Journal of Global Optimization **7**: 229–259, 1995.

Partly Convex Programming and Zermelo's Navigation Problems*

S. ZLOBEC
McGill University, Department of Mathematics and Statistics, Burnside Hall, 805 Sherbrooke Street West, Montreal, Quebec, Canada H3A 2K6.

(Received: 14 February 1994; accepted: 31 March 1995)

Abstract. Mathematical programs, that become convex programs after "freezing" some variables, are termed partly convex. For such programs we give saddle-point conditions that are both necessary and sufficient that a feasible point be globally optimal. The conditions require "cooperation" of the feasible point tested for optimality, an assumption implied by lower semicontinuity of the feasible set mapping. The characterizations are simplified if certain point-to-set mappings satisfy a "sandwich condition".
The tools of parametric optimization and basic point-to-set topology are used in formulating both optimality conditions and numerical methods. In particular, we solve a large class of Zermelo's navigation problems and establish global optimality of the numerical solutions.

Mathematics Subject Classifications (1991). 90C25, 90C30, 90C90.

Key words: Global optimum, local optimum, saddle point, point-to-set mapping, Zermelo's navigation problem.

1. Introduction

Consider the mathematical program

$$(P) \qquad \underset{\text{s.t.}}{\text{Min}} \quad f(z)$$
$$f^i(z) \leqslant 0, \quad i \in \mathcal{P} = \{1, \ldots, m\},$$

where $f, f^i : R^N \to R, i \in \mathcal{P}$ are continuous functions. After splitting the variable $z \in R^N$ into $z = (x, \theta)$, where $x \in R^n$, $\theta \in R^p$, $n + p = N$, the program (P) is rewritten as

$$(P, \theta) \qquad \underset{\text{s.t.}}{\text{Min}} \quad f(x, \theta)$$
$$f^i(x, \theta) \leqslant 0, \quad i \in \mathcal{P}.$$

If, for every θ, the functions $f(\cdot, \theta)$, $f^i(\cdot, \theta) : R^n \to R$, $i \in \mathcal{P}$ are convex, then (P) is termed a *partly convex program* (abbreviated: PCP) relative to the splitting $z = (x, \theta)$.

* Research partly supported by NSERC of Canada.

Many real life problems can be formulated as PCPs. They include the classical navigation problems of Zermelo [23, also see Section 8 below], optimal design of multi-stage heat exchangers in chemical engineering (see [2, 28]), the problem of how to increase capacities of machines to obtain a higher profit in a textile mill (see [5, 28]), pooling and blending in oil refinery (see [7]) and problems from chemical physics to determine a configuration of clusters of atoms and molecules that minimizes the Lennard–Jones interaction potential (see [12]).

Note that every convex program is a PCP relative to an arbitrary splitting. What makes a general PCP essentially different from convex programs is nonconvexity of the feasible set and the fact that local and global optima may not coincide. Moreover (as in Zermelo's problems) the feasible set of a PCP may consist of disjoint subsets. The main objective of this paper is to study and characterise *global* and *local* optimality of a feasible point in PCP and apply the results to the Zermelo navigation problems. Let us recall that the first and second order optimality conditions from the literature provide only partial results, i.e., either necessary or sufficient conditions for local optimality but not both. Practical results on characterizing global optimality are virtually nonexistent. (See [4, 9].)

In order to formulate optimality conditions we associate, with every feasible point, a region that permits testing of global optimality by a saddle-point condition. The size of that region is a measure of "cooperation" of the feasible point. Global optimality is characterized in Section 3. Points that are local (but not global) optima are identified as nonoptimal. The characterization is simplified when three point-to-set mappings satisfy a "sandwich condition", introduced and studied in Section 4. An application to convex programming is given in Section 5. An exact penalty function for a PCP is given in Section 6. The global optima of the exact penalty function and of the PCP coincide. Optimal solutions of PCP can also be calculated by methods adapted from input optimization. We formulate such methods in Section 7 and use them in Section 8 to solve classic navigation problems of Zermelo. To characterize local optimality of $z^* = (x^*, \theta^*)$ we strengthen the local cooperation requirement by assuming lower semicontinuity of the feasible set mapping and uniqueness of the optimal solution x^* of the convex program (P, θ^*), in Section 9. Finally, the results are applied to partly linear programs with equality constraints, in Section 10.

