General Monotone Polarity

In the sequel we suppose $Z := X \times Y$, where X, Y are two sets and $c : X \times Y \to \mathbb{R}_{-\infty}$ is considered as a coupling function. We provide $Z \times Z$ with the coupling b given by

$$\begin{gathered} b(w,z) := b\left((u,v),(x,y)\right) := c(u,y) + c(x,v) \\ w := (u,v), z := (x,y) \in X \times Y, \end{gathered} \tag{6}$$

so that, for every $z := (x,y) \in X \times Y$, one has

$$c(x,y) = \frac{1}{2} b\left(z,z\right) = \frac{1}{2} b\left((x,y),(x,y)\right).$$

In the classical case, X, Y are two normed vector spaces in duality and $c : (x,y) \mapsto \langle x, y\rangle$ is the given pairing. Even in the classical case, b is not the usual coupling on Z, but it enables us to consider Z as paired with itself, a decisive advantage. As mentioned above, the Fenchel-Moreau conjugate of $f \in \mathcal{F}$ with respect to this coupling is given by

$$f^b(x,y) := -\inf\{f(u,v) - b\left((u,v),(x,y)\right) : (u,v) \in X \times Y\}. \tag{7}$$

The following fact is noteworthy.

Lemma 4. *One has $c^b \geq c$.*

Proof. We note that for any $r \in \mathbb{R}_{-\infty}$ we have $r - 2r = -r$. Given $(x,y) \in X \times Y$, taking $f = c$ and $(u,v) = (x,y)$ in relation (7), we get $-c^b(x,y) \leq c(x,y) - 2c(x,y) = -c(x,y)$. □

A similar proof for an arbitrary subset S of Z (not just $S = Z$) yields the following result by taking $(x,y) \in S$ and by plugging $(u,v) := (x,y)$ in relation (7).

Proposition 4. *For any S in $\mathcal{P} := \mathcal{P}(Z)$ the function $f_S := J(S)^b := c_S^b$ satisfies $f_S \geq c$ on S.*

The following concept has been introduced by S. Rolewicz [49], [50].

Definition 1. *A multimapping $S : X \rightrightarrows Y$ (identified with its graph in $X \times Y$) is said to be c-monotone if for any $u, x \in X$, $v \in S(u)$, $y \in S(x)$ one has*

$$c(u,y) + c(x,v) \leq c(u,v) + c(x,y). \tag{8}$$

In the classical case, we just say that S is monotone and the preceding relation can be written

$$\langle u - x, v - y\rangle \geq 0.$$

Let us note that S is c-monotone if, and only if, $f_S(z) \leq c(z)$ for every $z \in S$, or, equivalently, in view of Proposition 4 if, and only if, $f_S(z) = c(z)$ for every $z \in S$. Other characterizations generalizing [44, Prop. 4] and using the function $g_S := c_S^{bb}$ introduced there in the special case just mentioned.

Proposition 5. *For any multimapping S the following assertions are equivalent:*
(a) S is c-monotone;
(b) $f_S \leq c_S$;
(c) $f_S \leq g_S$;
(d) $g_S(w) + g_S(z) \geq b(w,z)$ for every $w, z \in Z$.

Proof. The implication (a)⇒(b) is a reformulation of the observation preceding the statement since $f_S \leq c_S$ means that $f_S(z) \leq c(z)$ for every $z \in S$. The implication (b)⇒(c) stems from the fact that the conjugacy is antitone. For (c)⇒(d) we observe that since $f_S = c_S^b = c_S^{bbb} = g_S^b$ for any $z \in Z$ the relation $g_S(z) \geq f_S(s)$ can be written $g_S(z) \geq b(w,z) - g_S(w)$ for every $w \in Z$. To prove (d)⇒(a) we note that since $c_S \geq c_S^{bb} = g_S$, for $w := (u,v) \in S$, $z := (x,y) \in S$, we have

$$c(u,v) + c(x,y) = c_S(w) + c_S(z) \geq g_S(w) + g_S(z) \geq b(w,z) = c(u,y) + c(x,v),$$

so that S is monotone.
Using (c) and taking $w = z$ in (d) one gets $2g_S(z) \geq b(z,z) = 2c(z)$ and the next corollary.

Corollary 1. *For any monotone multimapping $S : X \rightrightarrows Y$ one has $g_S \geq f_S \geq c$.*

Before describing a polarity associated with this notion, let us present some examples.
Example 1. Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be an arbitrary function with nonempty domain $\operatorname{dom} f$. The *c-subdifferential* of f is the multimapping $\partial^c f : X \rightrightarrows Y$ with domain included in $\operatorname{dom} f$ given for $u \in \operatorname{dom} f$ by

$$v \in \partial^c f(u) \Leftrightarrow \forall x \in X \quad f(x) + c(u,v) \geq f(u) + c(x,v). \tag{9}$$

Then $\partial^c f$ is c-monotone: given $v \in \partial^c f(u)$, $y \in \partial^c f(x)$, writing the analogous relation

$$f(u) + c(x,y) \geq f(x) + c(u,y)$$

and adding sides by sides with the inequality in (9), we obtain (8) after simplification, noting that $f(u)$ and $f(x)$ are finite.
Example 2. Let X be a Banach space, let $Z := X \times X$ and let $c : Z \to \mathbb{R}$ be the (upper) semi-scalar product given by

$$c(x,y) := (x,y)_+ := \lim_{t \to 0_+} \frac{1}{2t} \left(\|x + ty\|^2 - \|x\|^2 \right) = \sup_{x^b \in J_X(x)} \langle x^b, y \rangle,$$

where $J_X := (1/2)\partial \|\cdot\|^2$ is the duality mapping. A c-monotone multimapping $S : X \rightrightarrows X$ is a multimapping satisfying, for any $u, x \in X$, $v \in S(u)$, $y \in S(x)$,

$$(u,v)_+ + (x,y)_+ \geq (u,y)_+ + (x,v)_+.$$

In view of the sublinearity of the function $(\cdot,\cdot)_+$ with respect to its second variable, such a multimapping satisfies

$$(u, v - y)_+ + (x, y - v)_+ \geq 0.$$

Such a condition is close to the definition of accretivity ([3], [12], [61]) which is

$$(u - x, v - y)_+ \geq 0$$

for any $u, x \in X$, $v \in S(u)$, $y \in S(x)$. For that reason we say that a c-monotone multimapping is pseudo-accretif. Note that when the norm of X is Gâteaux-differentiable, the three preceding relations coincide and can be written

$$\langle j(u) - j(x), v - y \rangle \geq 0,$$

where $j : X \to X^b$ is the derivative of $(1/2) \|\cdot\|^2$, so that $j \circ S^{-1} : X \rightrightarrows X^b$ is monotone and pseudo-accretivity coincides with accretivity. Such a fact indicates that a number of results about maximal monotone operators may be extended to maximal pseudo-accretive operators.

Example 3. Let X be a lattice and let Y be the space of modular functions on X, a function $f : X \to \mathbb{R}$ being called *modular* if it satisfies

$$\forall u, x \in X \qquad f(u \wedge x) + f(u \vee x) = f(u) + f(x).$$

Let $c : X \times Y \to \mathbb{R}$ be the evaluation mapping given by $c(x, y) := y(x)$. When X has a smallest element, another choice consists in taking for Y the space of modular functions which are null at that element. In both cases, it seems of interest to study c-monotone operators from X to Y. The case of the subdifferential of a function $\varphi : X \to \mathbb{R}$ is studied in [18] when X is a distributive sublattice of $\mathcal{P}(E)$, where E is a finite set. See also [32], [43], [59] for some relationships with duality. The fact that one disposes of the Frank's discrete separation theorem ([32, p. 17, 111]), an analogue of the convex sandwich theorem and of studies of the relationships between such results and Fenchel-like duality is encouraging.

Example 4. Suppose X and Y are *m-convex* sets in the sense that X and Y are provided with maps $m_X : X^2 \to X$, $m_Y : Y^2 \to Y$. These maps can be considered as operations in X and Y respectively, as in the preceding example. When the spaces are metric spaces, such a notion has been widely studied in connection with metric convexity and geodesics (see [20] and its references): the space (X, d) is said to be mid-convex if for any $(x, x') \in X^2$ there exists a point $m := m_{(x,x')} \in X$ such that $d(m, x) = d(m, x') = (1/2)d(x, x')$. Let $f : Z \to \overline{\mathbb{R}}$ be m-convex in the sense that for $z = (x, y)$, $z' = (x', y') \in Z$ and $m_Z(z, z') := (m_X(x, x'), m_Y(y, y'))$ one has

$$f(m_Z(z, z')) \leq \frac{1}{2} f(z) + \frac{1}{2} f(z').$$

When Z is a normed vector space, when $m_Z(z, z') := (1/2)(z + z')$ and when f is continuous, then such a map is convex in the usual sense. If $f \geq c$ and if c is m-concave (i.e. $-c$ is m-convex) in each of its two variables, then $S := \{z : f(z) = c(z)\}$ is c-monotone. In fact, for $z = (x, y)$, $z' = (x', y') \in S$ one has

$$\frac{1}{2}c(x,y) + \frac{1}{2}c(x',y') = \frac{1}{2}f(x,y) + \frac{1}{2}f(x',y')$$
$$\geq f(m_Z(z,z')) \geq c(m_Z(z,z'))$$
$$\geq \frac{1}{4}c(x,y) + \frac{1}{4}c(x,y') + \frac{1}{4}c(x',y) + \frac{1}{4}c(x',y'),$$

so that $c(x,y) + c(x',y') \geq c(x',y) + c(x,y')$: S is c-monotone. This example slightly generalizes an argument in [29] and [41, Lemma 3], X and Y having no linear structure here. We refer to [41, Examples 1-4] for particular cases of this example.

Example 5. Let X and Y be arbitrary sets, and let $F : X \rightrightarrows Y$ be a relation. A number of duality schemes, in particular the radiant and shady dualities ([36, Example 4.2]), the sublevel duality ([36, Example 4.3]), are obtained by taking the coupling $c : X \times Y \to \mathbb{R}_{-\infty}$ given by

$$c(x,y) := -\iota_{F(x)}(y) := -\iota_F(x,y).$$

In such a case, the conjugate of a function $f \in \overline{\mathbb{R}}^X$ is given by

$$f^c(y) := -\inf\{f(x) : x \in F^{-1}(y)\}.$$

The simplicity of this conjugacy justifies its interest (see [37], [42], [58]...). In particular, introducing the polarity $P_c : \mathcal{P}(X) \to \mathcal{P}(Y)$ given by

$$P_c(A) := \{y \in Y : A \cap F^{-1}(y) = \varnothing\},$$

the conjugate f^c of f can be simply described by its sublevel sets given by

$$[f^c \leq r] = P_c([f < -r]) \qquad\qquad r \in \mathbb{R}.$$

A multimapping $S : X \rightrightarrows Y$ is c-monotone iff the following equivalent implications hold:

$$u,x \in X,\ (S(u) \cap F(x)) \setminus F(u) \neq \varnothing \quad \Rightarrow \quad S(x) \cap F(u) = \varnothing,$$
$$u,x \in X,\ S(u) \cap F(x) \neq \varnothing \quad \Rightarrow \quad S(x) \cap F(u) \subset F(x).$$

In several cases of interest F is defined by the way of a function $a : X \times Y \to \overline{\mathbb{R}}$ by

$$F(x) := \{y \in Y : a(x,y) > 0\};$$

then the preceding implications amount to

$$u,x \in X,\ v \in S(u),\ y \in S(x),\ a(u,y) > 0,\ a(x,v) > 0$$
$$\Rightarrow \quad a(u,v) > 0,\ a(x,y) > 0.$$

When X is a n.v.s. and Y is its dual space, with $a(x,y) := \langle x,y \rangle - 1$, one gets a condition in terms of half-spaces.

Example 6. Let W, X and Y be arbitrary sets, and let $p : X \times Y \to W$ be a map considered as an operation $(x,y) \mapsto xby := p(x,y)$. Given a function $d : W \to \mathbb{R}_{-\infty}$ one may consider the coupling function $c : X \times Y \to \mathbb{R}_{-\infty}$ given by

$$c(x,y) := d(p(x,y)) \qquad\qquad (x,y) \in X \times Y.$$

Two important cases have been considered ($\mathbb{P}$ denoting the set of positive numbers) in [26] and [53] respectively (see also [1]):
(a) for $W = X = Y = \mathbb{R}^n$, $p(x,y) := x + y$, $d(v) := \inf_{1\leq i\leq n} v_i$ for $v := (v_1, ..., v_n) \in \mathbb{R}^n$;
(b) for $W = X = Y = \mathbb{P}^n$, $p(x,y) := xy$, $d(v) := \inf_{1\leq i\leq n} v_i$ for $v := (v_1, ..., v_n) \in \mathbb{R}^n$, the product xy being the vector with components $x_i y_i$ for $x = (x_i)$, $y = (y_i)$ in $\mathbb{R}^n$.
Both cases correspond to separation properties (for closed downward sets and closed normal sets respectively) and have fruitful applications in mathematical economics.
Another example is provided by the following proposition in which $c : Z \to \mathbb{R}_{-\infty}$ is an arbitrary function and b is given by relation (6).

Proposition 6. *For any $f \in \mathcal{F} := \overline{\mathbb{R}}^Z$ the set $S := \{z \in Z : f(z) \leq c(z), f^b(z) \leq c(z)\}$ is c-monotone.*

Proof. Given $w := (u,v),\ z := (x,y) \in S$, we have $f(u,v) \leq c(u,v)$, hence

$$-c(x,y) \leq -f^b(x,y) \leq c(u,v) - [c(u,y) + c(x,v)]\,.$$

Taking in (2) $r := c(u,v)$, $s := c(x,y)$, $t := c(u,y) + c(x,v)$, we get $c(u,y) + c(x,v) \leq c(u,u) + c(x,y)$. □
Let us compare the polarity described in Proposition 2 with the polarity $S \mapsto S^\mu$ introduced to us by J.-E. Martínez-Legaz ([29]); it is given by

$$S^\mu := \{(x,y) \in Z : \forall (u,v) \in S,\ c(u,y) + c(x,v) \leq c(u,v) + c(x,y)\} \qquad (10)$$

for $S \in \mathcal{P}$.
The map $S \mapsto S^\mu$ is a polarity since for any family $(S_i)_{i\in I}$ in $\mathcal{P}$ it clearly satisfies

$$(\bigcup_{i\in I} S_i)^\mu = \bigcap_{i\in I} S_i^\mu.$$

Since $S \subset S^{\mu\mu}$, the following properties ensue:

$$S \subset T \Rightarrow T^\mu \subset S^\mu,$$
$$S^{\mu\mu\mu} = S^\mu.$$

Moreover, $S \mapsto S^\mu$ is designed in such a way that

$$S \subset S^\mu \Leftrightarrow S \text{ is } c\text{-monotone}.$$

Let us denote by $\mathcal{M}$ the class of c-monotone multimappings and by $\mathcal{M}'$ the class of *comonotone* multimappings defined by

$$S \in \mathcal{M}' \Leftrightarrow S^{\mu} \subset S.$$

We say that a multimapping S is *maximal c-monotone* if any c-monotone multimapping $T \supset S$ coincides with S. Observing that for a c-monotone operator S, one has

$$(x, y) \in S^{\mu} \Leftrightarrow S \cup \{(x, y)\} \text{ is } c\text{-monotone},$$

one gets, for a c-monotone operator S :

$$S \text{ is maximal } c\text{-monotone} \Leftrightarrow S = S^{\mu} \Leftrightarrow S \in \mathcal{M}' \cap \mathcal{M}.$$

It will be convenient to introduce the *Fitzpatrick map* $L : S \mapsto f_S := (\iota_S + c)^b$.

Definition 2. *The Fitzpatrick map* $L : \mathcal{P} \to \mathcal{F}$ *is given by* $L(S) = C(J(S))$, *where* $C : \mathcal{F} \to \mathcal{F}$ *is the conjugacy (or Legendre-Fenchel transform)* $C : f \mapsto f^b := f^b$.

In view of (4), it is a morphism of sup-lattices when, as usual, $\mathcal{P}$ is ordered by the inclusion and $\mathcal{F}$ is endowed with the pointwise order.

Proposition 7. *Taking for* $C : \mathcal{F} \to \mathcal{F}$ *the conjugacy* $f \mapsto f^b$, *the polarity* $P := M \circ L := M \circ C \circ J$ *on* $\mathcal{P}$ *coincides with the "monotone polarity"* $S \mapsto S^{\mu}$*: for any* $S \in \mathcal{P}$ *one has*

$$M(L(S)) = S^{\mu}. \tag{11}$$

Proof. For $S = \varnothing$, relation (11) is obvious, both sides being Z. Let $f_S := L(S)$ with $S \neq \varnothing$. Then, for any $(x, y) \in S^{\mu}$, $(u, v) \in S$, using (10) and taking $r := c(u, u)$, $s := c(x, y)$, $t := c(u, y) + c(x, v)$ in (3), we get

$$c(u, v) - [c(u, y) + c(x, v)] \geq -c(x, y),$$

so that we have

$$-f_S(x, y) := \inf\{c(u, v) - [c(u, y) + c(x, v)] : (u, v) \in S\} \geq -c(x, y).$$

Thus $(x, y) \in M(f_S)$ by definition of M in (5).
Conversely, if $(x, y) \notin S^{\mu}$ there exists $(u, v) \in S$ such that

$$c(u, v) + c(x, y) < c(u, y) + c(x, v). \tag{12}$$

Then $c(u, y) + c(x, v)$ is finite and either $c(x, y) = -\infty$ and the relation

$$-c(x, y) > c(u, v) - [c(u, y) + c(x, v)]$$

is trivial, else $c(x, y)$ is finite and this relation follows from (12). In both cases one gets $f_S(x, y) > c(x, y)$ and $(x, y) \notin M(f_S)$. Thus $M(f_S) = S^{\mu}$. □
Replacing S by S^{μ}, we get the next consequence.

Corollary 2. *For any $S \in \mathcal{P}$ one has*

$$M(L(S^\mu)) = S^{\mu\mu}.$$

Corollary 3. *(a) For any c-monotone operator S one has $S \subset S^{\mu\mu} \subset M(L(S))$.*
(b) For any $S \in \mathcal{P}$ one has $M(L(S)) = S$ if, and only if, S is maximal c-monotone.

Proof. (a) For $S \in \mathcal{M}$ we have seen that $S \subset S^{\mu\mu}$ and $S \subset S^\mu$, so that $S^{\mu\mu} = M(L(S^\mu)) \subset M(L(S))$.
(b) For any $S \in \mathcal{P}$ one has $M(L(S)) = S$ if, and only if, $S^\mu = S$ if, and only if, S is maximal c-monotone.

5 Representations

Let us study more closely the Fitzpatrick map and some classes of b-convex functions, a function f on Z being called a b-convex function if $f^{bb} = f$. Then we write $f \in \Gamma_b(Z)$.

Proposition 8. *For any S in $\mathcal{P}$, the function $f_S := L(S)$ satisfies $f_S \mid_{S^\mu} = c \mid_{S^\mu}$. If S belongs to the family $\mathcal{M}$ of nonempty c-monotone subsets of Z, then $f_S := L(S)$ belongs to the set*

$$\mathcal{F}_S := \{f \in \Gamma_b(Z) : f \geq c,\ f \mid_S = c \mid_S\}.$$

Moreover, S is maximal c-monotone, if, and only if, $f_S := L(S)$ belongs to the set

$$\mathcal{H}_S := \{f \in \Gamma_b(Z) : f \geq c,\ S = M(f)\}.$$

Proof. The first part of the proof of Proposition 7 has shown that $f_S(x, y) \leq c(x, y)$ for any $(x, y) \in S^\mu$. Since $f_S \geq c$ by Proposition 4, we get $f_S \mid_{S^\mu} = c \mid_{S^\mu}$. When S is c-monotone, the relation $f_S \in \mathcal{F}_S$ is a consequence of the inclusion $S \subset S^\mu$.
Finally, the last corollary has shown that if S is maximal c-monotone then $S = M(f_S)$ and that conversely, if $f_S \in \mathcal{H}_S$ then S is maximal monotone. □
When computing the c-regularization of a function is easier than computing its c-conjugate, one may introduce the c-regularized function $g_S := (c + \iota_S)^{bb}$ of $c_S := c + \iota_S$.

Proposition 9. *For any c-monotone operator S the function g_S satisfies $f_S \leq g_S \leq c_S$ and one has $f_S = g_S = c$ on S so that g_S belongs to $\mathcal{F}_S$.*
If S is maximal c-monotone one has $S = M(g_S) = \{z \in Z : g_Z(z) = c(z)\}$.

Proof. The inequality $g_S := c_S^{bb} \leq c_S$ is a general fact for any conjugacy. Since for a c-monotone operator S we observed that $f_S \mid_S = c \mid_S$, we have $f_S \leq c_S$ hence $g_S = f_S^b \geq c_S^b = f_S$. It follows that $f_S = g_S = c_S = c$ on S.
Now suppose S is maximal c-monotone. Then, since $f_S \leq g_S$ we have $M(g_S) \subset M(f_S) = S$; but since $g_S = c$ on S, we also have $M(g_S) = S$. □
It follows that any function f such that

$$f_S \leq f \leq g_S$$

belongs to the class $\mathcal{H}_S$ when S is maximal monotone. Examples show that it may be more convenient to deal with such a representative function than with the particular representatives function f_S and g_S. For instance, when S is the subdifferential of a closed convex function φ, one may take $f(x, y) := \varphi(x) + \varphi^b(y)$, since f_S and g_S may be difficult to compute (see [4], [9])

6 Questions and Observations

Let us conclude with some questions which may stimulate further research (since answers are not always available).
1) What is the image of the polarity $S \mapsto S^\mu$? An answer can be provided in the general framework of polarities: this image is the family $\mathcal{T} := \{T \in \mathcal{P} : T^{\mu\mu} = T\}$ of $\mu\mu$-closed subsets since $S^{\mu\mu\mu} = S^\mu$ for any $S \in \mathcal{P}$ and since any $T \in \mathcal{T}$ is of the form $T = S^\mu$ for $S := T^\mu$
2) What is the image of $\mathcal{M}$ by L?
3) What is the image $L(\partial^c f)$ of the c-subdifferential of a function f by L? The recent papers [4], [9] deal with such a question in the case of the classical coupling; even in that case the question is not trivial!
4) Is the relationship between the polarity P and the conjugacy $C : f \mapsto f^b$ richer than what is described above?
5) What is the operation obtained from an operation in $\mathcal{M}$ (such as sum, parallel addition...) by transporting it into $\mathcal{F}$? What is the operation obtained from an operation in $\mathcal{F}$ (such as sum, infimal convolution...) by transporting it into $\mathcal{P}$? Partial answers are provided in [44], [56] and [7].
6) If $f \in \Gamma(Z)$, then $M(f)$ is closed for the convergence which is the product of the bounded weak convergence with the strong convergence. What more can be said from a topological viewpoint?)
7) What are the generating functions of the preceding dualities?
8) What can be said about the corresponding subdifferentials?
9) Can one get special properties of the representations corresponding to the existence of operations such as $\cap$ and $\cup$ in $\mathcal{P}$ or convexification in $\mathcal{F}$?
10) What is the image of $\Gamma_b(Z)$ by M?