The classification of feasible points by their "cooperation" (Definition 2.1), the sandwich condition (Definition 4.1), and the characterizations of global optimality (Theorems 3.1, 4.4 and 6.1) appear to be new. The result on local optima (Theorem 9.1) was recently given in (not readily available) conference proceedings [28] (see also [27]) and it is included here for the sake of completeness. The numerical methods for solving PCP are presented here in their primitive forms primarily as a means for calculating points to be checked for local and global optimality. They appear to work well for Zermelo's problems. More general and sophisticated formulations are currently being studied in [18] using the recent results from [19] and [26].

The idea to use point-to-set mappings in a study of global optimality is not new. Indeed, it is well known (see, e.g., [21, 22]) that a local minimum of a function is also global, if the feasible set mapping, relative to the right-hand side perturbations, is lower semi-continuous. The importance of lower semicontinuity in characterizations of local optima in parametric optimization has been demonstrated in, e.g., [25]. This paper reconfirms that, in order to characterize global and local optima in nonconvex optimization, in addition to linear algebra and calculus, one also needs some basic tools from point-to-set topology.

2. Cooperation of Feasible Points

We will study optimality of a feasible point

$$z \in Z = \{z \in R^N : f^i(z) \leqslant 0, \;\; i \in \mathcal{P}\}$$

using the notion of "cooperative" feasible points. With each feasible point we will associate a region of the feasible set Z where the point can be tested for global optimality using a saddle-point condition. First, let us introduce some tools.

Consider a PCP in the form (P, θ). For every fixed θ define the feasible set

$$F(\theta) = \{x : f^i(x, \theta) \leqslant 0, \;\; i \in \mathcal{P}\}.$$

Let $\mathcal{F}$ denote all those θ for which $F(\theta)$ exists, i.e.

$$\mathcal{F} = \{\theta : F(\theta) \neq \phi\}.$$

An important tool in our study is the "minimal index set of active constraints" defined, for every $\theta \in \mathcal{F}$, by

$$\mathcal{P}^{=}(\theta) = \{i \in \mathcal{P} : x \in F(\theta) \Rightarrow f^i(x, \theta) = 0\}.$$

If we denote the set of active constraints at $x \in F(\theta)$ by

$$\mathcal{P}(x, \theta) = \{i \in \mathcal{P} : f^i(x, \theta) = 0\}$$

then one can identify

$$\mathcal{P}^{=}(\theta) = \bigcap_{x \in F(\theta)} \mathcal{P}(x, \theta).$$

This set identifies the constraints that are active on the entire feasible set $F(\theta)$. (For an algorithm that calculates $\mathcal{P}^{=}(\theta)$ at a fixed θ, see, e.g., [1, 4].)

For every feasible point $z^* = (x^*, \theta^*) \in Z$ we first construct the set

$$S(\theta^*) = \{\theta \in \mathcal{F} : \mathcal{P}^{=}(\theta) \subset \mathcal{P}^{=}(\theta^*)\}$$

and then consider the point-to-set mapping $\Phi : R^N \to Z$, defined by

$$\Phi : z^* \to \Phi(z^*) = \{z = (x, \theta) : x \in F(\theta), \theta \in S(\theta^*)\}.$$

Note that $\Phi(z^*) \neq \emptyset$ and $\Phi(z^*) \subset Z$ for every $z^* \in Z$.

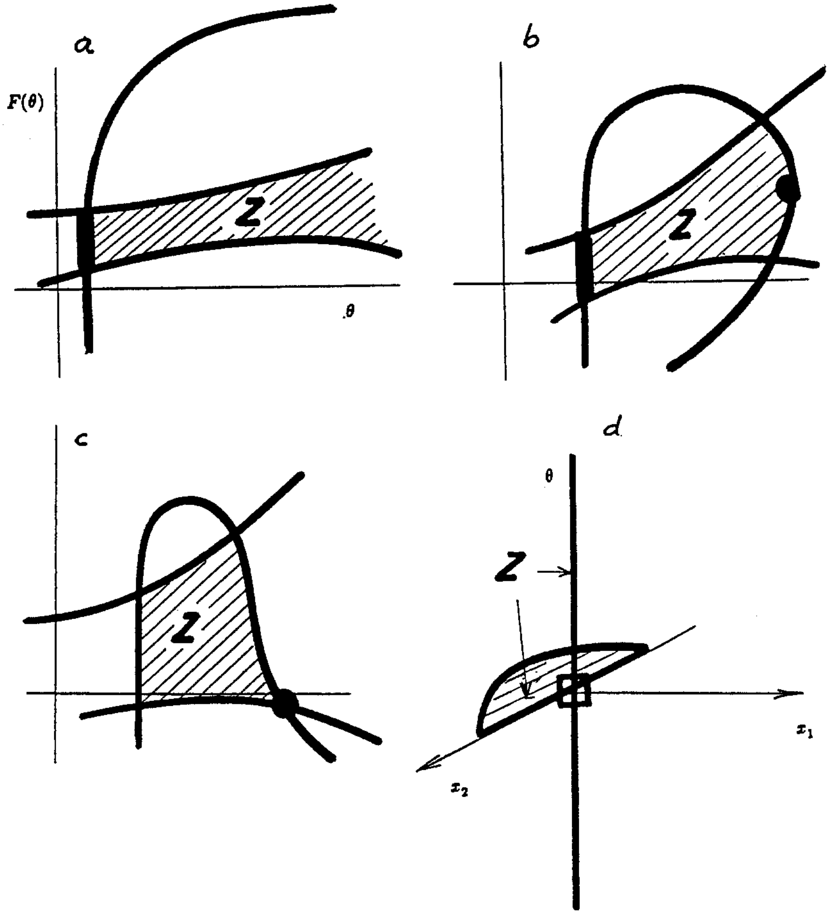

Fig. 1. Cooperation of feasible points.

DEFINITION 2.1. Consider a PCP in the form (P, θ). We say that a feasible point $z^* = (x^*, \theta^*)$ is *globally cooperative* if

$$\Phi(z^*) = Z.$$

The point z^* is *locally cooperative* if $\mathcal{F} \cap N(\theta^*) \subset S(\theta^*)$, in which case

$$\Phi(z^*) = \{z = (x, \theta) : x \in F(\theta), \theta \in S(\theta^*) \cap N(\theta^*)\}$$

where $N(\theta^*)$ is some neighbourhood of θ^*.

Finally, the point is *uncooperative* if θ^* is not the only point in $\mathcal{F}$ and if $S(\theta^*) = \{\theta^*\}$, i.e.,

$$\Phi(z^*) = \{z = (x, \theta^*) : x \in F(\theta^*)\}. \qquad \blacksquare$$

Examples of three types of cooperation are given further below in the text. Some typical situations are depicted in Figure 1. In (a), (b) and (c) every feasible point is locally cooperative. The heavily marked points, and only these points, are globally cooperative. The origin in (d), and only this point, is uncooperative. Every other feasible point is locally cooperative. (See Example 3.3 for details.)

Remarks. (i) It is well known (see [25, 26]) that for lower semicontinuous point-to-set feasible set mapping at the θ level: $F : \theta \to F(\theta)$, the set $S(\theta^*)$ contains a neighbourhood of θ^*. Therefore, for such programs, every feasible point is at least locally cooperative. If F is not lower semicontinuous at $\theta^* \in \mathcal{F}$, then such a neighbourhood may not exist in which case every point $z^* \in \{(x, \theta^*) : x \in F(\theta^*)\}$, is noncooperative. (A point may be cooperative also in the absence of lower semicontinuity.) Recall that a point-to-set mapping $\Gamma : R^p \to R^n$ is said to be lower semicontinuous at $\theta^* \in R^p$ if for every open set $\mathcal{A} \subset R^n$ such that $\mathcal{A} \cap \Gamma(\theta^*) \neq \emptyset$ there exists a neighbourhood $N(\theta^*)$ of θ^* such that $\mathcal{A} \cap \Gamma(\theta) \neq \emptyset$ for every $\theta \in N(\theta^*)$. (See, e.g., [3, 25].)