References

1. M. Akian and I. Singer, Topologies on lattice ordered groups, separation from closed downward sets and conjugations of type Lau, Optimization 52, No.6 (2003), 629-672.
2. E.J. Balder, An extension of duality-stability relations to nonconvex problems, SIAM J. Control Optim. 15 (1977), 329-343.
3. V. Barbu, *Nonlinear semigroups and differential equations in Banach spaces*, Editura Academiei, Bucharest, English transl. Noordhoff, Leyden (1976)
4. H.H. Bauschke, D.A. Maclaren and H.S. Sendov, Fitzpatrick functions: inequalities, examples and remarks on a problem of Fitzpatrick, J. Convex Anal, to appear.
5. G. Birkhoff, *Lattice Theory*, American Mathematical Society, Providence (1966).
6. J. Borwein, Maximality of sums of two maximal monotone operators, to appear, Proc. Amer. Math. Soc.
7. J. Borwein, Maximal monotonicity via convex analysis, to appear, J. Convex Anal.
8. R.M. Bot, S.I. Grad and G. Wanka, Maximal monotonicity for the precomposition with a linear operator, preprint, 2005.
9. R.S. Burachik and S. Fitzpatrick, On a family of convex functions associated to subdifferentials, J. Nonlin. and Convex Anal. 6 (2005), 165-171.
10. R. S. Burachik and B. F. Svaiter, Maximal monotone operators, convex functions and a special family of enlargements, Set-Valued Anal. 10, No.4 (2002), 297-316.
11. R. S. Burachik and B. F. Svaiter, Maximal monotonicity, conjugation and the duality product, Proc. Am. Math. Soc. 131, No.8 (2003), 2379-2383.
12. F.E. Browder, *Nonlinear operators and nonlinear equations of evolution in Banach spaces*, Proceedings of Symposia in Pure Mathematics. Vol. XVIII, Part 2. American Mathematical Society, Providence, R.I. (1976).
13. S. Dolecki and S. Kurcyusz, On Φ-convexity in extremal problems, SIAM J. Control Optim. 16 (1978), 277-300.
14. A. Eberhard, M. Nyblom and D. Ralph, Applying generalised convexity notions to jets, in *Generalized Convexity, Generalised Monotonicity: Recent Results*, J.-P. Crouzeix et al. (eds) Kluwer, Dordrecht (1998), 111-157.
15. S. Fitzpatrick, Representing monotone operators by convex functions, *Workshop and Miniconference on Functional Analysis and Optimization* (Canberra, 1988), Austral. Nat. Univ., Canberra, (1988), 59–65.
16. F. Flores-Bazán, On a notion of subdifferentiability for non-convex functions, Optimization 33 (1995), 1-8.
17. F. Flores-Bazán and J.E. Martínez-Legaz, Simplified global optimality conditions in generalized conjugation theory, in *Generalized Convexity, Generalized Monotonicity*, J.-P. Crouzeix, J.E. Martínez-Legaz and M. Volle, eds, Kluwer, Dordrecht (1998), 303-329.
18. S. Fujishige, On the subdifferential of a submodular function, Math. Prog. 29 (1984), 348-360.
19. R. Goebel and R.T. Rockafellar, Generalized conjugacy in Hamilton-Jacobi theory for fully convex Lagrangians, J. Convex Anal. 9, No.2, 463-473 (2002).
20. M. Gromov, *Metric Structures for Riemannian and non-Riemannian Spaces*, Progress in Mathematics #152, Birkhäuser, Boston, Mass. (1999).
21. A.D. Ioffe, Abstract convexity and non-smooth analysis, Kusuoka, S. et al.(eds.), Adv. Math. Econ. 3 (2001), 45-61.

22. E. Krauss, A representation of maximal monotone operators by saddle functions, Rev. Roum. Math. Pures Appl. 30, 823-836 (1985).
23. E. Krauss, A representation of arbitrary maximal monotone operators via subgradients of skew-symmetric saddle functions, Nonlinear Anal., Theory Methods Appl. 9, 1381-1399 (1985).
24. E. Krauss, Maximal monotone operators and saddle functions. I., Z. Anal. Anwend. 5, 333-346 (1986).
25. J.-E. Martínez-Legaz, Fenchel duality and related properties in generalized conjugation theory, Southeast Asian Bull. Math. 19, (1995), 99-106.
26. J.-E. Martínez-Legaz, A.M. Rubinov and I. Singer, Downward sets and their separation and approximation properties, J. Glob. Optim. 23, No.2 (2002), 111-137.
27. J.-E. Martínez-Legaz and I. Singer, Dualities between complete lattices, Optim. 21 (1990), 481-508.
28. J.-E. Martínez-Legaz and I. Singer, Subdifferentials with respect to dualities, ZOR, Math. Methods Oper. Res. 42, No.1, 109-125 (1995).
29. J.-E. Martínez-Legaz and B.F. Svaiter, Monotone operators representable by l.s.c. convex functions, Set-Valued Anal. 13, No.1, 21-46 (2005).
30. J. E. Martinez-Legaz and M. Théra, A convex representation of maximal monotone operators, J. Nonlinear Convex Anal. 2 (2001), 243–247.
31. J.-J. Moreau, Inf-convolution, sous-additivité, convexité des fonctions numériques, J. Math. Pures et Appl. 49 (1970), 109-154.
32. K. Murota, *Discrete Convex Analysis*, SIAM Monographs on Discrete Mathematics and Applications, SIAM, Philadelphia (2003).
33. O. Ore, Galois connexions, Trans. Amer. Math. Soc. 55 (1944), 493-513.
34. D. Pallaschke and S. Rolewicz, *Foundations of Mathematical Optimization, Convex Analysis without Linearity*, Kluwer, Dordrecht (1998).
35. J.-P. Penot, Duality for radiant and shady programs, Acta Math. Vietnamica 22 (2) (1997), 541-566.
36. J.-P. Penot, What is quasiconvex analysis?, Optimization 47 (2000), 35-110.
37. J.-P. Penot, Duality for anticonvex programs, J. of Global Optim. 19 (2001), 163-182.
38. J.-P. Penot, Augmented Lagrangians, duality and growth conditions, J. Nonlinear Convex Anal. 3 (3) (2002), 283-302.
39. J.-P. Penot, Rotundity, smoothness and duality, Control and Cybernetics 32 (4) (2003), 711-733.
40. J.-P. Penot, Is convexity useful for the study of monotonicity? in "*Nonlinear Analysis and Applications*", R. P. Agarwal and D. O'Regan, eds., Kluwer, Dordrecht (2003), 807-822.
41. J.-P. Penot, Autoconjugate functions and representations of monotone operators, Bull. Aust. Math. Soc. 67 (2003), 277-284.
42. J.-P. Penot, The bearing of duality on microeconomics, Adv. Math. Economics 7 (2005), 113-139.
43. J.-P. Penot, Unilateral analysis and duality, in *GERAD, Essays and Surveys in Global Optimization*, G. Savard et al eds., Springer, New York (2005), 1-37.
44. J.-P. Penot, The relevance of convex analysis for the study of monotonicity, Nonlinear Anal. TMA, 58, (7-8), (2004), 855-871.
45. J.-P. Penot, Natural closure, natural compositions and natural sums of monotone operators, preprint, Univ. Pau (2002).

46. J.-P. Penot, C. Zălinescu, On the convergence of maximal monotone operators, Proc. Amer. Math. Soc. (to appear).
47. J.-P. Penot, C. Zălinescu, Some problems about the representation of monotone operators by convex functions, ANZIAM J. 47 (2005), 1-20.
48. R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Springer, Berlin, 1997.
49. S. Rolewicz, Convex analysis without linearity, Control and Cybern. 23 (1994), 247-256.
50. S. Rolewicz, Duality and convex analysis in the absence of linear structure, Math. Jap. 44, No.1, 165-182 (1996).
51. A.M. Rubinov, *Abstract Convexity and Global Optimization*, Kluwer, Dordrecht (2000).
52. A.M. Rubinov, Radiant sets and their gauges, in *Quasidifferentiability and Related Topics*, V. Demyanov and A.M. Rubinov, eds. Kluwer, Dordrecht (2000).
53. A.M. Rubinov, B.M. Glover and V. Jeyakumar, A general approach to dual characterizations of solvability of inequality systems and applications, J. Convex Anal. 2(1995), 309-344.
54. A. Schrijver, *Combinatorial Optimization. Polyhedra and Efficiency,* Algorithms and Combinatorics 24. Springer, Berlin (2003).
55. S. Simons, C. Zalinescu, A new proof for Rockafellar's characterization of and maximal monotone operators, Proc. Amer. Math. Soc. 132 (2004), 2969-2972.
56. S. Simons, C. Zalinescu, Fenchel duality, Fitzpatrick functions and maximal monotonicity, J. Nonlinear Convex Anal. 6 (1), (2005), 1-22.
57. I. Singer, *Abstract Convex Analysis*, Canadian Math. Soc. Series of Monographs and Advanced Texts. Wiley, New York, New York (1997).
58. P.T. Thach, A nonconvex duality with zero gap and applications, SIAM J. Optim. 4 (1) (1994), 44-64.
59. D.M. Topkis, *Supermodularity and Complementarity*, Princeton Univ. Press, Princeton (1998).
60. M. Volle, Conjugaison par tranches, Annali Mat. Pura Appl. 139, pp. 279-312 (1985).
61. I.I. Vrabie, *Compactness methods for nonlinear evolutions*, Pitman Monographs and Surveys in Pure and Applied Mathematics, 32, Longman, Harlow (1987).
62. C. Zalinescu, A new proof of the maximal monotonicity of the sum using the Fitzpatrick function, in *Variational Analysis and Applications*, F. Giannessi and A. Maugeri, eds., Springer, New York, 2005, 1159-1172 .

On Variational-like Inequalities with Generalized Monotone Mappings*

Vasile Preda[1], Miruna Beldiman[2], and Anton Bătătorescu[3]

[1] University of Bucharest, Faculty of Mathematics and Computer Science, 14 Academiei str., 10014 Bucharest, Romania. preda@fmi.unibuc.ro
[2] University of Bucharest, 14 Academiei str., 10014 Bucharest, Romania. miruna.m@gmail.com
[3] University of Bucharest, 14 Academiei str., 10014 Bucharest, Romania. batator@fmi.unibuc.ro

Summary. We consider two new classes of generalized relaxed α-monotone and semimonotone functions and using the KKM technique we prove the existence of solutions for variational-like inequalities relative to these types of mappings in Banach spaces. Several examples and special cases are also considered.

Key words: Variational-like inequality, generalized monotone mapping, KKM Theorem, coercivity, semicontinuity.

1 Introduction

The monotonicity property of a map, together with continuity, convexity and coercivity, has a very important role in many fields like optimization and mathematical programming problems (Luc [19]), equilibrium problems (Ansari, Konnov, Yao [2]), game theory and variational inequality theory (Glowinski, Lions, Tremolieres [8]). Many authors obtained interesting generalizations of this notion and used them for establish the existence conditions for some types of variational inequalities (Hadjisavvas, Schaible [10], Hartman, Stampacchia [11], Giannessi [6], Giannessi, Maugeri [7], Kassay, Kolumban [13], Schaible [24]).
In 1995, Konnov and Yao proved in [14] some results about the existence of solutions for vector variational inequalities with C_x-pseudomonotone set-valued mappings, which were been extended later by Ansari, Siddiqi and Yao in [1]. Also, Konnov [15, 16, 17] obtained some combined relaxation methods for solving variational inequalities which involve different classes of generalized monotone functions.

*This research was partially supported by Grant CNCSIS no. 27694/2005

In 1997, Verma [25], studied a class of nonlinear variational inequalities with p-monotone and p-Lipschitz maps in reflexive Banach spaces. In 1999, Chen [4] obtained the existence of solution for a class of variational inequalities with semi-monotone single-valued maps in nonreflexive Banach spaces. In 2003, Fang and Huang [5], considered two classes of variational-like inequalities with generalized monotone and semi-monotone mappings. Using the KKM technique, they proved the existence of the solutions for these variational-like inequalities with relaxed $\eta-\alpha$ monotone mappings in reflexive Banach spaces. In this case, the following problems were considered:

$$\text{Find } x \in K \text{ such that } \langle Tx, \eta(y,x)\rangle + f(y) - f(x) \geq 0, \ \forall\, y \in K,$$

and

$$\text{Find } x \in K \text{ such that } \langle Tx, \eta(y,x)\rangle + f(y) - f(x) \geq \alpha(y-x), \ \forall\, y \in K,$$

where T, η, f and α will be defined later in section 3.
The solvability of variational-like inequalities with relaxed $\eta-\alpha$ semimonotone mappings in arbitrary Banach spaces were also studied by means of the Kakutani-Fan-Glicksberg fixed-point theorem. In this case, the following problems were considered:

$$\text{Find } u \in K \text{ such that } \langle A(u,u), \eta(v,u)\rangle + f(v) - f(u) \geq 0, \ \forall\, v \in K,$$

and

$$\text{Find } u \in K \text{ such that } \langle A(u,u), \eta(v,u)\rangle + f(v) - f(u) \geq \alpha(v-u), \ \forall\, v \in K,$$

where A, η, f, and α will also be defined later in section 4.
In this way, some previous results concerning variational inequalities were extended, among such references we mention Chang, Lee, Chen, [3], Chen [4], Goeleven, Motreanu [9], Hartman, Stampacchia [11], Siddiqi, Ansari, Kazmi [22], and Verma [25, 26].
Kang, Huang and Lee extended in 2003 [12] these notions for the case of set-valued mappings.
In this paper, we define two classes of generalized relaxed α-monotone and semi-monotone mappings and show by several examples the importance of these types of functions.
Using the KKM technique, we state the existence theorems for variational-like inequalities with generalized relaxed α-monotone mappings in reflexive Banach spaces. Further, by employing the Kakutani-Fan-Glicksberg fixed-point theorem, we establish also the solvability of variational-like inequalities with generalized relaxed αsemimonotone mappings in arbitrary Banach spaces.
Our paper extends and improves, at least, some known results relative to:

- more general variational inequality classes as presented before;
- more general classes of monotone, respectively semimonotone mappings;

- the convexity and the lower semicontinuity assumptions for the mappings $y \longmapsto \langle Tz, \eta(y,x)\rangle$, $y \longmapsto \langle A(z,w), \eta(y,x)\rangle$ and f are replaced by more general one, which extend even the usual convexity and semicontinuity assumptions of the mappings $y \longmapsto \langle Tz, \eta(y,x)\rangle + f(y) - f(x)$, and $y \longmapsto \langle A(z,w), \eta(y,x)\rangle + f(y) - f(x)$. Our results are sustained by some significant examples given in sections 2 and 5.

2 Definitions and Some Preliminaries

We consider in the sections 2 and 3 the real reflexive Banach space E and its dual space E^*, and let K be a nonempty subset of E. We consider the mapping $T : K \to E^*$ and the functions $\Psi : K \times K \times K \to \mathbb{R}$ and $\alpha : E \times E \to \mathbb{R}$.

Definition 1. *Ψ is general relaxed α-monotone if for any $x, y \in K$ we have*

$$\Psi(y,x;y) - \Psi(y,x;x) \geq \alpha(x,y)$$

where $\lim\limits_{t \searrow 0} \frac{\alpha(x,x+t(y-x))}{t} = 0$.

Definition 2. *If $\Psi(x,y;z) = \langle Tz, \eta(x,y)\rangle$, where $\eta : K \times K \to E$, we say that the mapping T is general $\eta - \alpha$ monotone.*

Remark 1.

($i1$) If $\Psi(y,x;z) = \langle Tz, \eta(y,x)\rangle$ with $\alpha(x,y) = \beta(y-x)$, where $\beta : K \to \mathbb{R}$ with $\beta(tz) = t^p \beta(z)$ for $t > 0$, $p > 1$ and $\eta : K \times K \to E$, the Definition 1 reduces to relaxed $\eta - \alpha$ monotonicity of mapping T (see [5]).

($i2$) In the case of ($i1$), if $\eta(x,y) = x - y$ for all $x, y \in K$, the Definition 1 reduces to

$$\langle Ty - Tx, y - x\rangle \geq \beta(y-x), \ \forall\, x, y \in K,$$

and T is said to be relaxed α monotone (see also [5]).

($i3$) In the case of ($i2$), if $\beta(z) = k\|z\|^p$, where $k > 0$ is a constant, then Definition 1 reduces to

$$\langle Ty - Tx, y - x\rangle \geq k\|x-y\|^p, \ \forall\, x, y \in K,$$

and T is said to be p -monotone (see [4, 25]).

($i4$) We see that every monotone mapping is relaxed $\eta - \alpha$ monotone with $\eta(x,y) = x - y$ for all $x, y \in K$ and $\alpha \equiv 0$.

Example 1. If we consider $E = E^* = \mathbb{R}$, $K = (-\infty, +\infty)$,

$$\Psi(x,y;z) = \langle Tz, \eta(x,y)\rangle = -z\eta(x,y)$$

$$\eta(y,x) = \begin{cases} c(y-x), & x \leq y \\ -c(y-x), & x > y \end{cases}$$

where $c > 0$ is a constant, then Ψ is general relaxed α-monotone with $\alpha(x,y) = \beta(y-x)$,

$$\beta(z) = \begin{cases} -cz^2, & z < 0 \\ cz^2, & z \geq 0 \end{cases}$$

i.e., the mapping T is relaxed $\eta - \alpha$ monotone.

Example 2. If we consider $E = E^* = \mathbb{R}$, $K = (-\infty, +\infty)$,

$$\Psi(x,y;z) = \langle Tz, \eta(x,y)\rangle = -z^2\eta(x,y)$$

$$\eta(y,x) = \begin{cases} -c\left(y^2 - x^2\right), & y^2 \geq x^2 \\ c\left(y^2 - x^2\right), & y^2 < x^2 \end{cases}$$

where $c > 0$ is a constant, then Ψ is general relaxed α-monotone with

$$\alpha(x,y) = \begin{cases} -c\left(y^2 - x^2\right)^2, & y^2 < x^2 \\ c\left(y^2 - x^2\right)^2, & y^2 \geq x^2 \end{cases}$$

i.e., the mapping T is relaxed $\eta - \alpha$ monotone.

We see that in this case T is not relaxed $\eta - \alpha$ monotone with $\alpha(x,y)$ given by $\alpha(x,y) = \beta(y-x)$. Hence, the class of general relaxed $\eta - \alpha$ monotone mappings is more large than the class of relaxed $\eta - \alpha$ monotone mappings defined by Fang, Huang [5].

Example 3. For $K = (0, \pi)$ and $\alpha(x,y) = \left[\left(\sin^2 x\right)^x - \left(\sin^2 y\right)^y\right]^\tau$ where $\tau > 1$ is a constant, we have $\lim_{t \searrow 0} \frac{\alpha(x, x+t(y-x))}{t} = 0$. In this case we see that $\alpha(x,y)$ has not the form $\beta(\gamma(y) - \gamma(x))$, where $\gamma : K \to K$.

Definition 3. *$\Psi(y,x;\cdot)$ is hemicontinuous if for any fixed $x, y \in K$, the mapping $\mu : [0,1] \to (-\infty, +\infty)$ defined by $\mu(t) = \Psi(y,x;x+t(y-x))$ is continuous at 0^+.*

Definition 4. *We say that Ψ is coercive if there exists $y_0 \in K$ such that*

$$\lim_{\|x\| \to +\infty} \frac{\Psi(x,y_0;x) - \Psi(x,y_0;y_0)}{|\Psi(y_0,x;y_0)|} = +\infty$$

Remark 2.

(*i*1) We see that for $\Psi(y,x;z) = \langle Tz, \eta(y,x)\rangle$ such that $\Psi(y_0,x;y_0) = c\eta(y_0,x)$, where c is a non-zero real constant, then Definition 4 reduces to η -coercivity of the mapping T (see Schaible [23]).

(*i*2) For $\Psi(y,x;z) = \langle Tz, \eta(y,x)\rangle + f(z)$, where $f : K \to \mathbb{R} \cup \{+\infty\}$ is a proper function such that $\Psi(y_0,x;y_0) = c\eta(y_0,x)$, c is a non-zero real constant, then Definition 4 reduces to η -coercivity of the mapping T with respect to f, defined in [5]. In this case, if $f = \delta_K$, where δ_K is the indicator function of K, then Definition 4 coincides with the definition of η -coercivity in the sense of Schaible [23].

$(i3)$ For $\Psi(y,x;z) = \langle Tz, \eta(y,x)\rangle + \varphi(y,x)$ where $\varphi : K \times K \to \mathbb{R} \cup \{+\infty\}$ is a proper function such that $\Psi(y_0,x;y_0) = c\eta(y_0,x)$, with c a non-zero real constant, then Definition 4 reduces also to η -coercivity of the mapping T (see Schaible [23]).

Definition 5. *(Ky Fan [18]) A mapping $F : K \to 2^E$ is said to be a KKM mapping if for any $\{x_1,\ldots,x_n\} \subset K$, we have $co\{x_1,\ldots,x_n\} \subset \bigcup_{i=1}^{n} F(x_i)$, where 2^E denotes the family of all the nonempty subsets of E.*

Lemma 1. *(Ky Fan [18]) Let K be a nonempty subset of a Hausdorff topological vector space X and let $F : K \to 2^X$ be a KKM mapping. If $F(x)$ is closed in X for every $x \in K$ and compact for some $x \in K$, then $\bigcap_{x\in K} F(x) \neq \emptyset$.*

3 Variational-like Inequalities with General Relaxed α-Monotone Mappings

In this section we suppose that K is a nonempty closed convex subset of E and now we consider the following problems:

$$\text{Find } x \in K \text{ such that } \Psi(y,x;x) \geq 0,\ \forall\ y \in K \tag{1}$$

and

$$\text{Find } x \in K \text{ such that } \Psi(y,x;y) \geq \alpha(x,y),\ \forall\ y \in K. \tag{2}$$

Relative to problems (1) and (2) we have the following results.

Theorem 1. *We suppose:*

(i1) $\Psi(y,x;\cdot)$ is hemicontinuous for any fixed $x,y \in K$;
(i2) $\Psi(\cdot,x;z)$ is a convex function on K, for any fixed $x,z \in K$;
(i3) $\Psi(x,x;z) = 0$ for any $x,z \in K$;
(i4) Ψ is general relaxed α-monotone.

Then the problems (1) and (2) are equivalent.

Proof. Let x be a solution for (1). Then

$$\Psi(y,x;x) \geq 0, \ \text{ for any } y \in K. \tag{3}$$

According to (i4) $\Psi(y,x;\cdot)$ is α-monotone and then

$$\Psi(y,x;y) - \Psi(y,x;x) \geq \alpha(x,y) \ \text{ for any } y \in K,$$

and from (3) we get

$$\Psi(y,x;y) \geq \alpha(x,y) \ \text{ for any } y \in K,$$

i.e. x is a solution for problem (2).
Conversely, let $x \in K$ be a solution of problem (2) and let $y \in K$. We denote

$$y_t = (1-t)\,x + ty\ ,\quad t \in (0,1)$$

and by convexity of K we have $y_t \in K$. Since x is a solution of (2) we have

$$\Psi\,(y,x;y) \geq \alpha\,(x,y) \quad \text{for any } y \in K.$$

Hence, for $y = y_t$, $t \in (0,1)$, we get

$$\Psi\,(y_t,x;y_t) \geq \alpha\,(x,y_t)\,. \tag{4}$$

By (i2) we have that $\Psi\,(\cdot,x;y_t)$ is a convex function. Hence

$$\Psi\,(y_t,x;y_t) \leq t\Psi\,(y,x;y_t) + (1-t)\,\Psi\,(x,x;y_t)\,. \tag{5}$$

Using (4), (5) and (i3) we obtain

$$\Psi\,(y,x;y_t) \geq \frac{\alpha\,(x,y_t)}{t}\ ,\quad \text{for any } t \in (0,1)\,. \tag{6}$$

Since $\Psi\,(y,x;\cdot)$ is hemicontinuous (according to (i1)) and $\lim\limits_{t\searrow 0} \frac{\alpha(x,x+t(y-x))}{t} = 0$ (according to (i4)), letting $t \to 0$ in (6) we get $\Psi\,(y,x;x) \geq 0$, for all $y \in K$.