(ii) For the usual convex programs (the case $z = x$, $\theta = \phi$ in PCP), $S(\theta^*) = \mathcal{F}$ implying that every feasible point is globally cooperative.

(ii) For highly nonconvex programs (the case $z = \theta$, $x = \emptyset$ in PCP), we find that, at $z^* \in Z$,

$$\begin{aligned} \mathcal{P}^{=}(\theta) &= \{i \in \mathcal{P} : x \in R^N \Rightarrow f^i(\theta^*) = 0\} \\ &= \{i \in \mathcal{P} : f^i(\theta^*) = 0\} \\ &= \mathcal{P}(z^*) \end{aligned}$$

i.e., the set of active constraints. The set $S(\theta^*)$ is now

$$\{z \in Z : \mathcal{P}(z) \subset \mathcal{P}^{=}(z^*)\}.$$

This condition is trivially satisfied (by continuity of the constraints). Hence every feasible point of such programs is locally cooperative. (In this paper we will not study this case.)

(iv) If Slater's condition is satisfied at some θ^* of a feasible $z^* = (x^*, \theta^*)$ in a PCP, i.e., if

$$\exists \hat{x} \in R^n \ni f^i(\hat{x}, \theta^*) < 0, \quad i \in \mathcal{P}$$

then $\mathcal{P}^{=}(\theta^*) = \emptyset$, implying $P^{=}(\theta) = \emptyset$ for every θ in some neighbourhood of θ^*, by continuity of the constraints. In this case, again, z^* is at least locally cooperative.

(v) Local cooperation of a feasible point $z^* = (x^*, \theta^*)$ is guaranteed also in some other important situations, e.g., when the point-to-set mapping

$$F^{=} : \theta \to F^{=}(\theta) = \{x \in R^n : f^i(x, \theta) = 0, \ i \in \mathcal{P}^{=}(\theta)\}$$

is lower semicontinuous at θ^*. (See [17, 25].)

3. Characterizing Global Optimality

Global optimality of $z^* = (x^*, \theta^*)$ is described in terms of the existence of a saddle point of the Lagrangian

$$L_*^<(z,u) = f(z) + \sum_{i \in \mathcal{P} \setminus \mathcal{P}^=(\theta^*)} u_i f^i(z)$$

over a set in R^N determined by $\theta \in \mathcal{F}$ and $x \in F_*^=(\theta)$, where

$$F_*^=(\theta) = \{x : f^i(x,\theta) \leqslant 0, \ \ i \in \mathcal{P}^=(\theta^*)\}.$$

The results (except in the linear case, see Section 10) become trivial if $P = \mathcal{P}^=(\theta^*)$. Denote by c the cardinality of the set $\mathcal{P} \setminus \mathcal{P}^=(\theta^*)$ and by R_+^c the non-negative orthant in R^c. Our main result on optimality follows.

THEOREM 3.1. *Consider a partly convex program (P) in the form (P, θ) and its feasible point $z^* = (x^*, \theta^*)$, where x^* is an optimal solution of the convex program (P, θ^*). Assume that z^* is globally cooperative. Then z^* is a global minimum if, and only if, there exists a vector function*

$$\tilde{u} : \mathcal{F} \to \tilde{u}(\theta) \in R_+^c$$

such that

$$L_*^<(z^*, u) \leqslant L_*^<(z^*, \tilde{u}(\theta^*)) \leqslant L_*^<(z, \tilde{u}(\theta)) \tag{3.1}$$

for every

$$z = (x, \theta) \in \{F_*^=(\theta), \mathcal{F}\}$$

and every $u \in R_+^c$.