Remark 3. We notice that Theorem 2.1 of Fang, Huang [5], as well as Theorem 2.1 of Verma [25] are particular cases of Theorem 1.

Theorem 2. *Let K be a nonempty bounded closed convex subset of a real reflexive Banach space E; let E^* be the dual space of E. We assume that:*

(j1) $\Psi\,(y,x;\cdot)$ *is hemicontinuous for any fixed* $x,y \in K$;
(j2) $\Psi\,(\cdot,x;z)$ *is a convex and lower semicontinuous function on* K, *for any fixed* $x,z \in K$;
(j3) $\Psi\,(x,y;z) + \Psi\,(y,x;z) = 0$ *for all* $x,y,z \in K$;
(j4) $\Psi\,(y,x;\cdot)$ *is* α*-monotone with* $\lim\limits_{t\searrow 0} \frac{\alpha(x,x+t(y-x))}{t} = 0$;
(j5) $\alpha\,(\cdot,y)$ *is weakly lower semicontinuous for any fixed* $y \in K$, *i.e., for any sequence* $\{x_\nu\}_\nu$ *that converges to* x *in* $\sigma\,(E,E^*)$ *we have*

$$\alpha\,(x,y) \leq \lim_{\nu\to\infty} \inf \alpha\,(x_\nu,y)\ ,\quad \textit{for any } y \in K.$$

Then, problem (1) is solvable.

Proof. We define two set-valued mappings $F,G : K \to 2^E$ as follows:

$$F\,(y) = \{x \in K \mid \Psi\,(y,x;x) \geq 0\}\ ,\quad \forall\, y \in K,$$
$$G\,(y) = \{x \in K \mid \Psi\,(y,x;y) \geq \alpha\,(x,y)\}\ ,\quad \forall\, y \in K.$$

According to Theorem 1 we have

$$\bigcap_{y \in K} F(y) = \bigcap_{y \in K} G(y). \tag{7}$$

We shall prove that

$$\bigcap_{y \in K} G(y) \neq \emptyset. \tag{8}$$

We claim first that F is a KKM mapping. We proceed by contradiction and suppose that there exist $\{y_1, \ldots, y_n\} \subset K$ and $t_i > 0$, $i = 1, 2, ..., n$, $\sum_{i=1}^{n} t_i = 1$, such that $y^0 = \sum_{i=1}^{n} t_i y_i \notin \bigcup_{i=1}^{n} F(y_i)$. It follows $y^0 \notin F(y_i)$ for any $i = 1, 2, ..., n$, i.e.

$$\Psi\left(y_i, y^0; y^0\right) < 0\ , \quad \text{for } i = 1, 2, ..., n. \tag{9}$$

Using (j3) we obtain

$$\Psi(y, y; z) = 0 \text{ for any } y, z \in K.$$

Now, by (j2) and (9) it follows

$$0 = \Psi\left(y^0, y^0; y^0\right) = \Psi\left(\sum_{i=1}^{n} t_i y_i, y^0; y^0\right) \leq$$

$$\leq \sum_{i=1}^{n} t_i \Psi\left(y_i, y^0; y^0\right) < 0$$

which is a contradiction. Hence, F is a KKM mapping.

We prove now that G is also a KKM mapping. It is sufficient to prove that for any $y \in K$ we have

$$F(y) \subset G(y).$$

Let $y \in K$. For $x \in F(y)$ we have $\Psi(y, x; x) \geq 0$. Since Ψ is general relaxed α-monotone, we have

$$\Psi(y, x; y) \geq \Psi(y, x; x) + \alpha(x, y) \geq \alpha(x, y)$$

i.e. $x \in G(y)$, hence $F(y) \subset G(y)$ for all $y \in K$ and therefore G is a KKM mapping.

We prove now that $G(y)$ is weakly compact in K for each $y \in K$. Indeed, according to the definition of $G(y)$ and by (j2) we have that the mapping $x \mapsto \Psi(x, y; y)$ is weakly lower semicontinuous. Using the definition of G and the weakly lower semicontinuity of $\alpha(\cdot, y)$ for all $y \in K$, we conclude that $G(y)$ is weakly closed for all $y \in K$. Since K is a bounded closed and convex set, it follows that K is weakly compact, and so $G(y)$ is weakly compact in K for all $y \in K$. Using now Lemma 1 we get (9). Hence, by (7) and (8) we get

$$\bigcap_{y \in K} F(y) \neq \emptyset,$$

and therefore, there exists $x \in K$ such that

$$\Psi(y, x; x) \geq 0\ , \quad \text{for all } y \in K,$$

i.e., problem (1) is solvable and the theorem is proved.

Remark 4. Theorem 2 includes as particular cases, for example, the Theorem 2.2 of Fang and Huang [5] and Theorem 2.2 of Verma [25].

We consider now the case of unbounded closed convex sets.

Theorem 3. *Let K be a nonempty unbounded closed convex subset of a real Banach space E and let E^* be the dual space of E. We assume that (j1), (j2), (j3) and (j5) of Theorem 2 are fulfilled together with*

(j6) Ψ is coercive.

Then, problem (1) is solvable.

Proof. For a positive real number r, we define

$$B_r = \{y \in E \mid \ \|y\| \leq r\},$$

and we consider the following problem:

$$\text{Find } x_r \in K \cap B_r \text{ such that } \Psi(y, x_r; x_r) \geq 0, \text{ for all } y \in K \cap B_r. \tag{10}$$

According to Theorem 2 we have that the problem (10) has a solution $x_r \in K \cap B_r$. We show that there exists $r' > 0$ such that $\|x_{r'}\| < r'$. If $\|x_r\| = r$ for any $r > 0$, then we chose r_0 such that $r_0 > \|y_0\|$, where y_0 is given by the coercivity condition. In this case we have

$$\Psi(y_0, x_{r_0}; x_{r_0}) \geq 0 \tag{11}$$

On the other hand, by (j3) we can write

$$\Psi(y_0, x_{r_0}; x_{r_0}) = -\Psi(x_{r_0}, y_0; x_{r_0}) =$$

$$= -[\Psi(x_{r_0}, y_0; x_{r_0}) - \Psi(x_{r_0}, y_0; y_0)] - \Psi(x_{r_0}, y_0; y_0) =$$
$$= -[\Psi(x_{r_0}, y_0; x_{r_0}) - \Psi(x_{r_0}, y_0; y_0)] + \Psi(y_0, x_{r_0}; y_0) =$$
$$\leq -[\Psi(x_{r_0}, y_0; x_{r_0}) - \Psi(x_{r_0}, y_0; y_0)] + |\Psi(y_0, x_{r_0}; y_0)| =$$
$$= -|\Psi(y_0, x_{r_0}; y_0)| \left[\frac{\Psi(x_{r_0}, y_0; x_{r_0}) - \Psi(x_{r_0}, y_0; y_0)}{|\Psi(y_0, x_{r_0}; y_0)|} - 1\right]$$

Now, we can choose r large enough so that the last inequality and the coercivity of Ψ imply $\Psi(y_0, x_{r_0}; x_{r_0}) < 0$, which contradicts (11).

Thus we conclude that there exists r' such that $\|x_{r'}\| < r'$. It follows easily that for any $y \in K$ we can choose ε such that $0 < \varepsilon < 1$ and $x_{r'} + \varepsilon (y - x_{r'}) \in K \cap B_{r'}$. Using now (10), (j2) and (j3), we obtain for any $y \in K$,

$$0 \leq \Psi (x_{r'} + \varepsilon (y - x_{r'}), x_{r'}; x_{r'}) = \Psi (\varepsilon y + (1 - \varepsilon) x_{r'}, x_{r'}; x_{r'}) \leq$$

$$\leq \varepsilon \Psi (y, x_{r'}; x_{r'}) + (1 - \varepsilon) \Psi (x_{r'}, x_{r'}; x_{r'}) = \varepsilon \Psi (y, x_{r'}; x_{r'})$$

i.e., $\Psi (y, x_{r'}; x_{r'}) \geq 0$, for all $y \in K$.

Remark 5. We note that the above results remain also true in the case when we consider $\Psi : K \times K \times K \to \mathbb{R} \cup \{+\infty\}$ and we suppose that the mapping $x \longmapsto \Psi (x, y, z)$ is properly convex instead of convex on K, for any fixed $y, z \in K$.

Remark 6. Theorem 3 includes as particular cases the Theorem 2.3 of Fang and Huang [5] and Theorem 2.3 of Verma [25].

4 Variational-like Inequalities with General-Relaxed α^*-Semimonotone Mappings

In this section we consider an arbitrary Banach space E. We denote by E^* the dual space of E and by E^{**} the dual space of E^*. Let K be a nonempty closed convex subset of E^{**}, and $\Psi^* : K \times K \times K \times K \to \mathbb{R}$.
We consider the following problem:

$$\begin{array}{l} \text{Find } u \in K \text{ such that} \\ \qquad \Psi^* (v, u; u, u) \geq 0 \ \text{ for all } v \in K. \end{array} \tag{12}$$

Definition 6. *Let $\alpha^* : E^{**} \times E^{**} \to \mathbb{R}$ be a mapping such that*

$$\lim_{t \searrow 0} \frac{\alpha^* (u, u + t (v - u))}{t} = 0.$$

We say that Ψ^* is general relaxed α^*-semimonotone if the following conditions hold:

(a) for each fixed $u \in K$, $\Psi^* (\cdot , \cdot ; u, \cdot)$ is general relaxed α^*-monotone, i.e.

$$\Psi^* (v, w; u, v) - \Psi^* (v, w; u, w) \geq \alpha^* (v, w), \ \ \textit{for all } v, w \in K;$$

(b) for each fixed $u, v, w \in K$, $\Psi^* (v, u; \cdot , w)$ is completely continuous, i.e., for any sequence $\{z_n\}_n$ that converges to z_0 in $\sigma (E^{**}, E^*)$, the sequence $\{\Psi^* (v, u; z_n, w)\}_n$ converges to $\Psi^* (v, u; z_0, w)$.

Let $A : K \times K \to E^*$, $\eta^* : K \times K \to E^{**}$ and $\beta^* : E^{**} \to \mathbb{R}$, such that $\beta^*(tz) = t^p \beta^*(z)$ for all $t > 0$, $z \in E^{**}$, where the real number $p > 1$ is a constant.
If $\Psi^*(x, y; z, w) = \langle A(z, w), \eta^*(x, y) \rangle$, then the Definition 6 reduces to the following

Definition 7. *(Fang, Huang [5, Definition 3.1]) The mapping A is said to be relaxed $\eta^* - \beta^*$ semimonotone if the following conditions hold:*

- *for each fixed $u \in K$, the mapping $A(u, \cdot)$ is relaxed $\eta^* - \beta^*$ monotone, i.e.,*

$$\langle A(u, v) - A(u, w), \eta^*(u, v) \rangle \geq \beta^*(v - w), \text{ for all } v, w \in K;$$

- *for each fixed $v \in K$, the mapping $A(\cdot, v)$ is completely continuous, i.e. for any sequence $\{u_n\}_n$ that converges to u_0 in $\sigma(E^{**}, E^*)$ implies that $\{A(u_n, v)\}_n$ converges to $A(u_0, v)$ in the norm topology of E^*.*

If $f : K \to \mathbb{R}$ is convex and lower semicontinuous and

$$\Psi^*(x, y; z, w) = \langle A(z, w), \eta^*(x, y) \rangle + f(x) - f(y),$$

then the problem (12) reduces to the problem (9) of Fang, Huang [5].

Theorem 4. *Let E be a real Banach space and let $K \subset E^{**}$ be a nonempty bounded closed convex set. We assume that:*

(i1) Ψ^* *is general relaxed α^*-semimonotone mapping;*
(i2) $x \longmapsto \Psi^*(x, y; z, w)$ *is a convex and lower semicontinuous mapping, for any fixed $y, z, w \in K$;*
(i3) $\Psi^*(x, y; z, \cdot) : K \to R$ *is finite dimensional continuous for any fixed $x, y, z \in K$, i.e., $\Psi^*(x, y; z, \cdot) : K \cap F \to R$ is continuous for any finite dimensional subspace $F \subset E^{**}$;*
(i4) $\alpha^*(v, \cdot)$ *is convex and lower semicontinuous for any fixed $v \in K$;*
(i5) $\Psi^*(x, y; z, w) + \Psi^*(y, x; z, w) = 0$ *for any $x, y, z, w \in K$.*

Then, the problem (12) is solvable.

Proof. Let $F \subset E^{**}$ be a finite dimensional subspace with $K_F = F \cap K \neq \emptyset$. For each $w \in K$ we consider the following problem:

$$\begin{array}{l} \text{Find } u_0 \in K_F \text{ such that} \\ \qquad \Psi^*(v, u_0; w, u_0) \geq 0 \text{ for all } v \in K. \end{array} \tag{13}$$

By (i2) - (i4) and Theorem 2 it follows that there exists a solution $u_0 \in K_F$ of the problem (13).
Let $T : K_F \to 2^{K_F}$ be a set valued mapping defined by

$$Tw = \{u \in K_F \mid \Psi^*(v, u; w, u) \geq 0 \text{ for all } v \in K_F\}.$$

By Theorem 1 we have that, for each fixed $w \in K_F$,

$$Tw = \{u \in K_F \mid \Psi^*(v, u; w, v) \geq \alpha^*(v, u) \text{ for all } v \in K_F\}.$$

From (i2) follows that $x \longmapsto \Psi^*(x, y; z, w)$ is also weakly lower semicontinuous, hence T has nonempty bounded closed and convex set values.
Since T is upper semicontinuous by (i1), using the Kakutani-Fan-Glicksberg fixed point theorem, we obtain that T has a fixed point $w_0 \in K_F$, i.e.

$$\Psi^*(v, w_0; w_0, w_0) \geq 0 \text{ for all } v \in K_F. \tag{14}$$

Now, we define

$$\mathcal{F} = \{F \subset E^{**} \mid F \text{ is a finite dimensional subspace, } F \cap K \neq \emptyset\}$$

and for each $F \in \mathcal{F}$,

$$W_F = \{u \in K \mid \Psi^*(v, u; u, v) \geq \alpha^*(v, u) \text{ for all } v \in K_F\}.$$

By (14) and Theorem 1, we have that W_F is nonempty and bounded. Now, if we denote by $\overline{W}_F$ the $\sigma(E^{**}, E^*)$-closure of W_F in E^*, then $\overline{W}_F$ is $\sigma(E^{**}, E^*)$-compact in E^{**}.
We know that $W_{\bigcap_i F_i} \subset \bigcap_i W_{F_i}$ for $F_i \in \mathcal{F}$, $i \in \{1, \ldots, N\}$. Therefore $\{\overline{W}_F \mid F \in \mathcal{F}\}$ has the finite intersection property and $\bigcap_{F \in \mathcal{F}} \overline{W}_F \neq \emptyset$.
Let $u \in \bigcap_{F \in \mathcal{F}} \overline{W}_F$. For each $v \in K$, we consider $F \in \mathcal{F}$ such that $v \in K_F$ and $u \in K_F$. Then, there exists a sequence $\{u_n\}_n \subset W_F$ which converges to u in $\sigma(E^{**}, E^*)$. From the definition of W_F, we have

$$\Psi^*(v, u_n; u_n, v) \geq \alpha^*(v, u_n).$$

It follows that

$$\Psi^*(v, u; u, v) \geq \alpha^*(v, u) \quad \text{for all } v \in K.$$

Using (i1) through condition (b) of the definition, (i4), and the technique used in Theorem 1 we obtain

$$\Psi^*(v, u; u, u) \geq 0 \text{ for all } v \in K,$$

i.e., u is a solution of problem (12) and the theorem is proved.

Remark 7. We notice that Theorem 3.1 of Fang, Huang [5] and Theorems 2.1 - 2.3 of Chen [4] can now be obtained as corollaries from Theorem 4.

Theorem 5. *Let E be a real Banach space and let $K \subset E^{**}$ be a nonempty unbounded closed convex set. We assume that conditions (i1) - (i5) of Theorem 4 are fulfilled together with* *(i6) there exists a point $y_0 \in K$ such that*

$$\lim_{\|x\| \to \infty} \Psi^* (x, y_0; x, x) > 0.$$

Then, the problem (12) is solvable.

Proof. From Theorem 4 we know that the problem

$$\Psi^* (v, u; u, u) \geq 0 \;\text{ for all } v \in K \cap B_r \tag{15}$$

has a solution $u_r \in K \cap B_r$, where $B_r \subset E^{**}$ is the closed ball centered in 0 with radius r. We choose r large enough such that $y_0 \in B_r$. By (15) we obtain that for $v = y_0$ and $u = u_r$,

$$\Psi^* (y_0, u_r; u_r, u_r) \geq 0.$$

It follows from (i1) and (i6) that $\{u_r\}$ is bounded. Now, using the technique of Theorem 1 we get

$$\Psi^* (v, u_r; u_r, v) \geq \alpha^* (v, u_r) \;\text{ for all } v \in K,$$

and, for $u_r \to u$ in $\sigma (E^{**}, E^*)$ when $r \to \infty$,

$$\Psi^* (v, u; u, v) \geq \alpha^* (v, u) \;\text{ for all } v \in K.$$

We apply again the technique used in Theorem 1 and obtain that u is a solution of problem (12).

Remark 8. The above results remain also true in the case when we consider $\Psi^* : K \times K \times K \times K \to \mathbb{R} \cup \{+\infty\}$ and we suppose that the mapping $x \longmapsto \Psi^* (x, y; z, w)$ is properly convex on K, for any fixed $y, z, w \in K$.

Remark 9. We observe that Theorem 3.2 of Fang, Huang [5] and Theorems 2.4 - 2.6 of Chen [4] become corollaries of Theorem 5.

5 Some Examples

In this section we consider some examples that justify the extensions presented in sections 4 and 5. We recall that in the examples of section 2 different forms of the function α were emphasized. In the next examples new forms of the functions α and α^* are given.
In the first three examples we consider the case of the reflexive Banach space $E = \mathbb{R}$ for which we show that, according to section 3, the following can happen:

- the mapping $y \longmapsto \langle Tz, \eta (y, x) \rangle$ is convex for any $y, z \in K$ whereas $y \longmapsto \varphi (y, x) = f (y) - f (x)$ is not convex with respect to y, but the mapping $y \longmapsto \Psi (y, x; z) = \langle Tz, \eta (y, x) \rangle + \varphi (y, x)$ is convex;

- the mapping $y \longmapsto \langle Tz, \eta(y,x)\rangle$ is not convex for any $y, z \in K$ whereas $y \longmapsto \varphi(y,x) = f(y) - f(x)$ is convex with respect to y, but the mapping $y \longmapsto \Psi(y,x;z) = \langle Tz, \eta(y,x)\rangle + \varphi(y,x)$ is convex;
- the mapping $y \longmapsto \Psi(y,x;z)$ is convex but Ψ has not the above form.

In all these instances the mapping $y \longmapsto \Psi(y,x;z)$ verifies the hypotheses of theorems 1 - 3. We notice that, for example, the results of the corresponding theorems of Fang, Huang [5] do not apply for the cases of our examples.
In the last two examples we consider other Banach spaces than $\mathbb{R}$.

Example 4. Let us consider $E = E^* = \mathbb{R}$, the closed bounded convex set $K = [-1,1]$, and the mappings

$$\begin{aligned} &T : K \to E, \ \ Tz = -|z|, \\ &\eta : K \times K \to E, \ \ \eta(y,x) = y^2 - x^2, \\ &\varphi : K \times K \to E, \ \ \varphi(y,x) = y^4 + y^2 - x^4 - x^2. \end{aligned}$$

We have: $\langle Tz, \eta(y,x)\rangle = -|z|\left(y^2 - x^2\right)$, which is not convex with respect to y, φ is convex, but $\Psi(y,x;z) = -|z|\left(y^2 - x^2\right) + y^4 + y^2 - x^4 - x^2$ is convex with respect to y on K for any $x, z \in K$.
If we take $\alpha(x,y) = \left(y^2 - x^2\right)(|x| - |y|)$ then Ψ verifies the assumptions $(i1)-(i4)$ of Theorem 1 and $(j1)-(j5)$ of Theorem 2.

Example 5. We consider again $E = E^* = \mathbb{R}$, the closed unbounded convex set $K = [1,\infty)$, and the mappings

$$Tz = z, \ \ \eta(y,x) = y^2 - x^2, \ \ \varphi(y,x) = x^2 - y^2.$$

We have: $\langle Tz, \eta(y,x)\rangle = z\left(y^2 - x^2\right)$, which is convex on K with respect to y, φ is not convex on K with respect to y, but $\Psi(y,x;z) = \langle Tz, \eta(y,x)\rangle + \varphi(y,x)$ is convex on K with respect to y.
For $\alpha(x,y) = \left(y^2 - x^2\right)(|y| - |x|)$ the hypotheses of Theorem 3 are fulfilled, which also suppose a coercivity assumption.

Example 6. Let $E = E^* = \mathbb{R}$, $K = [1,4]$ or $K = [1,\infty)$. We consider $\Psi(y,x;z) = \left(y^2 - x^2\right)\exp(z) + z(y - x)$, which has not the form

$$\langle Tz, \eta(y,x)\rangle + \varphi(y,x).$$

If we define $\alpha(x,y) = \left(y^2 - x^2\right)(\exp(y) - \exp(x)) + (y - x)^2$, then the mapping Ψ verifies the hypotheses of Theorems 1 and 2 for $K = [1,4]$, and the coercivity hypothesis of Theorem 3 for $K = [1,\infty)$.

Example 7. Let us consider

$$E = \ell_2 = \left\{ x = (x_n)_{n\geq 1} \;\middle|\; \|x\| = \left(\sum_{n\geq 1} x_n^2\right)^{1/2} < \infty, \ x_n \in \mathbb{R}, \ n \geq 1 \right\},$$

which is a reflexive Banach space [21]. We consider

$$K = \left\{ x \in \ell_2 \; \middle| \; |x_n| \leq 3^{-n}, \; \forall \, n \geq 1 \right\},$$

which is a closed, convex and bounded set in ℓ_2. Since $u : K \to \mathbb{R}$, $u(x) = \sum_{n \geq 1} 2^n x_n$ is a continuous and affine mapping on K [21], the mapping $\Psi : K \times K \times K \to \mathbb{R}$ defined by

$$\Psi(y, x; z) = \|z\| \sum_{n \geq 1} 2^n (y_n - x_n) + \sum_{n \geq 1} (y_n - x_n) \exp(\|z\|)$$

is hemicontinuous and convex as required in the Theorems 1 and 2. It is also easy to show that the other assumptions of the mentioned theorems are fulfilled for

$$\begin{aligned} \alpha(x, y) = (\|y\| - \|x\|) \sum_{n \geq 1} 2^n (y_n - x_n) + \\ + (\exp(\|y\|) - \exp(\|x\|)) \sum_{n \geq 1} (y_n - x_n). \end{aligned}$$

Example 8. Let E be the space of square-integrable random variables defined on some fixed probability space. Hence, if $X \in E$ then $\mathcal{E}X^2 < \infty$, where $\mathcal{E}X^2$ is the usual expectation of the random variable X^2. Almost surely equal random variables are regarded as identical. Considering the scalar product

$$\langle X, Y \rangle = \mathcal{E}(XY), \quad \text{for } X, Y \in E,$$

the space E becomes a Hilbert space (see, for example, Neveu [20]). The induced norm will be $\|X\| = \left(\mathcal{E}X^2\right)^{1/2}$. According to [21] E is a reflexive Banach space, i.e., $E = E^*$. We take $K = \{X \in E \mid 1 \leq \mathcal{E}X \leq 2\}$ which is a nonempty convex, closed and nonempty set in E. For $X, Y, Z \in E$ let us define

$$\begin{aligned} \Psi(Y, X; Z) = \left[\mathcal{E}\left(Y^2 - Y\right) - \mathcal{E}\left(X^2 - X\right)\right] \sin^2 \|Z\| + \\ + (\mathcal{E}Y - \mathcal{E}X) \ln(\|Z\| + 1) \end{aligned}$$

and

$$\begin{aligned} \alpha(X, Y) = \left[\mathcal{E}\left(Y^2 - Y\right) - \mathcal{E}\left(X^2 - X\right)\right] \left(\sin^2 \|Y\| - \sin^2 \|X\|\right) + \\ + (\mathcal{E}Y - \mathcal{E}X) \ln \frac{\|Y\| + 1}{\|X\| + 1}. \end{aligned}$$

Since for $Y_1, Y_2 \in E$ and $0 \leq \lambda \leq 1$ we have $\mathcal{E}Y_1^2 + \mathcal{E}Y_2^2 \geq 2\mathcal{E}(Y_1 Y_2)$ and

$$\begin{aligned} \mathcal{E}(\lambda Y_1 + (1 - \lambda) Y_2)^2 = \mathcal{E}\left(\lambda^2 Y_1^2 + (1 - \lambda)^2 Y_2^2 + 2\lambda(1 - \lambda) Y_1 Y_2\right) \leq \\ \leq \lambda^2 \mathcal{E}Y_1^2 + (1 - \lambda)^2 \mathcal{E}Y_2^2 + \lambda(1 - \lambda)\left(\mathcal{E}Y_1^2 + \mathcal{E}Y_2^2\right) = \\ = \lambda \mathcal{E}Y_1^2 + (1 - \lambda) \mathcal{E}Y_2^2 \end{aligned}$$

We see that $\Psi(\,\cdot\,, X, Z)$ is a convex mapping on K for any fixed $X, Z \in K$. Also, it is easy to show that the other assumptions of the Theorems 1 and 2 are fulfilled.
We notice that if $K = \{X \in E \mid \mathcal{E}X \geq 1\}$ the above defined mappings Ψ and α satisfy the conditions of Theorem 3.

Example 9. Let $\mathbb{C}$ be the set of complex numbers and

$$E = c_0 = \left\{ x = (x_n)_{n \geq 1} \;\middle|\; x_n \in \mathbb{C},\ \lim_{n \to \infty} x_n = 0 \right\}.$$

We have

$$E^* = c_0^* = \ell_1 = \left\{ x = (x_n)_{n \geq 1} \;\middle|\; x_n \in \mathbb{C},\ |x|_1 \equiv \sum_{n \geq 1} |x_n| < \infty \right\}$$

and

$$E^{**} = \ell_1^* = c_0^{**} = \ell_\infty = \left\{ x = (x_n)_{n \geq 1} \;\middle|\; x_n \in \mathbb{C},\ \|x\|_\infty \equiv \sup_{n \geq 1} |x_n| < \infty \right\}$$

(see [21]). We note that $\ell_\infty^* \neq \ell_1$.
Let $a > 0$ be a real number and we define the set

$$K = \{x \in E \mid x_n \in \mathbb{R},\ 0 \leq x_n \leq a,\ \forall\, n \geq 1\},$$

which is a convex, closed and bounded set in E^{**}. For $x, y, z, w \in E^{**}$ we define

$$\Psi(x, y; z, w) = \left(\sum_{n \geq 1} \frac{x_n}{2^n} - \sum_{n \geq 1} \frac{y_n}{2^n} \right) (\|z\| + \|w\| + \exp \|w\|)$$

and

$$\alpha(x, y) = \left(\sum_{n \geq 1} \frac{x_n}{2^n} - \sum_{n \geq 1} \frac{y_n}{2^n} \right) (\|x\| - \|y\| + \exp \|x\| - \exp \|y\|)$$

We remark that Ψ and α satisfy the assumptions of Theorem 4. The assumptions of Theorem 5 will also be satisfied if we define the set

$$K = \{x \in E \mid x_n \in \mathbb{R},\ x_n \geq 0,\ \forall\, n \geq 1\}.$$

As final remarks we notice that, by using the spaces $E = c_0$ and $E^{**} = \ell_\infty$, we can construct similar examples like the Examples 4 and 5. Also, one can use the Banach spaces ℓ_p and L^p, with $1 < p < \infty$, for further examples.

Acknowledgment. The authors are thankfully to the referees for their valuable remarks, especially for the suggestion to include some examples, which confirm the significance of the proposed extensions.

References

1. Ansari QH, Siddiqi AH, Yao JC (2000) Generalized vector variational-like inequalities and their scalarizations. In: Giannessi F (ed) Vector Variational Inequalities and Vector Equilibria. Kluwer Academic Publishers, Dordrecht, Holland
2. Ansari QH, Konnov IV, Yao JC (2002) Characterization of solutions for vector equilibrium problems. Journal of Optimization Theory and Applications 113: 435–447
3. Chang SS, Lee BS, Chen YQ (1995) Variational inequalities for monotone operators in nonreflexive Banach spaces. Applied Mathematics Letters 8: 29–34
4. Chen YQ (1999) On the semimonotone operator theory and applications. Journal of Mathematical Analysis and Applications 231: 177–192
5. Fang YP, Huang NJ (2003) Variational-like inequalities with generalized monotone mappings in Banach spaces. Journal of Optimization Theory and Applications 118: 327–337
6. Giannessi F (ed) (2000) Vector Variational Inequalities and Vector Equilibria, Kluwer Academic Publishers, Dordrecht, Holland
7. Giannessi F, Maugeri A (1995) Variational Inequalities and Network Equilibrium Problems. Plenum Publishing Corporation, New York
8. Glowinski R, Lions JL, Tremolieres R (1981) Numerical Analysis of Variational Inequalities. North-Holland, Amsterdam, Holland
9. Goeleven D, Motreanu D (1996) Eigenvalue and dynamic problems for variational and hemivariational inequalities. Communications on Applied Nonlinear Analysis, 3: 1–21
10. Hadjisavvas N, Schaible S (1996) Quasimonotone variational inequalities in Banach spaces. Journal of Optimization Theory and Applications 90: 95–111
11. Hartman GJ, Stampacchia G (1966) On some nonlinear elliptic differential functional equations. Acta Mathematica 115: 271–310
12. Kang MK, Huang NJ, Lee BS (2003) Generalized pseudo-monotone set-valued variational-like inequalities. Indian Journal of Mathematics 45: 251–264
13. Kassay G, Kolumban I (2000) Variational inequalities given by semi - pseudomonotone mappings. Nonlinear Analysis Forum, 5: 35–50
14. Konnov IV, Yao JC (1997) On the generalized vector variational inequality problems. Journal of Mathematical Analysis and Applications 206: 42–58
15. Konnov IV (2001) Combined relaxation methods for variational inequalities. Lecture notes in economics and mathematical systems, Vol. 495, Springer, Berlin
16. Konnov IV (1998) A combined relaxation method for variational inequalities with nonlinear constraints. Mathematical Programming 80: 239–252
17. Konnov IV (2002) A class of combined relaxation methods for decomposable variational inequalities. Optimization 51: 109–125
18. Ky Fan (1961) A generalization of Tychonoff's fixed-point theorem, Mathematische Annalen, 142: 305–310
19. Luc DT (1989) Theory of Vector Optimization. Lecture Notes in Economics and Mathematical Systems, Springer Verlag, New York
20. Neveu J (1964) Bases Mathematiques du Calcul des Probabilites. Masson, Paris
21. Reed M, Simon B (1972) Methods of Modern Mathematical Physics, I. Functional Analysis. Academic Press, New York and London
22. A.H. Siddiqi, Q.H. Ansari, K.R. Kazmi (1994) On nonlinear variational inequalities. Indian Journal on Pure and Applied Mathematics 25: 969–973

23. Schaible S (1995) Generalized monotonicity - concepts and uses. In: Giannessi F, Maugeri A (eds) Variational Inequalities and Network Equilibrium Problems, Plenum Publishing Corporation, New York
24. Schaible S (1998) Criteria for generalized monotonicity. In: Giannessi F et al. (eds) New Trends in Mathematical Programming, Kluwer Academic Publishers, Dordrecht
25. Verma RU (1997) Nonlinear variational inequalities on convex subsets of Banach spaces. Applied Mathematics Letters 10: 25–27
26. Verma RU (1998) On monotone nonlinear variational inequalitiy problems. Comentationes Mathematicae Universitatis Carolinae, 39: 91–98

Almost Pure Nash Equilibria in Convex Noncooperative Games*

Tadeusz Radzik[1] and Wojciech Połowczuk[2]

[1] Institute of Mathematics, Wrocław University of Technology, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland, t_radzik@dami.pl
[2] Institute of Mathematics, Wrocław University of Technology, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland, wojciech.polowczuk@pwr.wroc.pl

Summary. This paper considers n-person non-coalitional games with finite players' strategy spaces and payoff functions having some concavity or convexity properties. For such games it is shown that there are two-point Nash equilibria in them, that is equilibria in players' strategies with support consisting of at most two points. The structure of such simple equilibria is discussed in different cases. The results obtained in the paper can be seen as a discrete counterpart of Glicksberg's theorem and other known results about the existence of pure (or "almost pure") Nash equilibria in continuous concave (convex) games with compact convex spaces of players' pure strategies.

Key words: Noncooperative games, matrix games, Nash equilibrium, convex payoffs, two-point strategies.

1 Introduction

The assumption of concavity/convexity of payoff functions is very often used, both in theoretical considerations and practical applications of noncooperative games. This kind of properties allow to look for players' strategies with a very simple structure (unrandomized) and creating equilibria of games. One of the most important concepts of optimal solution for noncooperative games is a *Nash equilibrium*. In particular, a Nash equilibrium realized in pure (unrandomized) strategies is very convenient for the players. Classical result in this field (Glicksberg[4]) says about the existence of a Nash equilibrium in pure strategies in n-person non-zero-sum games with continuous quasi-concave payoffs. There are other similar results on the existence of pure Nash equilibria in games ([12], [1], [8], [6]). However the basic parts in assumptions of those

*This paper was supported by Grant No. 342209, Wrocław University of Technology.

models are different types of convexity of players' payoff functions. The second feature of games considered in those models is that the spaces of players' strategies are uncountable. In many situations, however, goods have a discrete structure and are nondivisible, (e.g. people, cars), and therefore we cannot always consider continuous strategy spaces for the players. In such cases finite games arise and thereby the payoff functions are defined on finite sets. In this connection two fundamental questions arise:
(1) How to define convexity of payoff functions in finite games?
(2) Will such convex/concave finite game still possess pure or approximately pure Nash equilibria?
Just this kind of problems are discussed in this paper.
The organization of the paper is as follows. In Section 2 we present background results which are an inspiration for our further considerations. In Section 3 we give a rich review of recent results for convex finite two-person games. In Section 4 we discuss the case of n-person non-zero-sum finite games. There we give a theorem which is a discrete counterpart of Glicksberg theorem for infinite convex games.

2 Background Results

In this section we recall four background theorems, essential for our further considerations. First we need to fix some notation. We will start with the definition of an n-person non-zero-sum game G_n in the following normal form,

$$G_n = \langle N, \{X_i\}_{i\in N}, \{F_i\}_{i\in N}\rangle, \tag{1}$$

where

1. $N = \{1, 2, \ldots, n\}$ is a finite set of players;
2. for each $i \in N$, X_i is a space of pure strategies x_i of i-th Player;
3. for each $i \in N$ and $x = (x_1, x_2, \ldots, x_n) \in \Pi_{i\in N} X_i$, $F_i(x)$ is the payoff function of Player i, in the situation when players use pure strategies $x_1, x_2, \ldots, x_n$, respectively.

One of the most important concepts of optimal solution for such games is a *pure Nash equilibrium* or equivalently, a *Nash equilibrium in pure strategies*. It is defined as any *strategy profile* x^* consisting of players' pure strategies of the form $x^* = (x_1^*, x_2^*, \ldots, x_n^*) \in \prod_{i\in N} X_i$ satisfying the inequalities

$$F_i(x^*) \geq F_i(x_1^*, \ldots, x_{i-1}^*, x_i, x_{i+1}^*, \ldots, x_n^*) \quad \text{for } i \in N \text{ and } x_i \in X_i. \tag{2}$$

When all these inequalities hold up to an $\epsilon > 0$, we say about *pure ϵ-Nash equilibrium*. Such solutions in pure strategies have possibly the simplest structure and thereby, they are very desirable in practical applications.
A *mixed strategy* of player $i \in N$ in the game G_n is any probability distribution μ_i over the space X_i. It happens very often that there is no Nash equilibrium

in pure strategies. In such situations one can look for a Nash equilibrium in mixed strategies. Then for each $i \in N$, Player i's space X_i of his pure strategies is extended to the space $\mathcal{X}_i$ of his mixed strategies, and domain $\prod_{i \in N} X_i$ of his payoff function F_i is extended (as expected values with respect to the product distribution) to richer domain $\prod_{i \in N} \mathcal{X}_i$. Then, when each player $i \in N$ acts according to his mixed strategy μ_i, the *mixed strategy profile* is $\mu = (\mu_1, \mu_2, \ldots, \mu_n)$ and the payoff functions F_i are treated as defined on the space of the players' mixed strategy profiles and for each $i \in N$ equal to

$$F_i(\mu_1, \mu_2, \ldots, \mu_n) = \int_{\prod_{i \in N} \mathcal{X}_i} F_i(x_1, x_2, \ldots, x_n) d\mu_1(x_1) d\mu_2(x_2) \ldots d\mu_n(x_n) .$$

Now a *mixed Nash equilibrium* (or a *Nash equilibrium in mixed strategies*) is any mixed strategy profile $\mu^* = (\mu_1^*, \mu_2^*, \ldots, \mu_n^*) \in \prod_{i \in N} \mathcal{X}_i$ satisfying analogous inequalities to (2). It is known that any n-person game with all spaces X_i finite always has a mixed Nash equilibrium (Nash[5]).
To begin with, we will quote the most important result for infinite games. A first background theorem is basic for *n-person non-zero-sum games.* We recall here that by definition, a real-valued function $f(x)$ on a convex set X is *quasi-concave*, when for each real c, the set $\{x : f(x) \geq c\}$ is convex. Of course, every concave function is quasi-concave.

Theorem 1. (Debreu[2], Glicksberg[4], Fan[3]) *Let X_i, $i \in N$, be convex and compact subsets of some euclidean spaces $\mathcal{R}^{m_i}$, $m_i \geq 1$. If every function $F_i(x_1, \ldots, x_n)$ is continuous on $\prod_{i \in N} X_i$ and quasi-concave in x_i, then the n-person non-zero-sum game $G_n = \langle N, \{X_i\}_{i \in N}, \{F_i\}_{i \in N} \rangle$ possesses a pure strategy Nash equilibrium.*

The next two results we recall (Theorems 2–3) concern *two-person zero-sum games.* Such games are defined by the normal form

$$G_2^z = \langle \{1, 2\}, \{X, Y\}, \{F_1, -F_1\} \rangle, \tag{3}$$

where X and Y are strategy spaces of Players 1 and 2, and $F_1(x, y)$ and $-F_1(x, y)$ are payoff functions of Players 1 and 2, respectively. It appears that for two-person zero-sum games we can say much more (in comparison to Theorem 1) about the situations when Nash equilibria exist. We assume for the next two theorems that $X \subset \mathcal{R}^m$ and $Y \subset \mathcal{R}^n$, $m, n \geq 1$.

Theorem 2. (Sion[12]) *Let X, Y be convex sets with X compact. Assume that $F_1(x, y)$ is an upper semicontinuous function in x and quasi-concave in x for each y, and quasi-convex in y for each x. Then for any $\epsilon > 0$ the two-person zero-sum game G_2^z possesses a pure ϵ-Nash equilibrium.*

Theorem 3. *(*Bohnenblust, Karlin and Shapley[1]) *Let X, Y be compact sets with X convex. Assume that $F_1(x, y)$ is a continuous function on $X \times Y$ and concave in x for each y. Then the two-person zero-sum game G_2^z possesses* *a Nash equilibrium* (μ^*, ν^*) *with a pure strategy* μ^* *for Player 1, and with a mixed strategy* ν^* *for Player 2, being a probability measure concentrated in at most* n *points of* Y.

In the last two theorems, a *two-person non-zero-sum game* with the strategy spaces $X = Y = [0, 1]$ is considered. Let us denote such a game by

$$G_2^{nz} = \langle\{1, 2\}, \{[0, 1], [0, 1]\}, \{F_1, F_2\}\rangle, \tag{4}$$

where $F_1(x, y)$ and $F_2(x, y)$ are payoff functions of Players 1 and 2, respectively. It appears that such a special form of the players' strategy spaces as the unit intervals, ensures the existence of a pure ϵ-Nash equilibrium under much weaker assumptions (on the payoff functions) in comparison to the ones of the previous theorems. Practically, the only restriction on the payoff functions is the convexity (concavity) of Player 1's payoff function $F_1(x, y)$ in variable x for each $y \in [0, 1]$, while the second payoff function can be quite arbitrary (without any "continuity" assumptions). This is described below in Theorems 4 and 5. The payoff functions $F_1(x, y)$ and $F_2(x, y)$ are assumed there to be bounded and bounded from above on $[0, 1] \times [0, 1]$, respectively. Throughout the paper we shall often use the symbol δ_t as

δ_t – *a degenerate probability distribution concentrated at point* t.

Theorem 4. (Radzik[9]) *Let* $F_1(x, y)$ *be concave in* x *for each* y. *Then for any* $\varepsilon > 0$, *the two-person non-zero-sum game* G_2^{nz} *has an* ε*-Nash equilibrium of the form* $(\mu_1^*, \mu_2^*) = (\alpha\delta_a + (1 - \alpha)\delta_b, \beta\delta_c + (1 - \beta)\delta_d)$, *for some* $0 \leq \alpha, \beta, a, b, c, d \leq 1$ *with* $|a - b| < \varepsilon$.

Theorem 5. (Radzik[9]) *Let* $F_1(x, y)$ *be convex in* x *for each* y. *Then for any* $\varepsilon > 0$, *the two-person non-zero-sum game* G_2^{nz} *has an* ε*-Nash equilibrium of the form* $(\mu_1^*, \mu_2^*) = (\alpha\delta_0 + (1-\alpha)\delta_1, \beta\delta_c + (1-\beta)\delta_d)$, *for some* $0 \leq \alpha, \beta, c, d \leq 1$, *where* α *is independent of* ε.

In all the theorems given above, the convexity and/or concavity of the players' payoff functions play a remarkable role in their assumptions, and, besides, the players' strategy spaces are infinite (uncountable). However, in many situations, the players' strategy spaces are finite and thereby, those theorems say nothing about possible existence of Nash equilibria and the above theorems cannot be applied. Hence, a very essential question is whether there are possible "discrete" counterparts of Theorems 1–5, that is, analogs of those theorems with the players' finite strategy spaces. How to define the convexity/concavity of payoff functions on finite sets? Do such discrete counterparts preserve the propositions about the existence of Nash equilibria of similar "simple" form? In the next sections we just study these problems both for zero-sum and non-zero-sum games.

3 Two-person Finite Games

In this section we discuss two-person *finite games*, both in zero-sum and non-zero-sum version. According to (1), any two-person non-zero-sum game can be described by the triplet $G_2 = \langle\{1,2\},\{X,Y\},\{F_1,F_2\}\rangle$, where X and Y are pure strategy spaces for Players 1 and 2, respectively, and $F_1(x,y)$ and $F_2(x,y)$ are their payoff functions. A game Γ is called *finite* when the spaces X and Y are finite sets.
This section consists of two parts. The first subsection contains all the needed definitions and notions. The second subsection is devoted to a review of our most interesting results (see papers of Radzik and Połowczuk: [8], [10], [6] and [7]) for two-person zero-sum and non-zero-sum finite games. In the literature such games are generally called matrix and bimatrix games, respectively.

3.1 Definitions and Preliminary Results

For the rest of this section we will consider two-person non-zero-sum finite games $\Gamma = \langle\{1,2\},\{X,Y\},\{F_1,F_2\}\rangle$ with strategy spaces of the form

$$X = \{1,2,\ldots,m\} \quad \text{and} \quad Y = \{1,2,\ldots,n\}$$

for two naturals m and n, and with payoff functions F_1 and F_2 for Players 1 and 2, respectively.
Let $A = [a_{ij}]$ and $B = [b_{ij}]$ denote two matrices of size $(m \times n)$ such that $a_{ij} = F_1(i,j)$ and $b_{ij} = F_2(i,j)$ for all i and j. We will denote such a *bimatrix game* by $\Gamma(A,B)$ with A and B as the players' payoff matrices. We will also use the notation $(m \times n)$–game $\Gamma(A,B)$ to emphasize the dimension of the payoff matrices A and B in the game.
Therefore, the game $\Gamma(A,B)$ is played in such a way that Player 1 chooses any row $i, 1 \le i \le m$ and simultaneously, Player 2 chooses any column $j, 1 \le j \le m$, and then the players payoffs are equal to a_{ij} and b_{ij}, respectively. Of course, any pair (i,j) of players' pure strategies in game $\Gamma(A,B)$ can be identified with the pair (δ_i,δ_j).
One can easily see that the pair (i,j) is a pure Nash equilibrium in bimatrix game $\Gamma(A,B)$ if a_{ij} and b_{ij} are the biggest elements in the j-th column of matrix A and in the i-th row of matrix B, respectively.
When $\Gamma(A,B)$ is a zero-sum game, that is, when $B = -A$, then the game $\Gamma(A,-A)$ is called a *matrix game* and shortly denoted by $\Gamma(A)$.
It is clear that the pair (i,j) is a pure Nash equilibrium in matrix game $\Gamma(A)$ if a_{ij} is the smallest element in the i-th row and the biggest in the j-th column of matrix A. Then such a pair (i,j) is also called a *saddle point* of matrix A.
Now we will define two special types of players' mixed strategies with a very simple structure. Such strategies play an important role in our paper.

Definition 1. *A (mixed) strategy μ_1 of Player 1 in game $\Gamma(A,B)$ is a two-point strategy if it is of the form $\mu_1 = \alpha\delta_a + (1-\alpha)\delta_b$ with some $0 \le \alpha \le 1$*

and $1 \le a, b \le m$.
If $b = a+1$, *then* μ_1 *is a two-adjoining-point strategy. Strategies for Player 2 are defined analogously.*

Now we give the definitions of several types of concavity/convexity for finite games, which are basic for our paper.