Proof. Without loss of generality, let us assume that the first c indices are $\mathcal{P} \setminus \mathcal{P}^=(\theta^*)$. (Necessity:) Let $z^* = (x^*, \theta^*)$ be a global minimum, where x^* is an optimal solution of (P, θ^*). For every $\theta \in \mathcal{F}$ construct the set

$$K_1(\theta) = \left\{ y : y \geqslant \begin{bmatrix} f(x,\theta) \\ f^1(x,\theta) \\ \cdots\cdots \\ f^c(x,\theta) \end{bmatrix} \quad \text{for at least one } x \in F_*^=(\theta) \right\}.$$

Also construct

$$K_2 = \left\{ y : y < \begin{bmatrix} f(z^*) \\ 0 \\ \cdots\cdots \\ 0 \end{bmatrix} \right\} \quad \text{in } R^{c+1}.$$

Clearly both sets are convex. Moreover, for every $\theta \in \mathcal{F}$:

$$K_1(\theta) \cap K_2 = \theta.$$

Otherwise, for some $\bar{\theta} \in \mathcal{F}$ and $\bar{x} \in F_*^=(\bar{\theta})$, we would have

$$f^i(\bar{x}, \bar{\theta}) < 0, \quad i \in \mathcal{P} \backslash \mathcal{P}^=(\theta^*)$$

i.e., we would have a feasible point $\bar{z} = (\bar{x}, \bar{\theta})$, such that $f(\bar{z}) < f(z^*)$. This violates *global* optimality of z^*! Hence we conclude that, for every $\theta \in \mathcal{F}$, we can separate $K_1(\theta)$ from K_2, i.e., there exist $\tilde{u}_i = \tilde{u}_i(\theta) \geqslant 0$, $i = 0, 1, \ldots, c$, not all zero, such that

$$\tilde{u}_0 f(z^*) \leqslant \tilde{u}_0 f(x, \theta) + \sum_{i=1}^{c} \tilde{u}_i f^i(x, \theta) \tag{3.2}$$

for every $x \in F_*^=(\theta)$. (Non-negativity of the multipliers follows by the usual arguments.) The crucial step now is to show that $\tilde{u}_0 = \tilde{u}_0(\theta) > 0$ for every $\theta \in \mathcal{F}$. Indeed, if $\tilde{u}_0(\bar{\theta}) = 0$ for some $\bar{\theta} \in \mathcal{F}$, then (3.2) would imply

$$\sum_{i=1}^{c} \tilde{u}_i(\bar{\theta}) f^i(x, \bar{\theta}) \geqslant 0 \tag{3.3}$$

for every $x \in F_*^=(\bar{\theta})$. But we can choose

$$\bar{x} = \bar{x}(\bar{\theta}) \in F(\bar{\theta}) \subset F_*^=(\bar{\theta})$$

such that

$$f^i(\bar{x}, \bar{\theta}) < 0, \quad i \in \mathcal{P} \backslash \mathcal{P}^=(\bar{\theta}). \tag{3.4}$$

Since z^* is globally cooperative, we know that

$$\mathcal{P}^=(\bar{\theta}) \subset \mathcal{P}^=(\theta^*). \tag{3.5}$$

Now (3.4) and (3.5) imply

$$f^i(\bar{x}, \bar{\theta}) < 0, \quad i \in \mathcal{P} \backslash \mathcal{P}^=(\theta^*).$$

This contradicts (3.3), because $\tilde{u}_i(\bar{\theta}) \geqslant 0$, $i \in \mathcal{P} \backslash \mathcal{P}^=(\theta^*)$ and not all equal to zero. Hence $\tilde{u}_0(\theta)$ must be positive for every $\theta \in \mathcal{F}$. The rest of the proof is standard: Without loss of generality we can assume that $\tilde{u}_0(\theta) = 1$, in which case (3.2) becomes

$$f(z^*) \leqslant f(z) + \sum_{i=1}^{c} \tilde{u}_i f^i(z) \tag{3.6}$$

for some $\tilde{u}_i = \tilde{u}_i(\theta) \geqslant 0$, $i = 1, \ldots, c$. After specifying $z = z^*$ in (3.6), we find that

$$\sum_{i=1}^{c} \tilde{u}_i(\theta^*) f^i(z^*) \geqslant 0. \tag{3.7}$$

By feasibility of z^*, the inequality also points in the opposite direction. Hence (3.7) holds with equality. This implies that (3.6) can be written as

$$L_*^<(z^*, \tilde{u}(\theta^*)) \leqslant L_*^<(z, \tilde{u}(\theta)).$$

The "missing" left-hand side inequality

$$L_*^<(z^*, u) \leqslant L_*^<(z^*, \tilde{u}(\theta^*))$$

for every $u \geqslant 0$, follows by feasibility of z^*.