Definition 2. *A bimatrix game* $\Gamma(A,B)$ *is concave if there exist two functions* $F_1(x,y)$ *and* $F_2(x,y)$ *on the unit square, concave in* x *for each* y *and concave in* y *for each* x*, respectively, and if there are two strictly increasing sequences* $\{x_i\}_{i=1}^m$ *and* $\{y_j\}_{j=1}^n$ *in* $[0,1]$*, such that* $F_1(x_i,y_j) = a_{ij}$ *and* $F_2(x_i,y_j) = b_{ij}$ *for all* i *and* j.
If game $\Gamma(A,B)$ *satisfies this definition only with respect to the function* $F_1(x,y)$*, then it is a column-concave game.*
The properties of quasi-concavity, and convexity of game $\Gamma(A,B)$ *are defined analogously.*

Remark 1. One can easily see that the two-person non-zero-sum game G_2^{nz} of the form (4), under the assumptions of Theorem 4, can be "discretized" to an $(m \times n)$–bimatrix game $\Gamma(A,B)$ with column-concavity property. On the other hand, the same game G_2^{nz} under the assumptions of Theorem 5 leads, after discretization, to a bimatrix game with column-convexity property. Similarly, the two-person case of a game from Theorem 1 leads to a quasi-concave bimatrix game. On the other hand, after discretization of games from Theorems 2 and 3 we get matrix games which are quasi-concave and column-concave, respectively.

Remark 2. All Theorems 1–5 say about the existence (in some infinite games) of a pure Nash equilibria or two-point Nash equilibria. Hence, a very natural supposition is that the corresponding bimatrix or matrix games, should also have Nash equilibria with a similar simple structure, that is, in pure or two-point strategies. This will be discussed in our further considerations.

For a given game $\Gamma(A,B)$ it is rather difficult to check directly whether it is concave or not. It appears, however, that there exists an alternative (equivalent) characterization of concavity for bimatrix games, which allows us to check without difficulty, if a game has this property. The proof of this result is identical with the one for two-person zero-sum games, given in [10]. It can be written in the following form.

Proposition 1. *A game* $\Gamma(A,B)$ *is concave, if and only if there exist positive numbers* $\theta_1, \theta_2, \ldots, \theta_{m-1}$ *and* $\tau_1, \tau_2, \ldots \tau_{n-1}$ *such that*

$$\theta_1(a_{2j} - a_{1j}) \ge \theta_2(a_{3j} - a_{2j}) \ge \cdots \ge \theta_{m-1}(a_{mj} - a_{m-1,j}) \text{ for all } j \tag{5}$$

and

$$\tau_1(b_{i2} - b_{i1}) \ge \tau_2(b_{i3} - b_{i2}) \ge \cdots \ge \tau_{n-1}(b_{in} - b_{i,n-1}) \text{ for all } i. \tag{6}$$

When all the inequalities in (5) and in (6) are reverse, game $\Gamma(A,B)$ is a convex game. When only inequalities (5) hold, game $\Gamma(A,B)$ is column-concave (or column-convex when inequalities in (5) are reverse).

Remark 3. Note that (5) and (6) hold with positive $\theta_1, \ldots, \theta_{m-1}$ and $\tau_1, \ldots, \tau_{n-1}$ if and only if for each k and l, $1 \le k \le m-2$, $1 \le l \le n-2$ there are $\alpha_k > 0$ and $\beta_l > 0$ such that $\alpha_k(a_{k+1,j} - a_{kj}) \ge a_{k+2,j} - a_{k+1,j}$ for all j and $\beta_l(b_{i,l+1} - b_{il}) \ge b_{i,l+2} - b_{i,l+1}$ for all i. These two conditions are easily verifiable, allowing to check whether a game is concave. An analogous algorithm can be used in the "convex" case.

We now formulate a theorem allowing to check directly whether a bimatrix game is quasi-concave. It follows from the proof of Theorem 3.2 in [10].

Proposition 2. *A bimatrix game $\Gamma(A,B)$ is quasi-concave if and only if for each i and j, $1 \le i \le m$, $1 \le j \le n$, there exist natural k and l, $1 \le k \le n$, $1 \le l \le m$, such that*

$$\begin{cases} a_{1j} \le a_{2j} \le \cdots \le a_{lj} \ge a_{l+1,j} \ge \cdots \ge a_{mj} \\ b_{i1} \le b_{i2} \le \cdots \le b_{ik} \ge b_{i,k+1} \ge \cdots \ge b_{in} \, . \end{cases}$$

When $\Gamma(A,B)$ is a zero-sum game satisfying these inequalities, the functions F_1 and F_2 (in Def. 2) can be chosen with $F_2 = -F_1$.

The structure of inequalities in the above Proposition 2 encourages to define two other types of quasi-concavity of a bimatrix game. They will be basic for our further results.

Definition 3. *A bimatrix game $\Gamma(A,B)$ is strongly quasi-concave if for each i and j, $1 \le i \le m$, $1 \le j \le n$, there exist k, l, r and s, $1 \le k \le l \le n$, $1 \le r \le s \le m$, such that*

$$\begin{cases} a_{1j} < a_{2j} < \cdots < a_{rj} = a_{r+1,j} = \cdots = a_{sj} > a_{s+1,j} > \cdots > a_{mj} \\ b_{i1} < b_{i2} < \cdots < b_{ik} = b_{i,k+1} = \cdots = b_{il} > b_{i,l+1} > \cdots > b_{in} \, . \end{cases}$$

Definition 4. *A bimatrix game $\Gamma(A,B)$ is strictly quasi-concave if for each i and j, $1 \le i \le m$, $1 \le j \le n$, there exist k and r, $1 \le k \le n$, $1 \le r \le m$, such that*

$$\begin{cases} a_{1j} < a_{2j} < \cdots < a_{rj} > a_{r+1,j} > \cdots > a_{mj} \\ b_{i1} < b_{i2} < \cdots < b_{ik} > b_{i,k+1} > \cdots > b_{in} \, . \end{cases}$$

At the end we give a theorem about concavity of a bimatrix game, analogous to Proposition 2.

Proposition 3. *Any concave bimatrix game $\Gamma(A,B)$ is strongly quasi - concave.*

Proof. Fix j, $1 \leq j \leq n$, and let us put $c_i = \theta_i(a_{i+1,j} - a_{ij})$, $i = 1, 2, \ldots, m-1$. Therefore, by (5), the sequence $(c_1, c_2, \ldots, c_{m-1})$ is nonincreasing. Further, let $r = max\{i : \ c_i > 0\}$ (we put $r = 1$ if $\{i : \ c_i > 0\} = \emptyset$), and let $s = min\{i : \ c_i < 0\}$ (we put $s = m - 1$ if $\{i : \ c_i < 0\} = \emptyset$). Now, it is easily seen that for such r and s, the first line of inequalities in Definition 3 hold. Inequalities in the second line of Definition 3 can be shown analogously.

3.2 Equilibria in Two-person Finite Games

The four possible properties of bimatrix games $\Gamma(A, B)$ introduced in the previous subsection, that is, concavity, quasi-concavity, strong quasi-concavity and strict quasi-concavity, are basic for our subsequent considerations. They play an essential role in assumptions of the presented theorems about the existence of pure and two-point Nash equilibria in both zero-sum and non-zero-sum games.
We begin with the first group of three theorems about sufficient conditions for the existence of pure Nash equilibria in a bimatrix game $\Gamma(A, B)$ (their proofs and several counterexamples related to them one can find in [6]).
To express the theorems we need one more notation. Namely for any $(m \times n)$-matrix $W = [w_{rs}]$, let W_{kl}^{ij}, $1 \leq i \leq k \leq m$, $1 \leq j \leq l \leq n$, be its submatrix of the form:

$$W_{kl}^{ij} := \begin{bmatrix} w_{ij} & w_{i,j+1} & \cdots & w_{i,l} \\ w_{i+1,j} & w_{i+1,j+1} & \cdots & w_{i+1,l} \\ \vdots & \vdots & & \vdots \\ w_{kj} & w_{k,j+1} & \cdots & w_{kl} \end{bmatrix}.$$

Let $\Gamma(A, B)$ be a bimatrix game, where $A = [a_{rs}]$ and $B = [b_{rs}]$ are $(m \times n)$-matrices. By the subgame Γ_{kl}^{ij} of game $\Gamma(A, B)$, where $1 \leq i < k \leq m$ and $1 \leq j < l \leq n$, we mean the game $\Gamma(A_{kl}^{ij}, B_{kl}^{ij})$.

Theorem 6. *Assume that $\Gamma(A, B)$ is a strongly quasi-concave game and that every (2×2)-subgame (obtained by removing $m - 2$ rows and $n - 2$ columns of A and B) has a pure Nash equilibrium. Then the game Γ has a pure Nash equilibrium as well.*

Theorem 7. *Assume that $\Gamma(A, B)$ is a strongly quasi-concave game. If all $(2 \times l)$-subgames $\Gamma_{i+1,j+l-1}^{ij}$ and all $(k \times 2)$-subgames $\Gamma_{i+k-1,j+1}^{ij}$ of the game Γ ($1 \leq i < m$, $1 \leq j < n$, $k = 2, \ldots, m - i + 1$, $l = 2, \ldots, n - j + 1$) have pure Nash equilibria, then the game Γ also has a pure Nash equilibrium.*

The third theorem deals with strictly quasi-concave bimatrix games. It says that if such a game can be divided ("vertically" or "horizontally") into two subgames having pure Nash equilibria, then it also has a pure Nash equilibrium.

Theorem 8. *Let $\Gamma(A, B)$ be a strictly quasi-concave game. Assume that one of the following statements holds:*

1. *there exists* k*,* $1 < k < m$ *such that both subgames* $\Gamma_1 = \Gamma_{kn}^{11}$ *and* $\Gamma_2 = \Gamma_{mn}^{k1}$ *have pure Nash equilibria;*
2. *there exists* l*,* $1 < l < n$ *such that both subgames* $\Gamma_1 = \Gamma_{ml}^{11}$ *and* $\Gamma_2 = \Gamma_{mn}^{1l}$ *have pure Nash equilibria.*

Then the game Γ *has also a pure Nash equilibrium.*

Theorem 8 has an interesting corollary.

Corollary 1. *Assume that the game* $\Gamma(A, B)$ *is strictly quasi-concave. If all* (2×2)*-subgames of the form* $\Gamma_{i+1,j+1}^{ij}$ *of the game* Γ *have pure Nash equilibria, then also the game* Γ *has a pure Nash equilibrium.*

Remark 4. It is worth to mention that the zero-sum version of Theorem 6 is much stronger. Then the assumption about the strong quasi-concavity is not needed ([11]). The zero-sum version of Theorem 7 one can find in [8] (Theorem 2.2 there). It was shown there that the assumption of that theorem can be remarkably weakened, by taking $k \leq 3$ and $l \leq 3$. However in the non-zero-sum case it does not suffice, which is shown in the next example.

Example 1. According to what we said in the last remark, it is enough to assume that all (2×2), (2×3) and (3×2)-subgames of a strongly quasi-concave zero-sum game have pure Nash equilibria, and the consequence is that the entire game has a solution of the same type. But consider the two-person non-zero-sum strongly quasi-concave game with the payoff matrices:

$$A = \begin{bmatrix} 4 & 1 & 1 & 1 \\ 3 & 1 & 1 & 2 \\ 2 & 1 & 1 & 3 \\ 1 & 1 & 1 & 4 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \\ 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{bmatrix}.$$

Note that all (2×2), (3×2) and (2×3)-subgames have pure Nash equilibria but the entire game does not. It shows that the assumptions of Theorem 7 cannot be weakened.

Remark 5. The strong quasi-concavity in Theorem 6 cannot be replaced by quasi-concavity. Moreover, strict quasi-concavity in Theorem 8 cannot be replaced by the weaker property, strong quasi-concavity. This is widely discussed in Remark 2.2 of [8].

At the end of this section we give four theorems about the existence of two-point Nash equilibria in bimatrix games. Their proofs one can find in a recent paper of Połowczuk [7]. The zero-sum version of these results were earlier obtained by Radzik [10]. The theorems are described in terms of two-point and two-adjoining point strategies (see Definition 1).

Theorem 9. *Let* $\Gamma(A, B)$ *be a concave bimatrix game. Then, there exist a Nash equilibrium* (μ_1, μ_2) *in this game such that* μ_1 *and* μ_2 *are two-adjoining-point strategies.*

Theorem 10. *Let $\Gamma(A, B)$ be a column-concave bimatrix game. Then, there exist a Nash equilibrium (μ_1, μ_2) in this game such that μ_1 is a two-adjoining-point strategy and μ_2 is a two-point strategy.*

Theorem 11. *Let $\Gamma(A, B)$ be a convex bimatrix game. Then, there exist a Nash equilibrium (μ_1, μ_2) in this game such that μ_1 and μ_2 are two-point strategies of the form $\mu_1 = \lambda\delta_1 + (1-\lambda)\delta_m$ and $\mu_2 = \gamma\delta_1 + (1-\gamma)\delta_n$, for some $0 \le \lambda \le 1$ and $0 \le \gamma \le 1$.*

Theorem 12. *Let $\Gamma(A, B)$ be a column-convex bimatrix game. Then, there exist a Nash equilibrium (μ_1, μ_2) in this game such that μ_1 and μ_2 are two-point strategies with $\mu_1 = \lambda\delta_1 + (1-\lambda)\delta_m$ for some $0 \le \lambda \le 1$.*

Remark 6. One could ask about the procedure allowing to simply find a Nash equilibrium determined by Theorems 9–12. In [7] and [10] one can find such procedures for matrix and bimatrix games, respectively. It appears that for matrix games these procedures are much simpler than for bimatrix games.

Remark 7. One could see Theorem 9 as a discrete counterpart of a two-person "concave" version of Theorem 1. Similarly, Theorem 10 is a "discrete" counterpart of Theorem 4. As far as the last two theorems are concerned, they can be seen as discrete counterparts of Theorem 5.

Remark 8. Theorem 1 would suggest that the assumption of Theorem 9 on concavity of a bimatrix game $\Gamma(A, B)$ could be weakened to quasi-concavity of $\Gamma(A, B)$. However, it is rather a big surprise that Theorem 9 is not longer true when we replace concavity by quasi-concavity. Then a Nash equilibrium in two-adjoining-point strategies may not exist. This is shown in the next example.

Example 2. Consider a bimatrix game $\Gamma(A, B)$ with payoff matrices of the form:

$$A = \begin{bmatrix} 5 & 3 & -1 \\ -1 & 1 & 3 \\ -3 & -1 & 5 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} -1 & 0 & 2 \\ 2 & 1 & 0 \\ 3 & 2 & -1 \end{bmatrix}.$$

It is easily seen (by Proposition 2) that this game is quasi-concave. However, the only Nash equilibrium in this game consists of the following three-point strategies: $\mu^* = (1/2)\delta_1 + (1/4)\delta_2 + (1/4)\delta_3$ and $\nu^* = (1/4)\delta_1 + (1/4)\delta_2 + (1/2)\delta_3$, which contradicts our earlier supposition expressed in Remark 8. Therefore, in fact, the assumption on concavity cannot be replaced by quasi-concavity.

4 Concave n-Person Finite Games

In this section we discuss the question if there is a discrete n-person counterpart of Theorem 1. By Theorem 9 we know that such a discrete counterpart exists in a two-person concave version, and concavity cannot be weakened to quasiconcavity (Remark 8). This strongly suggests that the answer for that question is positive in the "concave case". Just this fact will be shown in our main theorem of this section (Theorem 13).

4.1 Main Theorem

Before expressing the theorem, we need to introduce some new definitions and notation.
An *n-person non-zero-sum finite game* $\varGamma_N$ will be denoted in the sequel by the following normal form

$$\varGamma_N \;=\; \langle N,\, \{E_i\}_{i\in N},\; \{H_i\}_{i\in N}\rangle\,, \tag{7}$$

where (1) $N = \{1, 2, \ldots, n\}$ is the set of players;
(2) for each $i \in N$, $E_i = \{1, 2, \ldots, k_i\}$ is a finite space of Player's i *pure strategies* e_i, with k_i a natural number;
(3) for each vector $e = (e_1, e_2, \ldots, e_n)$ of the players' pure strategies and for each i, $H_i(e)$ is the *payoff function* of Player i in situation e.
For simplicity we will use the notation

$$E_{-i} = \prod_{j=1}^{i-1} E_j \times \prod_{j=i+1}^{n} E_j\,, \qquad e_{-i} = (e_1, \ldots, e_{i-1}, e_{i+1}, \ldots, e_n)$$

and

$$(e_{-i}, t) = (e_1, \ldots, e_{i-1},\, t,\, e_{i+1}, \ldots, e_n)\,.$$

We now give a basic definition of convexity of payoff functions H_i in their variables which change over finite sets E_i, $i \in N$. It is consistent with Definition 2 and considers the properties being a discrete counterpart of the concavity of functions F_i assumed in Theorem 1.

Definition 5. *A payoff function H_i of player i in game $\varGamma_N$, $i \in N$, is concave (quasi-concave) in its variable if for $j = 1, 2, \ldots, n$ there are strictly increasing sequences $x^j = (x^j_1, x^j_2, \ldots, x^j_{k_j})$ in $[0, 1]$, and if there exists a continuous function $F(x_1, \ldots, x_n)$ on $[0, 1]^n$, concave (quasi-concave) in variable x_i, such that for all $(e_1, e_2, \ldots, e_n) \in E_1 \times E_2 \times \ldots E_n$, $F(x^1_{e_1}, x^2_{e_2}, \ldots, x^n_{e_n}) = H_i(e_1, e_2, \ldots, e_n)$.*

Remark 9. One could think that the above definition would be more natural if the set $[0, 1]^n$ was replaced by $conv(E_1 \times \cdots \times E_n)$ and the sequences x^j were taken as constant ones of the form $x^j = (1, 2, \ldots, k_j)$, for $j = 1, \ldots, n$. However this second approach is less general than given in Definition 5 and leads to much smaller classes of payoff functions H_i concave in their variables. As far as quasi-concave functions H_i are concerned, both approaches lead to the same class.

It is easily seen that for convexity properties expressed by Definition 5, we can, practically, repeat Propositions 1 and 2 for payoff function H_i of game Γ_N as follows.

Proposition 4. *A payoff function H_i of player i in game Γ_N, $i \in N$, is concave if and only if there exist positive numbers $\lambda_1, \lambda_2, \ldots, \lambda_{k_i-1}$ such that for each $e_{-i} \in E_{-i}$,*

$$\lambda_1[H_i(e_{-i},2) - H_i(e_{-i},1)] \geq \lambda_2[H_i(e_{-i},3) - H_i(e_{-i},2)] \geq \ldots \\ \geq \lambda_{k_i-1}[H_i(e_{-i},k_i) - H_i(e_{-i},k_i-1)]\,.$$

Proposition 5. *A payoff function H_i of player i in game Γ_N, $i \in N$, is quasi-concave if and only if for each $e_{-i} \in E_{-i}$, there exists natural l, $1 \leq l \leq k_i$ such that*

$$H_i(e_{-i},1) \leq \cdots \leq H_i(e_{-i},l) \geq H_i(e_{-i},l+1) \geq \cdots \geq H_i(e_{-i},k_i)\,. \quad (8)$$

Now we are ready to formulate the main theorem of this section which is a discrete counterpart of Glicksberg Theorem 1. Let Γ_N be an n-person non-zero-sum finite game of the form (7).

Theorem 13. *If each of the payoff functions $H_1, \ldots, H_n$ is concave in its variable then there exists a mixed Nash equilibrium $(\mu_1^*, \ldots, \mu_n^*)$ in the game Γ_N such that $\mu_1^*, \ldots, \mu_n^*$ are two-adjoining-point strategies.*

Remark 10. Theorem 1 would suggest that the assumption of Theorem 13 on concavity of payoff functions in their variables could be weakened to quasi-concavity. However in its discrete counterpart it is not true, it is shown in Example 2 for the simplest case $n = 2$.

4.2 Proof

Before we start proving Theorem 13, we must introduce some new notation. For the sets $\mathcal{N}$ of integers and real x we define

$$\lfloor x \rfloor = \max\{z \in \mathcal{N} \,|\, z \leq x\}, \qquad \lceil x \rceil = \min\{z \in \mathcal{N} \,|\, z \geq x\}.$$

Let us define the following auxiliary game G of the form

$$G = \left\langle N, \{E_i^G\}_{i \in N}, \{H_i^G\}_{i \in N} \right\rangle,$$

where the strategy spaces E_i^G and payoff functions H_i^G in the game G are defined by the following:

$$E_i^G = [1, k_i] \qquad \text{for} \qquad i = 1, 2, \ldots, n$$

and for $x \in E^G = \prod_{i=1}^n E_i^G$,

$$H_i^G(x) = \sum_{I \subset N} \left(\prod_{i \in I} \alpha_{x_i} \right) \left(\prod_{j \notin I} (1 - \alpha_{x_j}) \right) H(x_1^I, x_2^I, \ldots, x_n^I)$$

where

$$x_i^I = \begin{cases} \lfloor x_i \rfloor \text{ if } i \in I \\ \lceil x_i \rceil \text{ if } i \notin I, \end{cases} \quad \text{and} \quad \alpha_{x_i} = \lceil x_i \rceil - x_i.$$

Assume for a moment that in the game Γ_N, the players use any of their pure strategies described by e. One can easily check that for such a strategy profile, the values of the payoff functions in both games Γ_N and G are equal, that is, $H_i(e) = H_i^G(e)$ for all $i \in N$. On the other hand, the set of pure strategies in game G is much richer than the same in the game Γ_N. However, as it will appear (the next lemma) any profile of pure strategies in the game G is "equivalent" to some profile of two-adjoining-point strategies in Γ_N. Just this is formulated by the next lemma.

Lemma 1. *There exists a 1–1 correspondence between the set of pure strategies in game G and the set of two-adjoining-point strategies in game Γ_N such that, for every pure strategy profile of the form $x = (x_1, x_2, \ldots, x_n)$ in game G, the corresponding two-adjoining-point strategy profile $\pi = (\pi_1, \pi_2, \ldots, \pi_n)$ in game Γ_N satisfies: $H_i^G(x) = H_i(\pi)$ for every $i \in N$.*

Proof. For any $i \in N$ and any $x_i \in E_i^G$, let

$$a(x_i) = \alpha_{x_i} \delta_{\lfloor x_i \rfloor} + (1 - \alpha_{x_i}) \delta_{\lceil x_i \rceil},$$

where α_{x_i} is described in (9). We define the correspondence:

$$T(x_1, x_2, \ldots, x_n) = (a(x_1), a(x_2), \ldots, a(x_n)).$$

One can easily check that for any $j \in N$

$$H_j^G(x_1, x_2, \ldots, x_n) = H_j(T(x_1, x_2, \ldots, x_n)).$$

On the other hand, if there exist two $x_i^1, x_i^2 \in E_i^G$, such that $x_i^1 \neq x_i^2$, then $\lfloor x_i^1 \rfloor \neq \lfloor x_i^2 \rfloor$ or $\alpha_{x_i^1} \neq \alpha_{x_i^2}$, which implies that $a(x_i^1) \neq a(x_i^2)$, and consequently the correspondence T gives different values for different arguments. Moreover, the correspondence T is also of "onto" type . Namely, for every two-adjoining-point strategy $\pi_i = \alpha\delta_{y_i} + (1 - \alpha)\delta_{y_i+1}$ of player i in game Γ_N, there exists a pure strategy $x_i^\pi = \alpha y_i + (1 - \alpha)(y_i + 1)$ in game G, such that $a(x_i^\pi) = \pi_i$. Thereby the correspondence T is 1–1.
The next fact may be seen as a simple consequence of Lemma 1.

Lemma 2. *If there exists a pure-strategy Nash equilibrium in game G, then there also exists a two-adjoining-point Nash equilibrium in game Γ_N.*

Now we are ready to come back to the proof of Theorem 13. In view of Lemma 2, to complete the proof we only need to show that game G has a pure-strategy Nash equilibrium. First, note that the sets E_i^G for $i \in N$ are all nonempty, convex and compact.

Now, in view of Theorem 1, we have to show only that each function H_i^G for $i = 1, \ldots, n$ is concave in its variable, that is in variable x_i. Let us fix $i \in N$ and let x be variable in E^G. By Lemma 1 we know that there is a strategy profile π in game Γ consisting of only two-adjoining-point strategies, equivalent to x. We have the following sequence of equalities:

$$H_i^G(x) = H_i^G(x_{-i}, x_i) = H_i(\pi_{-i}, \pi_i)$$

$$= \alpha_{x_i} H_i(\pi_{-i}, \lfloor x_i \rfloor) + (1 - \alpha_{x_i}) H_i(\pi_{-i}, \lceil x_i \rceil),$$

where α_{x_i} is of the form (9). Denote the last convex combination by $B(\pi_{-i}, x_i)$. Now, one can easily see that the function $B(\pi_{-i}, x_i)$ is a linear function of variable x_i in each interval $[l, l+1]$, $l = 1, 2, \ldots, k_i - 1$. On the other hand, we have $B(\pi_{-i}, x_i) = H_i(\pi_{-i}, x_i)$ for any integer $x_i \in E_i$. By our concavity assumption it follows that also the values $H_i(\pi_{-i}, 1), H_i(\pi_{-i}, 2), \ldots, H_i(\pi_{-i}, k_i)$ satisfy the inequalities (8). Hence, we easily deduce that the function $B(\pi_{-i}, x_i)$ is quasi-concave in variable x_i, and thereby, also the function H_i^G is quasi-concave in variable x_i. This means that game G satisfies the assumptions of the Glicksberg Theorem 1, and thereby there exists a pure-strategy Nash equilibrium in this game, finally ending the proof of Theorem 13.

References

1. Bohnenblust HF, Karlin S, Shapley LS (1950) Games with continuous, convex pay-off. In: Kuhn HW, Tucker AW (eds) Contributions to the theory of games, I. Annals of mathematics Studies, Princeton University Press, Princeton 24: 181–192
2. Debreu G (1952) A social equilibrium existence theorem. Proceedings of the National Academy of Science USA 38: 886–893
3. Fan K (1952) Fixed point and minimax theorems in locally convex topological linear spaces. Proceedings of the National Academy of Science USA 38: 121–126
4. Glicksberg IL (1952) A further generalization of the Kakutani fixed point theorem with application to Nash equilibrium points. Proceedings of the American Mathematical Society 38: 170–174
5. Nash FJ (1950) Equilibrium points in N-person games. Proceedings National Academy of Science, U.S.A. 36: 48–49
6. Połowczuk W (2003) Pure Nash Equilibria in Finite Two-Person Non-Zero-Sum Games. International Journal of Game Theory, 32: 229–240
7. Połowczuk W (2004) On two-point Nash equilibria in bimatrix games with convexity properties. Applicationes Mathematicae
8. Radzik T (1991) Saddle Point Theorems. International Journal of Game Theory 20: 23–32

9. Radzik T (1993) Nash equilibria of discontinuous non-zero-sum two-person games. International Journal of Game Theory 21: 429–437
10. Radzik T (2000) Characterization of optimal strategies in matrix games with convexity properties. International Journal of Game Theory 29: 211–227
11. Shapley LS (1964) Some topics in two-person game. Advances in Game Theory, Ann. of Math. Studies 52: 1–28
12. Sion M (1958) On general minimax theorems. Pacific Journal of Mathematics 8: 171–176

A Spectral Approach to Solve Box-constrained Multi-objective Optimization Problems

Maria Cristina Recchioni[1]

Dipartimento di Scienze Sociali “D. Serrani”, Piazzale Martelli 8, 60121 Ancona
m.c.recchioni@univpm.it

Summary. This paper presents some first and second order conditions necessary for the Pareto optimality of box-constrained multi-objective optimization problems. These necessary conditions are related to the spectrum of a matrix defined via the gradient vectors and the Hessian matrices of the objective functions. These necessary conditions are used to develop two algorithms. The first one is built taking into account the first order necessary conditions and determines some critical points for the multi-objective problems considered. The second one is based on the second order necessary conditions and discards the critical points that do not belong to the local Pareto optimal front. Some numerical results are shown.

Key words: Multi-objective optimization problems, path following methods, dynamical systems, spectral analysis.

1 Introduction

Let $\mathbf{R}^n$ be the n-dimensional real Euclidean space, $x = (x_1, x_2, \ldots, x_n)^T \in \mathbf{R}^n$ be a generic vector, where the superscript T means transpose. Let x, $y \in \mathbf{R}^n$, we denote with $y^T x$ the Euclidean scalar product, $\|x\| = (x^T x)^{1/2}$ the Euclidean norm and we denote with the symbols $x < y$, $x \leq y$, x, $y \in \mathbf{R}^n$ the inequalities componentwise, that is: $x_i < y_i$, $x_i \leq y_i$, $i = 1, 2, \ldots, n$.
Let us define the box-constrained multi-objective optimization problem.
Let $B \subset \mathbf{R}^n$ be the following box:

$$B = \{\ x \in \mathbf{R}^n \mid l \leq x \leq u\ \} \tag{1}$$

where $l \in \mathbf{R}^n$, $u \in \mathbf{R}^n$, $l \leq u$ are two given vectors, and let $E \subset \mathbf{R}^n$ be an open set with $B \subset E$.
Let $F = (F_1, F_2, \ldots, F_s)^T$, $s \leq n$, be a vector valued function and let $F_i : E \subseteq \mathbf{R}^n \longrightarrow \mathbf{R}$, $i = 1, 2, \ldots, s$, be continuously differentiable functions in the open set E, we consider the following problem:

$$\min_{x \in B} F(x). \tag{2}$$

The solution of problem (2) is represented by the global Pareto optimal points.

Definition 1. *A point $x^* \in B$ is a "global" Pareto optimal point when it satisfies the following condition:*

$$\nexists x \in B \text{ with } F(x) \leq F(x^*) \text{ and } F(x) \neq F(x^*). \tag{3}$$

Definition 2. *A point x^* is called a "local" Pareto optimal point if there exists a neighbourhood $U \subseteq \mathbf{R}^n$ of x^* such that the following condition is satisfied:*

$$\nexists x \in B \cap U \text{ with } F(x) \leq F(x^*) \text{ and } F(x) \neq F(x^*). \tag{4}$$

In this paper we formulate some necessary conditions for Pareto optimality via the spectrum of a suitable matrix related to the jacobian matrix of the objective function F. The first order necessary conditions proposed allow us to formulate a suitable computational method (we will refer to it as algorithm A_1) based on sequences of feasible points $\{ x^k \}$, $k = 0, 1, \ldots$, $x_0 \in int\, B$. We will prove that these sequences have accumulation points that are critical points for problem (2). Remember that a point $\tilde{x}$ is an accumulation point of the sequence $\{ x^k \}$ if there exists a subsequence $\{ x^{k_j} \}$ such that $\lim_{j \to +\infty} x^{k_j} = \tilde{x}$. Algorithm A_1 is a kind of interior point method for vector optimization problems that does not require any "a priori" scalarization of the objectives F_i, $i = 1, 2, \ldots, s$ and that determines some critical points for the multi-objective optimization problem. Finally, using the second order conditions we develop an algorithm (we will refer to it as algorithm A_2) that is able to establish when a critical point determined by algorithm A_1 is not a Pareto minimal point.

The numerical experiments proposed in Section 4 suggest to us that the joint use of algorithm A_1 and algorithm A_2 allows us to approximate the entire local Pareto front of the vector optimization problems considered when the accumulation points of a sufficiently large number of sequences starting from initial points suitably distributed on the box are computed. This is only an empirical result.

We note that when we choose an "a priori" scalarization of the vector objective function without having some convexity property on F_i, $i = 1, 2, \ldots, s$, we cannot approximate the entire Pareto front minimizing the scalar function. To overcome this difficulty we should minimize several weighted sums of the objective functions trying to choose the weights in order to get the entire local Pareto front. This choice of the weights is not easy in practice since the dependence of the Pareto front on these weights is not obvious.

Recently, several papers have been devoted to formulate well known methods of the scalar optimization for vector optimization, such as steepest descent methods [1], [2], [3], proximal methods [4], differential inclusion techniques [5], genetic algorithms [6], [7], tabu search [8].

The results presented in this paper are based on some ideas introduced in [3] in the context of the vector optimization problems and in [9], [10], [11], [12] in the context of scalar optimization problems. We must remember that several scalarization procedures have been introduced to solve multi-objective optimization problems, see for example [13], [14], [15], [16], and, more recently, [17]. Only in the last years can we find papers that propose computational methods to solve vector optimization problems without using any "a priori" scalarization of the original vector function, see for example [1], [18], [19], [4], [5], [8], [20], [21].
In Section 2 we give the necessary conditions for Pareto minimal points of problem (2). In Section 3 we derive the computational method from the necessary conditions formulated in Section 2. In Section 4 we show some numerical results obtained applying the computational method introduced in Section 3 to solve some test problems.

2 Necessary Conditions for Pareto Optimal Fronts

Let $F = (F_1, F_2, \dots, F_s)^T$ be a vector valued function, whose components F_i, $F_i : E \subseteq \mathbf{R}^n \longrightarrow \mathbf{R}$, $i = 1, 2, \dots, s$, $s \le n$, are assumed to be continuously differentiable in the open set E (see Lemmas 1-4) and to be twice continuously differentiable in the open set E (see Lemma 5).
We denote with ∇F_i, $i = 1, 2, \dots, s$, the gradient vector of F_i, that is:

$$\nabla F_i(x) = \begin{pmatrix} \frac{\partial F_i}{\partial x_1}(x) \\ \vdots \\ \frac{\partial F_i}{\partial x_n}(x) \end{pmatrix}, \quad i = 1, 2 \dots, s, \tag{5}$$

with $J_F(x)^T = (\nabla F_1(x) | \nabla F_2(x) | \cdots | \nabla F_s(x)) \in \mathbf{R}^{n \times s}$ the transposed matrix of the Jacobian matrix $J_F(x) \in \mathbf{R}^{s \times n}$ of F at the point x and with $H_{F_i}(x) \in \mathbf{R}^{n \times n}$ the Hessian matrix associated with the function F_i, $i = 1, 2, \dots, s$.
Let $l \in \mathbf{R}^n$ and $u \in \mathbf{R}^n$ be the vector given in (1), we denote with $D(x) \in \mathbf{R}^{n \times n}$ the diagonal matrix defined by:

$$(D(x))_{i,j} = \begin{cases} (x_i - l_i)^2 (u_i - x_i)^2 & i = j \\ 0 & i \ne j \end{cases}, \; l \le x \le u, \tag{6}$$

and with $M_F(x) \in \mathbf{R}^{s \times s}$, $x \in B$ the following matrix:

$$M_F(x) = J_F(x) D(x) J_F(x)^T, \; x \in B. \tag{7}$$

Remark 1. We note that when $x_i = u_i$ or $x_i = l_i$ for any $v \in \mathbf{R}^n$ the i-th component of the vector $h = D(x)v$ is equal to zero. That is the vector $x + D(x)v$ belongs to the affine space generated by the face of B that contains x.

Lemma 1. *Let F_i, $i = 1, 2, \dots, s$, be continuously differentiable functions in E, then the matrix $M_F(x)$, $x \in B$ given in (7) is a positive semi-definite matrix.*

Proof. See [3] Lemma 2.3.

Let $0 \leq \lambda^1(x) \leq \lambda^2(x) \leq \dots \leq \lambda^s(x)$, $x \in B$ be the eigenvalues of the matrix $M_F(x)$ and $v^1(x)$, $v^2(x), \dots, v^s(x)$ be a corresponding orthonormal basis of eigenvectors. Let us denote with $\lambda_{min,F}(x) = \lambda^1(x)$ and $\lambda_{max,F}(x) = \lambda^s(x)$ the smallest and the largest eigenvalues of the matrix $M_F(x)$, $x \in B$ respectively and with $v_{min,F}(x) = v^1(x)$ and $v_{max,F}(x) = v^s(x)$, $x \in B$ the corresponding orthonormal eigenvectors.

Lemma 2. *Let $F_i : E \subset \mathbf{R} \to \mathbf{R}$, $i = 1, 2, \dots, s$, be continuously differentiable functions and let x^* be a Pareto local optimal point for problem (2) then we have:*

$$\lambda_{min,F}(x^*) = 0. \tag{8}$$

Proof. The proof of (8) is a consequence of the fact that in a Pareto local optimal point the matrix $M_F(x)$ must be singular (see [3] Lemma 2.4).

Let P_g, P_l be the sets of the global and local Pareto optimal points respectively and finally let V_{min} and V_{max} be the following sets:

$$V_{min} = \{\ x \in B \mid \lambda_{min,F}(x) = 0\ \}, \tag{9}$$

$$V_{max} = \{\ x \in B \mid \lambda_{max,F}(x) = 0\ \}. \tag{10}$$

It is easy to see that we have $P_g \subseteq P_l \subseteq V_{min}$ while, in general, we can not establish a relation between the global or local Pareto optimal fronts and the set V_{max}.
Let $\mathcal{F}(x)$, $x \in B$ be the set of the feasible directions at x that is:

$$\mathcal{F}(x) = \{\, h \in \mathbf{R}^n \mid \exists t > 0,\ x + th \in B\, \}, \tag{11}$$

an obvious necessary condition for the local Pareto optimality of $x \in B$ is:

$$Image_{\mathcal{F}(x)}(J_F(x)) \cap (-\mathbf{R}^s_+) = \emptyset, \tag{12}$$

where $Image_{\mathcal{F}(x)}(J_F(x))$ is the image of $\mathcal{F}(x)$ by the linear operator $J_F(x)$. Condition (12) is the analogous of the following necessary condition for unconstrained optimization problems (see [23]):

$$Image(J_F(x)) \cap (-\mathbf{R}^s_+) = \emptyset, \tag{13}$$

where $Image(J_F(x))$ is the image of $\mathbf{R}^n$ by the linear operator $J_F(x)$. In fact when condition (13) is not satisfied there exists a vector $h \in \mathbf{R}^n$ satisfying

$J_F(x)h < 0$ and this implies that the direction h is a descent direction, that is we can decrease all the objective functions at once.
Let us illustrate the results of Lemma 2 in the case of a bi-criteria unconstrained optimization problem. In this case the matrix $M_F(x)$ given in (7) has the following form:

$$M_F(x) = \begin{pmatrix} \|\nabla F_1(x)\|^2 & \nabla F_1(x)^T \nabla F_2(x) \\ \nabla F_1(x)^T \nabla F_2(x) & \|\nabla F_2(x)\|^2 \end{pmatrix}, \qquad (14)$$

Note that equation $\lambda_{min,F}(x) = 0$ implies the following equation:

$$\|\nabla F_1(x)\|^2 \, \|\nabla F_1(x)\|^2 (1 - \cos^2(\theta(x))) = 0, \qquad (15)$$

where $\theta(x)$ is the angle between the two gradient vectors $\nabla F_1(x)$, $\nabla F_2(x)$. Equation (15) implies that one of the following equations holds:

$$\begin{array}{l} \|\nabla F_1(x)\| = 0, \\ \|\nabla F_2(x)\| = 0, \\ \cos(\theta(x)) = \pm 1. \end{array} \qquad (16)$$

The first two equations are trivially necessary conditions for Pareto optimality. The third equation implies that $\theta(x) = \pi$ or $\theta(x) = 0$. Obviously when $\theta(x) = 0$ we can decrease all the objective functions, while when $\theta(x) = \pi$ the two gradient vectors point in opposite directions so we cannot decrease one of the objective functions without increasing the other one.
Note that when $\lambda_{min,F}(x) = 0$ and $\lambda_{max,F}(x) > 0$ if there exists an eigenvector $v_{max,F}(x)$ associated with $\lambda_{max}(x)$ satisfying the condition $v_{max,F}(x) > 0$ then the direction $h = -J_F(x)^T v_{max,F}(x)$ is a descent direction (i.e. $J_F(x)h = -\lambda_{max,F}(x) v_{max,F} < 0$) so that the point x does not belong to the local Pareto optimal front. The example shows that it is relevant to investigate the spectrum of the matrix $M_F(x)$ (see (7)) when the point $x \in B$ is a critical point, that is $x \in V_{min}$. Hence we investigate the feasible singular points of the matrix $M_F(x)$ in order to establish when a feasible singular point of the matrix $M_F(x)$ surely does not belong to the local Pareto optimal front.
Let $x \in V_{min}$ and $x \notin V_{max}$, and let j, $j = 2, 3, \ldots, s$ be the index of the first positive eigenvalue of $M_F(x)$, that is $\lambda^j(x) > 0$ and $\lambda^i(x) = 0$, $i = 1, 2, \ldots, j-1$. We introduce the following set:

$$\mathcal{C}_j(x) = \left\{ v \in \mathbf{R}^s \,|\, v = \sum_{i=j}^{s} c_i v^i(x),\; c_i \in \mathbf{R},\; i = j, j+1, \ldots, s, \right.$$
$$\left. \mathit{and} \sum_{i=j}^{s} c_i \lambda^i(x) v^i(x) > 0 \right\}, \quad x \in V_{min}, \qquad (17)$$

we have:

Lemma 3. *Let $F_i : E \subset \mathbf{R}^n \to \mathbf{R}$, $i = 1, 2, \dots, s$ be continuously differentiable functions in E with $B \subset E$, let $\tilde{x} \in B$ be a point such that $\lambda_{min,F}(\tilde{x}) = 0$ and let us assume that there exists an integer j, $1 < j \leq s$ such that we have $\lambda^j(\tilde{x}) > 0$. Let $\tilde{x} \in B$ be a local Pareto optimal point then we have:*

$$\mathcal{C}_j(\tilde{x}) = \emptyset \tag{18}$$

Proof. We note that for $\tilde{x} \in V_{min}$ the following set:

$$\mathcal{A}(\tilde{x}) = \{ h \in \mathbf{R}^n \,|\, h = -D(\tilde{x}) J_F(\tilde{x})^T v, \;\; v \in \mathbf{R}^s \}, \tag{19}$$

is a subset of the set $\mathcal{F}(\tilde{x})$ of the feasible direction at $\tilde{x}$ defined in (11) (see Remark 1). Hence if there exists $v \in \mathcal{C}_j(\tilde{x})$ moving along the direction $h = -D(\tilde{x}) J_F(\tilde{x})^T v \in \mathcal{A}(\tilde{x})$ we have:

$$F_k(\tilde{x} + th) = F_k(\tilde{x}) - t \sum_{i=j}^{s} c_i \lambda^i(\tilde{x}) v_k^i(\tilde{x}) + o(t) < F_k(\tilde{x}),\, t \to 0^+. \tag{20}$$

This concludes the proof.

An easy consequence of Lemma 3 is the following lemma:

Lemma 4. *Let $F_i : E \subset \mathbf{R}^n \to \mathbf{R}$, $i = 1, 2, \dots, s$ be continuously differentiable functions in E, $B \subset E$, let $\tilde{x} \in B$ be a point such that the matrix $M_F(\tilde{x})$ is singular that is $\lambda_{min,F}(\tilde{x}) = 0$, and $\lambda_{max,F}(\tilde{x}) > 0$. If there exists a one-sign eigenvector associated to some non-zero eigenvalues of the matrix $M_F(\tilde{x})$ then the point $\tilde{x}$ does not belong to the local Pareto optimal front.*

We conclude this section introducing a necessary condition involving the second order partial derivatives of the objective functions, so that we assume that the objective functions F_i, $i = 1, 2, \dots, s$, are twice continuously differentiable functions in the open set E.
For $h \in \mathbf{R}^n$ and $x \in V_{min}$ we define the matrix $D_x(h) \in \mathbf{R}^{n \times n}$ as follows:

$$D_x(h)_{i,i} = \begin{cases} 1, & x_i \in (l_i, u_i), \\ -h_i, & x_i = u_i, \\ h_i, & x_i = l_i, \end{cases} \qquad i = 1, 2, \dots, n, \tag{21}$$

$$D_x(h)_{i,j} = 0,\, i \neq j,\, i, j = 1, 2, \dots, n\,. \tag{22}$$

We note that for any $h \in \mathbf{R}^n$ the vector $\tilde{h} = D_x(h)h$ is a feasible direction at x so that the set $\mathcal{F}(x)$ of the feasible direction at x (see (11)) can be rewritten as follows:

$$\mathcal{F}(x) = \{ h \in \mathbf{R}^n \,|\, h = D_x(b)b, \;\; b \in \mathbf{R}^n \}, \;\; x \in B \tag{23}$$

and we define a set that contains the set of descent directions at x, that is:

$$\mathcal{D}(x) = \{ h \in \mathbf{R}^n \mid J_F(x) D_x(h) h \le 0 \}, \; x \in B. \tag{24}$$

Now we can state the second order necessary condition.
Let $g_x(h) \in \mathbf{R}^s$, $g_{i,x} : \mathbf{R}^n \to \mathbf{R}$, $i = 1, 2 \ldots, s$ be the function defined as follows:

$$g_{i,x}(h) = \nabla F_i(x)^T D_x(h) h + \frac{1}{2} h^T D_x(h) H_{F_i}(x) D_x(h) h, \; i = 1, 2, \ldots, s, \tag{25}$$

we have:

Lemma 5. *Let $F_i : E \subset \mathbf{R}^n \to \mathbf{R}$, $i = 1, 2, \ldots, s$ be twice continuously differentiable functions in E, $B \subset E$, let $\tilde{x}$ be a local Pareto optimal solution of problem (2) then we have:*

$$Image_{\mathcal{D}(\tilde{x})}(g_{\tilde{x}}) \cap (-\mathbf{R}^s_+) = \emptyset \tag{26}$$

where $Image_{\mathcal{D}(\tilde{x})}(g_{\tilde{x}})$ is the image of the set $\mathcal{D}(\tilde{x})$ (see (24)) by the map $g_{\tilde{x}}$.

Proof. Let us assume by contradiction that there exists $h \in \mathcal{D}(\tilde{x})$ such that we have $g_{\tilde{x}}(h) < 0$. This allows us to see that we have $g_{\tilde{x}}(th) < 0$, $t \in (0, 1)$. Hence using Taylor expansion with base point $\tilde{x}$ we obtain:

$$F_i(\tilde{x} + t\,h) = F_i(\tilde{x}) + g_{i,\tilde{x}}(th) + o(t^2 \|h\|^2) < F_i(\tilde{x}), \; t \to 0^+, \; i = 1, 2, \ldots, s, \tag{27}$$

where $o(\cdot)$ is the Landau symbol. Equation (27) is absurd. This concludes the proof.

As we will show in the next section, Lemma 5 allows us to formulate a computational method to test condition (26) and to discard a critical point $\tilde{x}$ when we find $h \in \mathcal{D}(\tilde{x})$ such that $g_{\tilde{x}}(h) < 0$. Future work will be needed to study how the class of functions satisfying (26) is related to well known kinds of generalized convexity (see [23], [24], [25], [26] for further details).

3 A Computational Method to Find Local Pareto Optimal Fronts

In this section we provide a practical tool to determine the local Pareto optimal solutions of problem (2) thanks to the use of Lemma 2 and Lemma 5. First of all we determine a set of critical points for problem (2) and later we discard the points of this set that do not match condition (26) of Lemma 5. In particular we determine a subset V^*_{min} of the set V_{min} (see (9)) defined as follows:

$$V^*_{min} = \{ x \in V_{min} \mid x \text{ is not a Pareto maximal point} \}. \tag{28}$$

We show that the points of the set V^*_{min} are the accumulation points of a suitable infinite sequence of feasible points. That is we formulate a kind of interior point steepest descent method for the box constrained multi-objective optimization problem (2) and then we investigate the points of the set V^*_{min} using Lemma 5.