(Sufficiency:) This part of the proof is straightforward and it does not depend on convexity or cooperation. Since z^* is feasible, it follows from the left-hand side inequality in (3.1) that

$$\sum_{i=1}^{c} \tilde{u}_i(\theta^*) f^i(z^*) = 0.$$

The right-hand side inequality in (3.1) now implies

$$\begin{aligned} f(z^*) &\leqslant f(z) + \sum_{i=1}^{c} \tilde{u}_i(\theta) f^i(z) \\ &\leqslant f(z) \end{aligned}$$

for every feasible z. This completes the proof. ∎

The above theorem is useful in the situations where the dimension of θ is low and the mapping $F_*^=$ is simple. The latter is the case, e.g., when Slater's condition holds at θ, in which case $F_c^+(\theta) = R^n$.

EXAMPLE 3.2. Consider the PCP:

$$\begin{aligned} \operatorname*{Min}_{\text{s.t.}} \quad & f = (z_1 + 1)^3 - z_2 \\ & f^1 = z_1 z_2 - 1 \leqslant 0 \\ & f^2 = 1 - z_1 \leqslant 0 \\ & f^3 = -z_2 \leqslant 0. \end{aligned}$$

We want to know whether, say, $z_1^* = 1$, $z_2^* = 1$ is a global minimum.

After identifying $z_1 = \theta$ and $z_2 = x$, the program is rewritten as a partly convex (in fact partly linear) program:

$$\begin{aligned} \underset{\text{s.t.}}{\text{Min}} \quad & f = (\theta + 1)^3 - x \\ & f^1 = x\theta - 1 \leqslant 0 \\ & f^2 = 1 - \theta \leqslant 0 \\ & f^3 = -x \leqslant 0. \end{aligned}$$

We find that

$$\mathcal{P}^{=}(\theta) = \begin{cases} \{2\}, & \text{if } \theta = 1 \\ \theta & \text{if } \theta > 1 \end{cases}$$

and hence $S(\theta^*) = \mathcal{F}$, so the point $z^* = (z_i^*)$ is globally cooperative. We also need the set

$$F_*^{=}(\theta) = \{x : 1 - \theta \leqslant 0\} = R \quad \text{for any } \theta \in \mathcal{F}.$$

The Lagrangian to be used is

$$\begin{aligned} L_*^{<}(z,u) &= f(z) + u_1 f^1(z) + u_3 f^3(z) \\ &= (1+\theta)^3 - x + u_1(x\theta - 1) + u_3(-x). \end{aligned}$$

The point $z_1^* = 1$, $z_2^* = 1$ is a global optimum, according to Theorem 3.1, if, and only if, the saddle-point inequalities (3.1) hold for every $\theta \geqslant 1$ and every x. The condition becomes, after substitution,

$$(1+\theta)^3 - \tilde{u}_1 + x(-1 + \tilde{u}_1\theta - \tilde{u}_3) \geqslant 7 \tag{3.8}$$

for every $\theta \geqslant 1$ and every x. The term in the parentheses must be equal to zero. Hence $\tilde{u}_3 = -1 + \tilde{u}_1\theta$, while $\tilde{u}_3 = 0$ since f^3 is a nonactive constraint. This yields

$$\tilde{u}_1 = \frac{1}{\theta} \quad \text{and} \quad \tilde{u}_3 = 0.$$

After substitution in (3.8) we conclude that, for this choice of multipliers, the saddle-point condition becomes

$$(1+\theta)^3 \geqslant 7 + \frac{1}{\theta} \quad \text{for} \quad \theta \geqslant 1,$$

clearly satisfied. This means that $z_1^* = 1$, $z_z^* = 1$, is a global optimum. ■

The following example shows that Theorem 3.1 does not generally hold for uncooperative points.