Definition 3. *Let $x \in B$ we say that a feasible direction $h \in \mathcal{F}(x)$ is a "descent direction" at x when we have $J_F(x)h < 0$.*

Definition 4. *Let $x \in B$ we say that a feasible direction $h \in \mathcal{F}(x)$ is a "scalarization compatible" at x when there exists $w \in \mathbf{R}^s$ such that we have:*

$$h = -D(x)J_F(x)^T w. \tag{29}$$

The set $\mathcal{A}(x)$, $x \in B$ given in (19) is the set of the "scalarization compatible" directions at x. Definition 4 is a modification of the definition proposed in [2] for the unconstrained case. The meaning of the term "scalarization compatible" is a consequence of the fact that a scalarization compatible direction can be interpreted as the projected gradient of a suitable scalarization of the vector objective function, that is:

$$h = -D(x)J_F(x)^T w = -D(x)\nabla_x w^T F(x), \tag{30}$$

where ∇_x denotes the gradient vector with respect to the variable x and the scalar product $w^T F(x) = \sum_{i=1}^{s} w_i F_i(x)$ is the scalarization of the vector function F at the point x.
We look for a vector valued function $h(x)$, $x \in B$ satisfying the following requirements:

(i) $h(x)$ must be a scalarization compatible direction at x, i.e.: equation (29) holds for some w;
(ii) $h(x)$ must be a descent direction at x,
(iii) $J_F(x)h(x)$ must be a continuous function on B.

Let be $x \in B$ and $x \notin V_{min}$, let $M_F(x) \in \mathbf{R}^{s\times s}$ be the matrix defined in (7) and let

$$M_F(x) = Q_F(x)\Lambda(x)Q_F(x)^T, \tag{31}$$

be its spectral decomposition, where:

$$\begin{aligned} \Lambda(x) &= Diag\left(\lambda_{min,F}(x), \lambda^2(x), \ldots, \lambda^{s-1}(x), \lambda_{max,F}\right) \in \mathbf{R}^{s\times s}, \\ Q_F(x) &= \left(v^1(x)|v^2(x)|\cdots|v^s(x)\right). \end{aligned} \tag{32}$$

Let $e \in \mathbf{R}^s$ be the vector whose components are equal to one (i.e. $e = (1,1,\ldots,1)^T \in \mathbf{R}^s$), we can define the vector valued function $h(x)$

$$h(x) = -\lambda_{min,F}(x)D(x)J_F(x)^T w(x), \tag{33}$$

where $w(x) \in \mathbf{R}^s$ is given by:

$$w(x) = Q_F(x)Diag\left(1, \frac{\lambda_{min,F}(x)}{\lambda_2(x)}, \frac{\lambda_{min,F}(x)}{\lambda_3(x)}, \ldots, \frac{\lambda_{min,F}(x)}{\lambda_{max,F}(x)}\right) Q_F(x)^T e\,. \tag{34}$$

It is easy to see that $h(x)$ is feasible direction at x satisfying the requirements (i), (ii) and (iii).

Using standard linear algebra we have:

$$\begin{aligned}J_F(x)h &= -\lambda_{min,F}(x)J_F(x)D(x)J_F(x)^T w = -\lambda_{min,F}(x)M_F(x)w\\ &= -\lambda_{min,F}(x)Q_F(x)\Lambda(x)Q_F(x)^T Q_F(x)\cdot\\ &\cdot Diag\left(1, \frac{\lambda_{min,F}(x)}{\lambda_2(x)}, \frac{\lambda_{min,F}(x)}{\lambda_3(x)}, \dots, \frac{\lambda_{min,F}(x)}{\lambda_{max,F}(x)}\right) Q_F(x)^T e\\ &= -\lambda_{min,F}(x)^2 e, \end{aligned} \tag{35}$$

hence, roughly speaking, moving along the direction $h(x)$ all the objective functions at once decrease by a quantity $\lambda_{min,F}(x)^2$.
Now we are able to define the algorithm to determine the set V^*_{min} (see (28)) and we refer to it as algorithm A_1.

Algorithm A_1

Step 1. Choose an initial point $x^0 \in int\, B$, a positive constant ϵ, $0 < \epsilon << 1$, a positive integer it_{max} and set $k = 0$.

Step 2. If $\lambda_{min,F}(x^k) < \epsilon$ (i.e.: $x^k \in V^*_{min}$) or $k > it_{max}$ stop. Otherwise compute:

$$w^k = Q_F(x^k)Diag\left(1, \frac{\lambda_{min,F}(x^k)}{\lambda_2(x^k)}, \frac{\lambda_{min,F}(x^k)}{\lambda_3(x^k)}, \dots, \frac{\lambda_{min,F}(x^k)}{\lambda_{max,F}(x^k)}\right) Q_F(x^k)^T e, \tag{36}$$

and $h^k = -\lambda_{min,F}(x^k)D(x^k)J_F(x^k)^T w^k$ then proceed with Step 3.

Step 3. Compute $x^{k+1} = x^k + t_k h^k$, where t_k is the step-size computed with the usual backtracking procedure:

(1a) Set $t = 1$, $\beta = \frac{1}{2}$

(2a) if $F(x^k + t\, h^k) \le F(x^k) + \beta\, t J_F(x^k)h^k$ and $l < x^k + th^k < u$ then $t_k = t$ otherwise $t = t/2$ (end of backtracking)

Step 4. Set $k = k + 1$ and go to Step 2.

Lemma 6. *Let $F_i : E \subseteq \mathbf{R}^n \to \mathbf{R}$, $i = 1, 2, \dots, s$ be continuously differentiable functions in E, $B \subset E$, let $\{x^k\}$, $k = 0, 1, \dots$ be an infinite sequence generated by algorithm A_1. We have:*

(a) $\{\, F(x^k)\,\}$, $k = 0, 1, \dots$ is a monotonically not increasing sequence;
(b) the points x^k, $k = 0, 1, \dots$ belong to the interior of B;
(c) the sequence $\{\, x^k \,\}$ has at least one feasible accumulation point $\tilde{x}$, furthermore each accumulation point of $\{\, x^k \,\}$ belongs to V^*_{min};
(d) each accumulation point $\tilde{x}$ of $\{\, x^k \,\}$ satisfies the following inequality:

$$F(\tilde{x}) \le F(x^k),\ k = 0, 1, 2, \dots. \tag{37}$$

Proof. Let us prove (a). By virtue of Step 3 of algorithm A_1 we have (see the backtracking procedure (2a)):

$$F(x^{k+1}) = F(x^k + t_k h^k) \le F(x^k) + \beta\, t_k J_F(x^k)h^k,\ k = 0, 1, \dots. \tag{38}$$

Equation (38) and the fact that h^k is a descent direction at x^k (see equation (35)) imply assertion (a).
Let us prove (b). Note that x^k belongs to the interior of the feasible region since x^0 belongs to $int\, B$, h^k is a feasible direction at x^k and the step-size t_k satisfies condition (2a).
Let us prove (c) and (d). Since B is a compact set then there exists at least an accumulation point $\tilde{x}$ of the sequence $\{x^k\}$ and it belongs to B by virtue of assertion (b). That is part of assertion (c) is proved.
Now we prove assertion (d). Let $\{x^{k_j}\}$ be the subsequence of $\{x^k\}$ such that:

$$\lim_{j \to +\infty} x^{k_j} = \tilde{x}. \tag{39}$$

From equation (35) since F_i, $i = 1, 2, \ldots, s$ are continuously differentiable functions in the open set E containing B we have that $\lambda_{min,F}(x)$ is a continuous function of $x \in B$ (see [22]) and we have:

$$\lim_{j \to +\infty} J_F(x^{k_j}) h^{k_j} = - \lim_{j \to +\infty} \lambda_{min,F}(x^{k_j})^2 e = -\lambda_{min,F}(\tilde{x})^2 e. \tag{40}$$

Furthermore for any k there exists j such that $k < k_j$ so that we have:

$$F(x^{k_j}) \leq F(x^k), k = 0, 1, \ldots. \tag{41}$$

Assertion (d) that is formula (37) follows from (41) taking the limits $j \to +\infty$.
Now we are able to conclude the proof of (c). Since the sequence $\{F(x^k)\}$ is bounded below, from Step 3 of algorithm A_1 we obtain:

$$F(x^{k+1}) \leq F(x^0) - \beta \left(\sum_{i=0}^{k} t_i \, \lambda_{min,F}(x^i)^2 \right) e \,, \quad k = 0, 1, \ldots. \tag{42}$$

Equation (42) and the fact that the sequence $\{F(x^k)\}$ is bounded below imply the following inequality:

$$\beta \sum_{i=0}^{+\infty} t_i \, \lambda_{min,F}(x^i)^2 < +\infty \tag{43}$$

and as a consequence we have:

$$\lim_{j \to +\infty} t_{k_j} \, \lambda_{min,F}(x^{k_j}) = 0\,. \tag{44}$$

Let us show that $\lambda_{min,F}(\tilde{x}) = 0$. Let us assume by contradiction that we have $\lambda_{min,F}(\tilde{x}) > 0$ and that j_0 exists such that for any $j > j_0$ we have that $\lambda_{min,F}(x^{k_j}) > 0$ with $x^{k_j} \in \, int\, B$. Moreover the following equation holds:

$$F(x^{k_j} + t h^{k_j}) \leq F(x^{k_j}) + t\beta J_F(x^{k_j}) h^{k_j}, \; x^{k_j} + t\, h^{k_j} \in \, int\, B, \; \forall t \in [0, \tilde{t}\,), \tag{45}$$

where $\tilde{t}$ is a positive constant sufficiently small. Now we prove that for $j > j_0$ we have:

$$2t_{k_j} > \min\{\,1\,,\,\tilde{t}\,\}. \tag{46}$$

Let us remember the backtracking procedure described in steps (1a), (2a). If $t_{k_j} = 1$, formula (46) holds. If $t_{k_j} < 1$, then $2t_{k_j}$ is not an acceptable step-size since condition (2a) is not satisfied or $x^{k_j} + 2t_{k_j} h^{k_j}$ is not a feasible point. Hence equation (46) must hold. The assumption $\lambda_{min,F}(\tilde{x}) > 0$ and equation (44) imply the following equation:

$$\lim_{j \to +\infty} 2t_{k_j} = 0. \tag{47}$$

Equation (47) contradicts (46), hence we have $\lambda_{min,F}(\tilde{x}) = 0$.
We conclude noting that $\tilde{x}$ is not a local Pareto maximal point since it is the accumulation point of a sequence of points where the vector function F is decreasing that is $\tilde{x} \in V^*_{min}$.

We conclude this section investigating the accumulation points of the sequences generated by algorithm A_1 when we choose several initial points x^0 in the interior of the box B.
Let $\tilde{x} \in V^*_{min}$ we consider the following multi-objective optimization problem:

$$\min_{h \in \mathbf{R}^n} g_{\tilde{x}}(h) \tag{48}$$

where the function $g_{\tilde{x}}$ is given in (25). We look for the Pareto optimal solutions of problem (48) using the path following method proposed in [3]. We discard the point $\tilde{x}$ if there exists h such that $g_{\tilde{x}}(h) < 0$ and

$$J_F(\tilde{x}) D_{\tilde{x}}(h) h \leq 0. \tag{49}$$

In the next section we call algorithm A_2 the procedure that solves problem (48) and that computes $g_{\tilde{x}}(h) < 0$ for each h solution of problem (48).

Lemma 7. *Let $F_i : E \subseteq \mathbf{R}^n \to \mathbf{R}$, $i = 1, 2, \ldots, s$ be twice continuously differentiable functions in E, $B \subset E$, let $J_{g_{\tilde{x}}}(h)$ be the Jacobian matrix of $g_{\tilde{x}}$ at the point h, then a necessary condition for h^* to be a local Pareto optimal solution of problem (48) is that the matrix $J_{g_{\tilde{x}}}(h) J_{g_{\tilde{x}}}(h)^T$ is singular at $h = h^*$.*

Proof. It follows using Lemma 4 in [3].

We conclude this section describing the basic steps of algorithm A_2. Let $\mathcal{A}$ be the set of the accumulation points determined by algorithm A_1, let $[-1,1]^n \subset \mathbf{R}^n$ be the cartesian product of n copies of the interval $[-1,1]$, let $M \in \mathbf{R}^{s \times s}$ be a matrix and let $det(M) \in \mathbf{R}$ and $adj(M) \in \mathbf{R}^{s \times s}$ denote the determinant of M and the adjoint matrix of M respectively. Algorithm A_2 is defined by the following steps:

Algorithm A_2

Step 0 Select a point $\tilde{x} \in \mathcal{A}$;
Step 1. Choose an initial direction $h^0 \in int\,[-1,1]^n$, a positive constant ϵ, $0 < \epsilon << 1$, a positive integer it_{max} and set $k = 0$.
Step 2. If $det\left(J_{g_{\tilde{x}}}(h^k) J_{g_{\tilde{x}}}(h^k)^T\right) < \epsilon$ or $k > it_{max}$ go to Step 5. Otherwise compute:

$$p^k = -J_{g_{\tilde{x}}}(h^k)^T\, adj\left(J_{g_{\tilde{x}}}(h^k) J_{g_{\tilde{x}}}(h^k)^T\right) e\,, \tag{50}$$

and proceed with Step 3.
Step 3. Compute $h^{k+1} = h^k + s_k p^k$, where s_k is the step-size computed with a standard backtracking procedure:
(1a) Set $s = 0.5$
(2a) if $-e < h^k + sp^k < e$ then $s_k = s$ otherwise $s = s/2$ (end of backtracking)
Step 4. Set $k = k + 1$ and go to Step 2.
Step 5. If $g_{\tilde{x}}(h^k) < 0$ discard $\tilde{x}$, otherwise accept $\tilde{x}$. Go to *Step* 0.

Note that we choose the step-size s_k in such a way the direction h^k belongs to $[-1,1]^n$. This is not a restrictive condition since the norm of the vector h^k is irrelevant for algorithm A_2. In fact algorithm A_2 exploits the existence of directions h where $g_{\tilde{x}}(h) < 0$.

4 Numerical Experiments

In this section we validate the numerical methods (i.e.: algorithm A_1 and algorithm A_2) proposed in Section 3 on four bi-criteria optimization problems. The numerical methods have been implemented in Matlab on a Pentium M 1.6GHz in double precision arithmetic. The computational time has been measured using the “cputime” Matlab function. For each test problem we consider N_{tot} sequences starting from N_{tot} initial guesses $x_{0,i} \in int\,B$, $i = 1, 2, \ldots, N_{tot}$. In particular, the starting points $x_{0,i}$, $i = 1, 2 \ldots, N_{tot}$ are chosen equally spatially distributed or randomly uniformly distributed on the box B. We note that algorithm A_1 is well suited for parallel computing since we can compute each of the N_{tot} sequences independently from the others.
Once determined the accumulation points of the N_{tot} sequences we apply algorithm A_2 to establish if some of the accumulation points determined by algorithm A_1 must be discarded.
We consider four test problems, the first one and the third one belong to a class of two-objective optimization problems proposed by K. Deb in [6]. The tests considered have an increasing degree of difficulty.

Test 1) We have two objective functions (i.e.: $s = 2$) and two spatial variables (i.e.: $n = 2$). The objective functions are given by:

$$f_1(x) = x_1 \qquad\qquad f_2(x) = \psi(x_2)/x_1\ , \tag{51}$$

where

$$\psi(x_2) = 2 - 0.8\ e^{-\left(\frac{x_2-0.6}{0.4}\right)^2} -\ e^{-\left(\frac{x_2-0.2}{0.04}\right)^2}, \tag{52}$$

and the box constraint is given by:

$$B = \{\ \ x = (x_1, x_2)^T \in \mathbf{R}^2, \quad |\ \ 0.1 \le x_1 \le 1, \quad 0 \le x_2 \le 1\ \ \}. \tag{53}$$

The function ψ has a global minimizer at $x_2 \approx 0.2$ and has a local minimizer at $x_2 \approx 0.6$. The global minimizer of the function ψ has a narrow attraction region when compared with the attraction region of its local minimizer. This feature makes it a very interesting test problem.

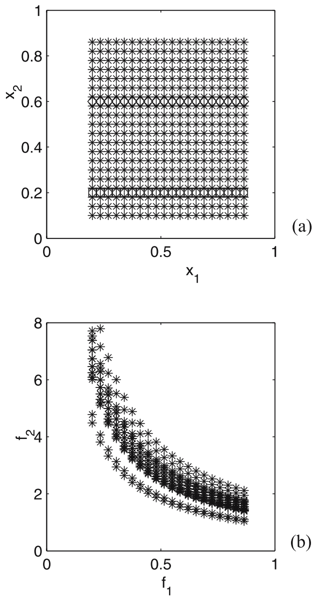

Fig. 1. Starting points (a); Objective functions values (b)

We have a convex global Pareto optimal front (see Figure 2(a)) corresponding to the global minimizer of the function ψ, that is the set given by:

$$Global : \mathcal{P}_g = \{(x_1, x_2)^T \in \mathbf{R}^2\ |\, x_2 \approx 0.2, \quad 0.1 \le x_1 \le 1\}\,, \tag{54}$$

and we have a convex local Pareto optimal front corresponding to the local minimizer of the function ψ, that is the set given by (see Figure 2(a)):

$$Local : \mathcal{P}_l = \{(x_1, x_2)^T \in \mathbf{R}^2\ \ |\ x_2 \approx 0.6,\ \ 0.1 \le x_1 \le 1\}\,. \tag{55}$$

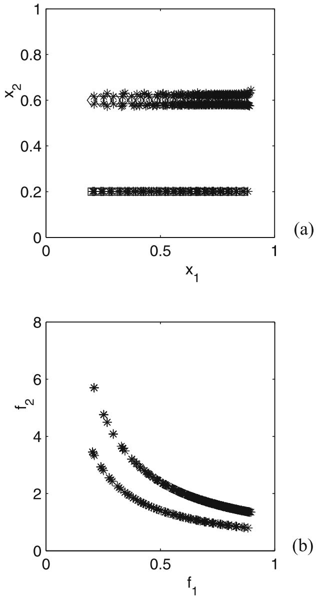

Fig. 2. Accumulation points (a); Objective functions values (b)

Figure 2(a) shows the numerical approximations of the local and global Pareto fronts determined by algorithm A_1 starting from $N_{tot} = 400$ points distributed in the interior of the box B as shown in Figure 1(a). Figures 1(b) and 2(b) show the value of the objective functions in the f_1-f_2 plane. The computational time required to compute the accumulation points of $N_{tot} = 400$ trajectories fulfilling the stop criterion described in Step 2 of algorithm A_1 when $\epsilon = 5 \cdot 10^{-7}$ and $it_{max} = 500$ is of about 36.96 seconds. We do not apply algorithm A_2 to this test problem since the set of points generated by algorithm A_1 consists only of the local Pareto optimal points.

Test 2) We consider two objective functions (i.e.: $s = 2$) and two spatial variables (i.e.: $n = 2$). The objective functions are given by:

$$f_1(x) = x_1^3 \qquad\qquad f_2(x) = (x_2 - x_1)^3, \tag{56}$$

and the box is given by:

$$B = \{ \ x = (x_1, x_2)^T \in \mathbf{R}^2 \ \ | \ -1 \leq x_1 \leq 1, \quad -1 \leq x_2 \leq 1 \ \}. \tag{57}$$

This test problem is very difficult to deal with since there are two sets of points where the gradient vectors of the objective functions are identically null but these points do not belong to the local Pareto optimal front. In fact we have the local and the global Pareto fronts given by:

$$Global : \mathcal{P}_g = \{(-1, 1)^T\}\,, \tag{58}$$

$$Local : \mathcal{P}_l = \{(-1, -1)^T\}\,, \tag{59}$$

and we have the following two sets of points that belong to V_{min} but do not belong to the Pareto local front:

$$\mathcal{B} = \{(x_1, x_2)^T \in B \mid x_1 = x_2,\ x_1 \neq -1\}\,, \tag{60}$$

$$\mathcal{Z} = \{(x_1, x_2)^T \in B \mid x_1 = 0\}\,. \tag{61}$$

Figure 3 and Figure 4 show the numerical results obtained applying algorithm A_1 to solve *Test 2*. Figure 3 shows the starting points ($N_{tot} = 100$) and Figure 4 shows the accumulation points of the sequence defined by algorithm A_1. Note that Figures 3(a) and 4(a) show the points in the x_1-x_2 plane and Figure 3(b) and 4(b) show the values of the objective functions in the plane f_1-f_2. The computational time required to compute the accumulation points of $N_{tot} = 100$ trajectories fulfilling the stop criterion described in Step 2 of algorithm A_1 when $\epsilon = 5 \cdot 10^{-7}$ and $it_{max} = 500$ is of about 16 seconds.

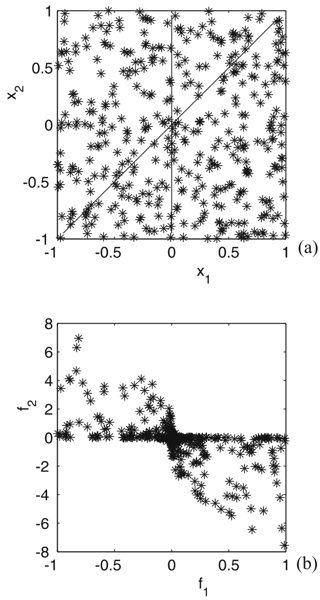

Fig. 3. Starting points (a); Objective functions values (b)

Finally we apply algorithm A_2 at each point determined by algorithm A_1 (see Figure 4(a)) in order to establish the points to discard. Algorithm A_2 shows that only three points are possible Pareto optimal points. Note that two points are local optimal Pareto points while the point $x_1 = 0$, $x_2 = 0$ is not

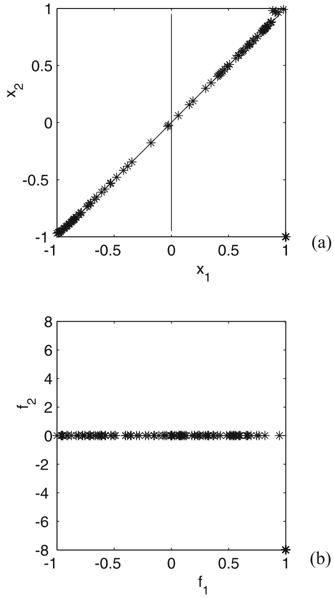

Fig. 4. Accumulation points (a); Objective functions values (b)

a local optimal Pareto point. Indeed algorithm A_2 is not able to discard this point since at this point all the gradient vectors and the Hessian matrices are identically null. Figure 5 shows the result of the selection due to algorithm A_2.

Test 3) We have $s = 2$, $n = 2$ and the following objective functions:

$$f_1(x) = v_1(x_1) \qquad\qquad f_2(x) = \psi(x_2) r(x_1, x_2),$$

$$r(x_1, x_2) = 1 - \left(\frac{v_1(x_1)}{\psi(x_2)}\right)^{\alpha} - \frac{v_1(x_1)}{\psi(x_2)} \sin(2\pi q v_1(x_1)),$$

where we choose $v_1(x_1) = x_1$, $\psi(x_2) = 1 + 10x_2$, $\alpha = 2$, $q = 4$. The box constraints are given by:

$$B = \{ \ x = (x_1, x_2)^T \in \mathbf{R}^2, \quad | \ 0 \leq x_1 \leq 1, \quad 0 \leq x_2 \leq 1 \ \}.$$

The Pareto-optimal front is not a connected set. Figure 6 shows the part of the graph $r(x_1, 0)$ versus x_1, $x_1 \in [0, 1]$ where $r(x_1, 0)$ is a non-increasing function of x_1. This figure shows also the Pareto front. In fact a point $(x_1, 0)$ belongs to the global Pareto optimal front when the point $(x_1, r(x_1, 0))$ belongs to the dashed line shown in Figure 6 and a point $(x_1, 0)$ is a local (non global) Pareto optimal point front when the point $(x_1, r(x_1, 0))$ belongs to the solid

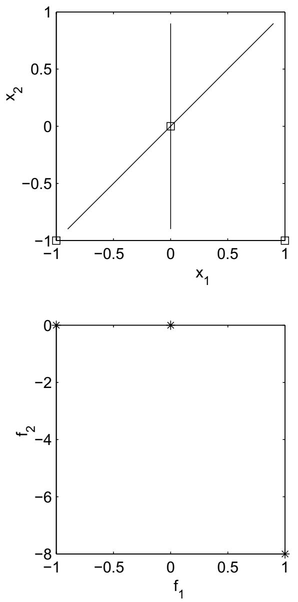

Fig. 5. Results of the selection due to algorithm A_2

line shown in Figure 6. Furthermore note that the points $(0, x_2)$, $x_2 \in [0, 1]$ are global minimizers of the function F_1, but only the point $(0, 0)$ belongs to the global Pareto-optimal front. Figure 7 shows the numerical results obtained

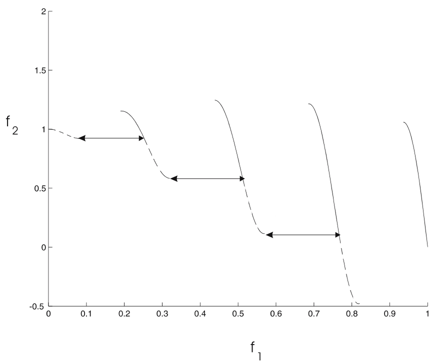

Fig. 6. $r(x_1, 0)$ versus x_1

applying algorithm A_1 to solve *Test 3*. As in the previous numerical experiments we choose the starting points ($N_{tot} = 1600$) equally distributed on a rectangular grid of B. Figure 7 shows the accumulation points of the sequence defined by algorithm A_1. As in the previous experiments Figure 7(a) shows the points in the x_1-x_2 plane and Figure 7(b) shows the values of the objective functions in the plane f_1-f_2. The computational time required to compute the accumulation points of $N_{tot} = 6400$ trajectories fulfilling the stop criterion described in Step 2 of algorithm A_1 when $\epsilon = 5 \cdot 10^{-7}$ and $it_{max} = 250$ is of about 334.88 seconds.

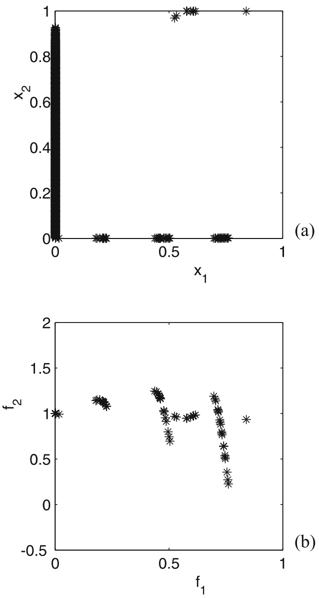

Fig. 7. Accumulation points (a); Objective functions values (b)

We apply algorithm A_2 starting from each point determined by algorithm A_1 (see Figure 7(a)) in order to establish the points to discard. The points selected by algorithm A_2 are shown in Figure 8. Similar results have been obtained for several choices of the parameters α and q.

Test 4) We have $s = 2$, $n = 100$ and the following objective functions (see [8] p.442-443):

$$f_1(x) = \left(\frac{1}{n}\sum_{i=1}^{n}\left[x_i^2 - 10\cos(2\pi\, x_i) + 10\right]\right)^{1/4}$$
$$f_2(x) = \left(\frac{1}{n}\sum_{i=1}^{n}\left[(x_i - 1.5)^2 - 10\cos(2\pi(x_i - 1.5)) + 10\right]\right)^{1/4}.$$

The box constraints are given by:

$$B = \{\ x = (x_1, x_2, \ldots, x_n)^T \in \mathbf{R}^n, \quad | \quad -5 \le x_i \le 5,\ i = 1, 2, \ldots, n\ \}.$$

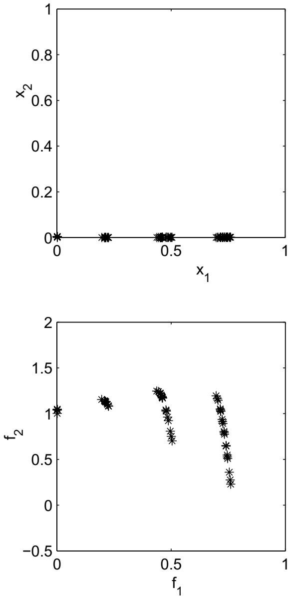

Fig. 8. Results of the selection due to algorithm A_2

The Pareto front determined by several kinds of genetic algorithms and with multi-start tabu search can be found in [8] Figure 13 pag. 443. Figure 9 shows the values of f_1 and f_2 on the Pareto front achieved by algorithm A_1. The computational time required to compute the accumulation points of $N_{tot} = 1000$ trajectories fulfilling the stop criterion described in Step 2 of algorithm A_1 when $\epsilon = 5 \cdot 10^{-12}$ and $it_{max} = 1000$ is of about 371.08 seconds.

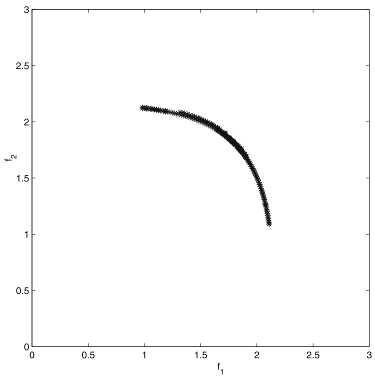

Fig. 9. Objective functions valued at the accumulation points determined by algorithm A_1

5 Conclusions

We conclude noting that algorithm A_1 has a satisfactory behavior when applied to solve multi-objective optimization problems. In fact in almost all the numerical experiments several trajectories converge to the local Pareto front. Moreover starting from a sufficiently large number of points algorithm A_1 is able to approximate the entire Pareto front. Finally when algorithm A_1 does not work well we can use algorithm A_2 to discard the points determined by the algorithm A_1 where it is possible to decrease all the objective functions. The computational cost of algorithm A_1 is essentially due to the computation of the spectral decomposition of the matrix $M_F(x^k)$ given in (31) made at each iteration k. Obviously the computational cost of algorithm A_1 increases heavily when the number s of the objective functions increases so that some techniques to avoid the computation of the spectral decomposition at each iteration should be studied. Probably something like the rank one corrections could be used when the number of the iterations becomes sufficiently large. However algorithm A_1 is well suited for parallel implementation since each trajectory can be computed independently from the others so that using a parallel machine we can reduce the computational cost required to follow several trajectories. The computational cost of algorithm A_2, when $\tilde{x} \in \mathcal{A}$ is fixed, is essentially imputable to the computation of the adjoint matrix $adj\left(J_{g_{\tilde{x}}}(h^k)J_{g_{\tilde{x}}}(h^k)^T\right)$ made at each iteration k. Hence algorithm A_2 can be executed with affordable computing resources when the number of the objective functions is sufficiently small, that is s is a few hundreds, and the number of points belonging to $\mathcal{A}$ to exploit is a few thousands. Future work is needed to generate a computational method that fuses the basic features of algorithm A_1 and algorithm A_2. In particular preliminary attempts should be made to avoid the use of algorithm A_2 on the entire set $\mathcal{A}$.

References

1. Fliege J, Svaiter BF (2000) Steepest descent methods for multicriteria optimization. Mathematical Methods of Operations Research 51:479-494
2. Grana Drummond LM, Svaiter BF (2005) A steepest descent method for vector omptimazion. Journal of Computational and Applied Mathematics 175:395-414
3. Recchioni MC (2003) A path following method for box-constrained multiobjective optimization with applications to goal programming problems. Mathematical Methods of Operations Research 58:69-85
4. Bonne H, Iusem AN, Svaiter BF (2005) Proximal methods in vector optimization. SIAM Journal on Optimization 4:953-970
5. Miglierina E (2004) Slow solutions of a differential inclusion and vector optimization. Set Valued Analysis 12:345-356
6. Deb K (1999) Multi-objective genetic algorithms: Problem difficulties and construction of test problems. Evolutionary Computation Journal 7:205-230
7. Deb K (2001) Nonlinear goal programming using multi-objective genetic algorithms. Journal of the Operational Research Society 52:291-302

8. Beausoleil RP (2006) “MOSS” mutiobjective scatter search applied to non-linear multiple criteria optimization. European Journal of Operational Research 169:426-449
9. Herzel S, Recchioni MC, Zirilli F (1991) A quadratically convergent method for linear programming. Linear Algebra and its Applications 152:255-289
10. Pacelli G, Recchioni MC (2000) Monotone variable - metric algorithm for linearly constrained nonlinear programming. Journal of Optimization Theory and Applications 104:255-279
11. Pacelli G, Recchioni MC (2001) An interior point algorithm for global optimal solutions and KKT points. Optimization Methods and Software 15:225-256
12. Recchioni MC, Scoccia A (2000) A stochastic algorithm for constrained global optimization. Journal of Global Optimization 16:257-270
13. Geoffrin AM (1968) Proper efficiency and the theory of vector maximization. Journal of Mathematical Analysis and Applications 22:618-630
14. Das I, Dennis JE (1997) A closer look at drawbacks of minimizing weighted sums of objectives for Pareto set generation in multi-criteria optimization problems. Structural Optimization 14:63-69
15. Das I, Dennis JE (1998) Normal-boundary intersection: a new method for generating Pareto optimal points in nonlinear multicriteria optimization problems. Siam Journal on Optimization 8:631-657
16. Tind J, Wiecek MM (1999) Augmented Lagrangian and Tchebycheff approaches in multiple objective programming. Journal of Global Optimization 14:251-266
17. Schandl B, Klamroth K, Wiecek MM (2002) Introducing oblique norms into multiple criteria programming. Journal of Global Optimization 23:81-97
18. Hanne T (2000) Global multiobjective optimization using evolutionary algorithms. Journal of Heuristics 6:347-360
19. Hillermeier C (2001) Generalized homotopy approach to multiobjective optimization. Journal of Optimization Theory and Applications 110:557-583
20. Zitzler E, Deb K, Thiele L, Coello Coello CA, Corne D Editors (2001) First International Conference on Evolutionary Multi-Criterion Optmization. Lecture Notes in Computer Science Vol. 1993, Springer, Berlin
21. Cantu-Paz E, Goos G, Hartmanis J, van Leeeuwen J (2003) Genetic and Evolutionary Computation - GECCO 2003. Lecture Notes in Computer Science Vol. 2723 and 2724, Springer, Berlin
22. Golub GH, Van Loan CF (1989) Matrix computation. The Johns Hopkins University Press, Baltimore Maryland
23. Luc DT (1988) Theory of vector optimization. Springer-Verlag, Berlin
24. Sawaragi Y, Nakayama H, Tanino T (1985) Theory of multiobjective optimization. Mathematics in Science and Engineering 176, Academic Press, Orlando FL
25. Giorgi G, Guerraggio A (1998) The notion of invexity in vector optimization: smooth and nonsmooth case. In: Crouzeix JP, Martinez Legaz JE and Volle M (eds.) Nonconvex optimization and its applications 27. Kluwer Academic Publishers, Dordrecht, 389-405
26. Aghezzaf B, Hachimi M (2000) Generalized invexity and duality in multiobjective programming problems. Journal of Global Optimization 18:91-101

Lecture Notes in Economics and Mathematical Systems

For information about Vols. 1–495
please contact your bookseller or Springer-Verlag

Vol. 496: P. Weiß, Unemployment in Open Economies. XII, 226 pages. 2001.

Vol. 497: J. Inkmann, Conditional Moment Estimation of Nonlinear Equation Systems. VIII, 214 pages. 2001.

Vol. 498: M. Reutter, A Macroeconomic Model of West German Unemployment. X, 125 pages. 2001.

Vol. 499: A. Casajus, Focal Points in Framed Games. XI, 131 pages. 2001.

Vol. 500: F. Nardini, Technical Progress and Economic Growth. XVII, 191 pages. 2001.

Vol. 501: M. Fleischmann, Quantitative Models for Reverse Logistics. XI, 181 pages. 2001.

Vol. 502: N. Hadjisavvas, J. E. Martínez-Legaz, J.-P. Penot (Eds.), Generalized Convexity and Generalized Monotonicity. IX, 410 pages. 2001.

Vol. 503: A. Kirman, J.-B. Zimmermann (Eds.), Economics with Heterogenous Interacting Agents. VII, 343 pages. 2001.

Vol. 504: P.-Y. Moix (Ed.), The Measurement of Market Risk. XI, 272 pages. 2001.

Vol. 505: S. Voß, J. R. Daduna (Eds.), Computer-Aided Scheduling of Public Transport. XI, 466 pages. 2001.

Vol. 506: B. P. Kellerhals, Financial Pricing Models in Continuous Time and Kalman Filtering. XIV, 247 pages. 2001.

Vol. 507: M. Koksalan, S. Zionts, Multiple Criteria Decision Making in the New Millenium. XII, 481 pages. 2001.

Vol. 508: K. Neumann, C. Schwindt, J. Zimmermann, Project Scheduling with Time Windows and Scarce Resources. XI, 335 pages. 2002.

Vol. 509: D. Hornung, Investment, R&D, and Long-Run Growth. XVI, 194 pages. 2002.

Vol. 510: A. S. Tangian, Constructing and Applying Objective Functions. XII, 582 pages. 2002.

Vol. 511: M. Külpmann, Stock Market Overreaction and Fundamental Valuation. IX, 198 pages. 2002.

Vol. 512: W.-B. Zhang, An Economic Theory of Cities.XI, 220 pages. 2002.

Vol. 513: K. Marti, Stochastic Optimization Techniques. VIII, 364 pages. 2002.

Vol. 514: S. Wang, Y. Xia, Portfolio and Asset Pricing. XII, 200 pages. 2002.

Vol. 515: G. Heisig, Planning Stability in Material Requirements Planning System. XII, 264 pages. 2002.

Vol. 516: B. Schmid, Pricing Credit Linked Financial Instruments. X, 246 pages. 2002.

Vol. 517: H. I. Meinhardt, Cooperative Decision Making in Common Pool Situations. VIII, 205 pages. 2002.

Vol. 518: S. Napel, Bilateral Bargaining. VIII, 188 pages. 2002.

Vol. 519: A. Klose, G. Speranza, L. N. Van Wassenhove (Eds.), Quantitative Approaches to Distribution Logistics and Supply Chain Management. XIII, 421 pages. 2002.

Vol. 520: B. Glaser, Efficiency versus Sustainability in Dynamic Decision Making. IX, 252 pages. 2002.

Vol. 521: R. Cowan, N. Jonard (Eds.), Heterogenous Agents, Interactions and Economic Performance. XIV, 339 pages. 2003.

Vol. 522: C. Neff, Corporate Finance, Innovation, and Strategic Competition. IX, 218 pages. 2003.

Vol. 523: W.-B. Zhang, A Theory of Interregional Dynamics. XI, 231 pages. 2003.

Vol. 524: M. Frölich, Programme Evaluation and Treatment Choise. VIII, 191 pages. 2003.

Vol. 525: S. Spinler, Capacity Reservation for Capital- Intensive Technologies. XVI, 139 pages. 2003.

Vol. 526: C. F. Daganzo, A Theory of Supply Chains. VIII, 123 pages. 2003.

Vol. 527: C. E. Metz, Information Dissemination in Currency Crises. XI, 231 pages. 2003.

Vol. 528: R. Stolletz, Performance Analysis and Optimization of Inbound Call Centers. X, 219 pages. 2003.

Vol. 529: W. Krabs, S. W. Pickl, Analysis, Controllability and Optimization of Time-Discrete Systems and Dynamical Games. XII, 187 pages. 2003.

Vol. 530: R. Wapler, Unemployment, Market Structure and Growth. XXVII, 207 pages. 2003.

Vol. 531: M. Gallegati, A. Kirman, M. Marsili (Eds.), The Complex Dynamics of Economic Interaction. XV, 402 pages, 2004.

Vol. 532: K. Marti, Y. Ermoliev, G. Pflug (Eds.), Dynamic Stochastic Optimization. VIII, 336 pages. 2004.

Vol. 533: G. Dudek, Collaborative Planning in Supply Chains. X, 234 pages. 2004.

Vol. 534: M. Runkel, Environmental and Resource Policy for Consumer Durables. X, 197 pages. 2004.

Vol. 535: X. Gandibleux, M. Sevaux, K. Sörensen, V. T'kindt (Eds.), Metaheuristics for Multiobjective Optimisation. IX, 249 pages. 2004.

Vol. 536: R. Brüggemann, Model Reduction Methods for Vector Autoregressive Processes. X, 218 pages. 2004.

Vol. 537: A. Esser, Pricing in (In)Complete Markets. XI, 122 pages, 2004.

Vol. 538: S. Kokot, The Econometrics of Sequential Trade Models. XI, 193 pages. 2004.

Vol. 539: N. Hautsch, Modelling Irregularly Spaced Financial Data. XII, 291 pages. 2004.

Vol. 540: H. Kraft, Optimal Portfolios with Stochastic Interest Rates and Defaultable Assets. X, 173 pages. 2004.

Vol. 541: G.-y. Chen, X. Huang, X. Yang, Vector Optimization. X, 306 pages. 2005.

Vol. 542: J. Lingens, Union Wage Bargaining and Economic Growth. XIII, 199 pages. 2004.

Vol. 543: C. Benkert, Default Risk in Bond and Credit Derivatives Markets. IX, 135 pages. 2004.

Vol. 544: B. Fleischmann, A. Klose, Distribution Logistics. X, 284 pages. 2004.

Vol. 545: R. Hafner, Stochastic Implied Volatility. XI, 229 pages. 2004.

Vol. 546: D. Quadt, Lot-Sizing and Scheduling for Flexible Flow Lines. XVIII, 227 pages. 2004.

Vol. 547: M. Wildi, Signal Extraction. XI, 279 pages. 2005.

Vol. 548: D. Kuhn, Generalized Bounds for Convex Multistage Stochastic Programs. XI, 190 pages. 2005.

Vol. 549: G. N. Krieg, Kanban-Controlled Manufacturing Systems. IX, 236 pages. 2005.

Vol. 550: T. Lux, S. Reitz, E. Samanidou, Nonlinear Dynamics and Heterogeneous Interacting Agents. XIII, 327 pages. 2005.

Vol. 551: J. Leskow, M. Puchet Anyul, L. F. Punzo, New Tools of Economic Dynamics. XIX, 392 pages. 2005.

Vol. 552: C. Suerie, Time Continuity in Discrete Time Models. XVIII, 229 pages. 2005.

Vol. 553: B. Mönch, Strategic Trading in Illiquid Markets. XIII, 116 pages. 2005.

Vol. 554: R. Foellmi, Consumption Structure and Macroeconomics. IX, 152 pages. 2005.

Vol. 555: J. Wenzelburger, Learning in Economic Systems with Expectations Feedback (planned) 2005.

Vol. 556: R. Branzei, D. Dimitrov, S. Tijs, Models in Cooperative Game Theory. VIII, 135 pages. 2005.

Vol. 557: S. Barbaro, Equity and Efficiency Considerations of Public Higer Education. XII, 128 pages. 2005.

Vol. 558: M. Faliva, M. G. Zoia, Topics in Dynamic Model Analysis. X, 144 pages. 2005.

Vol. 559: M. Schulmerich, Real Options Valuation. XVI, 357 pages. 2005.

Vol. 560: A. von Schemde, Index and Stability in Bimatrix Games. X, 151 pages. 2005.

Vol. 561: H. Bobzin, Principles of Network Economics. XX, 390 pages. 2006.

Vol. 562: T. Langenberg, Standardization and Expectations. IX, 132 pages. 2006.

Vol. 563: A. Seeger (Ed.), Recent Advances in Optimization. XI, 455 pages. 2006.

Vol. 564: P. Mathieu, B. Beaufils, O. Brandouy (Eds.), Artificial Economics. XIII, 237 pages. 2005.

Vol. 565: W. Lemke, Term Structure Modeling and Estimation in a State Space Framework. IX, 224 pages. 2006.

Vol. 566: M. Genser, A Structural Framework for the Pricing of Corporate Securities. XIX, 176 pages. 2006.

Vol. 567: A. Namatame, T. Kaizouji, Y. Aruga (Eds.), The Complex Networks of Economic Interactions. XI, 343 pages. 2006.

Vol. 568: M. Caliendo, Microeconometric Evaluation of Labour Market Policies. XVII, 258 pages. 2006.

Vol. 569: L. Neubecker, Strategic Competition in Oligopolies with Fluctuating Demand. IX, 233 pages. 2006.

Vol. 570: J. Woo, The Political Economy of Fiscal Policy. X, 169 pages. 2006.

Vol. 571: T. Herwig, Market-Conform Valuation of Options. VIII, 104 pages. 2006.

Vol. 572: M. F. Jäkel, Pensionomics. XII, 316 pages. 2006

Vol. 573: J. Emami Namini, International Trade and Multinational Activity, X, 159 pages, 2006.

Vol. 574: R. Kleber, Dynamic Inventory Management in Reverse Logisnes, XII, 181 pages, 2006.

Vol. 575: R. Hellermann, Capacity Options for Revenue Management, XV, 199 pages, 2006.

Vol. 576: J. Zajac, Economics Dynamics, Information and Equilibnum, X, 284 pages, 2006.

Vol. 577: K. Rudolph, Bargaining Power Effects in Financial Contracting, XVIII, 330 pages, 2006.

Vol. 578: J. Kühn, Optimal Risk-Return Trade-Offs of Commercial Banks, IX, 149 pages, 2006.

Vol. 579: D. Sondermann, Introduction to Stochastic Calculus for Finance, X, 136 pages, 2006.

Vol. 580: S. Seifert, Posted Price Offers in Internet Auction Markets, IX, 186 pages, 2006.

Vol. 581: K. Marti; Y. Ermoliev; M. Makowsk; G. Pflug (Eds.), Coping with Uncertainty, XIII, 330 pages, 2006.

Vol. 582: J. Andritzky, Sovereign Default Risks Valuation: Implications of Debt Crises and Bond Restructurings. VIII, 251 pages, 2006.

Vol. 583: I.V. Konnov, D.T. Luc, A.M. Rubinov (Eds.), Generalized Convexity and Related Topics, IX, 469 pages, 2006.

Vol. 584: C. Bruun, Adances in Artificial Economics: The Economy as a Complex Dynamic System. XVI, 296 pages, 2006.

Vol. 585: R. Pope, J. Leitner, U. Leopold-Wildburger, The Knowledge Ahead Approach to Risk, XVI, 218 pages, 2007 (planned).

Vol. 586: B. Lebreton, Strategic Closed-Loop Supply Chain Management. X, 150 pages, 2007 (planned).