EXAMPLE 3.3. Consider the program studied in [15]:

$$\begin{aligned}\underset{\text{s.t.}}{\text{Min}} \quad & f = z_1 z_3 \\ & f^1 = z_2 \leqslant 0 \\ & f^2 = (z_1^2 - z_2) z_3^2 \leqslant 0 \\ & f^3 = \max\{0, z_1^2 + z_2^2 - 1\} \leqslant 0.\end{aligned}$$

This is a PCP relative to the splitting $z_1 = x_1$, $z_2 = x_2$, $z_3 = \theta$. The program is rewritten in the form (P, θ). One finds that

$$\mathcal{P}^{=}(\theta) = \begin{cases} \{2, 3\} & \text{if } \theta = 0 \\ \{1, 2, 3\} & \text{if } \theta \neq 0. \end{cases}$$

The point $z_1^* = z_2^* = z_3^* = 0$ is a global minimum, but a saddle point does not exist. Inded, the Lagrangian, determined at $\theta^* = 0$, is

$$L_*^{<}(z, u) = x_1 \theta + u_1 x_2$$

and the saddle-point condition on $\{F_*^{=}(\theta), \mathcal{F}\}$ reduces to

$$u_1 \geqslant \frac{\theta}{|x_1|} \to \infty \quad \text{for a fixed} \quad \theta > 0 \quad \text{as} \quad x_1 \to 0.$$

The reason for the nonexistence of a saddle point is that $z^* = 0$ is not cooperative. ∎

The following example is interesting because it shows that the global optimality condition does not accept local optima as global (unless they are global optima).

EXAMPLE 3.4. Consider the PCP:

$$\begin{aligned}\underset{\text{s.t.}}{\text{Min}} \quad & f = z_1 z_2^2 \\ & z_1 \leqslant 1.\end{aligned}$$

After the splitting $z_1 = x$, $z_2 = \theta$, the problem is rewritten as

$$\begin{aligned}\underset{\text{s.t.}}{\text{Min}} \quad & f = x\theta^2 \\ & f^1 = x - 1 \leqslant 0.\end{aligned}$$

Since there is no θ in the constraint, the point $z_1^* = 1$, $z_2^* = 0$ is globally cooperative. The Lagrangian is here

$$L_*^{<}(z, u) = x\theta^2 + u_1(x - 1).$$

Consider the locally optimal point $z_1^* = 1$, $z_2^* = 0$. This point is also globally optimal, according to Theorem 3.1, if and only if, there exists $u_1 = u_1(\theta) \geqslant 0$ such that

$$(\theta^2 + u_1)x - u_1 \geqslant 0 \tag{3.9}$$

for every x and every θ. But this is possible if, and only if, $\theta^2 = 0$ for every $\theta \in R$, which is absurd.

Remarks. (i) If z^* is a global optimum of a PCP, but z^* is locally (rather than globally) cooperative, then the saddle point (3.1) exists, but relative to the set

$$z = (x, \theta) \in \{F_*^=(\theta), \mathcal{F} \cap N(\theta^*)\} \tag{3.10}$$

where $N(\theta^*)$ is some neighbourhood of θ^*. (This is a nontrivial remark since no neighbourhood of $x^* \in R^n$ is present in (3.9). The proof of this claim is similar to the proof of Theorem 3.1. The only essential difference is that this time one has to prove that $\tilde{u}_0(\theta) > 0$ for every θ sufficiently close to θ^*. If this was not true, there would exist a sequence $\theta^i \to \theta^*$ such that $\tilde{u}(\theta^i) = 0$. The inequality (3.3) with θ^i, instead of $\bar{\theta}$, is now violated by the existence of a sequence $x^i = x^i(\theta^i) \in F(\theta^i) \subset F_*^=(\theta^i)$ such that

$$f^i(x^i, \theta^i) < 0, \quad i \in \{\mathcal{P} \backslash \mathcal{P}^=(\theta^i)\} \supset \{\mathcal{P} \backslash \mathcal{P}^=(\theta^*)\}).$$

(ii) The above remark is important for convex programs. Namely, for an arbitrary splitting $z = (x, \theta)$ of a convex program (P), the feasible set mapping $F : \theta \to F(\theta)$ is lower semicontinuous. Hence any feasible z is locally cooperative and the local version of Theorem 3.1 applies, regardless of whether the Karush–Kuhn–Tucker (abbreviation below: KKT) conditions hold for (P). (See Section 5 for details.)

(iii) Often, a point z^* tested for global optimality will be "almost" globally co-operative (e.g., co-operative relative to the interior of the feasible set or all but a few feasible points.) In this case its global optimality can be established by continuity (lower semicontinuity) arguments at the disputed points. (For an illustration see Paragraph 8.3 below.) ■

An unpleasant feature of the characterization is that the function $\tilde{u}$ may turn out to be necessarily discontinuous (even at a globally cooperative point), implying discontinuity of the Lagrangian. This is demonstrated with the following example.

EXAMPLE 3.5. Consider the program

$$\begin{aligned} \underset{\text{s.t.}}{\text{Min}} \quad & f = z_1 \\ & f^1 = -z_1 z_2 \leqslant 0 \\ & f^2 = -z_1 - z_2 \leqslant 0. \end{aligned}$$

After the splitting $z_1 = x$, $z_2 = \theta$ we identify the program as a PCP. Its global optimal solution is

$$z_1^* = x^* = 0, \quad z_2^* = \theta^* = 0.$$

Since $\mathcal{P}^=(\theta^*) = \{1\}$ and $\mathcal{P}^=(\theta) = \emptyset$ if $\theta \neq \theta^*$, $\theta \in \mathcal{F}$, z^* is globally cooperative. The Lagrangian is

$$L_*^<(z, u) = x + u_2(-x - \theta)$$

and the saddle-point inequalities reduce to

$$x + \tilde{u}_2(\theta)(-x - \theta) \geqslant 0$$

for every $\theta \geqslant 0$ and every $x \in F_*^=(\theta)$, where

$$F_*^=(\theta) = \begin{cases} R, & \text{if } \theta = 0 \\ R_+, & \text{if } \theta > 0. \end{cases}$$

Hence it follows that the unique multiplier function is

$$\tilde{u}(\theta) = \begin{cases} 1, & \text{if } \theta = 0 \\ 0, & \text{if } \theta \neq 0, \end{cases}$$

and

$$L_*^<(x, \theta, u) = \begin{cases} 0, & \text{if } \theta = 0 \\ x, & \text{if } \theta \neq 0. \end{cases}$$

Note that, in this example, the mapping F is constant and hence lower semicontinuous at θ^* relative to $\mathcal{F}$.

Under suitable "input constraint qualifications" (see [17]), the reference set $F_*^=(\theta)$ in Theorem 3.1 can be significantly simplified (e.g., replaced by the constant set $F^=(\theta^*)$). A new such qualification is introduced in the next section.

4. Global Optimality under the Sandwich Condition

It is obvious that $F(\theta) \subset F_*^=(\theta)$, $\theta \in \mathcal{F}$ at any fixed $\theta^* \in \mathcal{F}$. The requirement that inserts the constant set $F^=(\theta^*)$ between $F(\theta)$ and $F_*^=(\theta)$, $\theta \in \mathcal{F}$ will be termed the "sandwich condition".

DEFINITION 4.1. Consider a PCP in the form (P, θ) and some $\theta^* \in \mathcal{F}$. If

$$F(\theta) \subset F^=(\theta^*) \subset F_*^=(\theta) \;\text{ for every } \theta \in \mathcal{F} \tag{4.1}$$

then the constraints are said to satisfy the *global sandwich condition* (abbreviated: GSC) at θ^*. If the inclusions (4.1) hold for every $\theta \in \mathcal{F} \cap N(\theta^*)$, where $N(\theta^*)$ is some neighbourhood of θ^*, then the constraints satisfy the *local sandwich condition* (abbreviated: LSC) at θ^*. ■

While GSC may appear rather restrictive, LSC is a rather weak condition. (Both GSC and LSC are satisfied in Examples 3.2, 3.3 and 3.4, but not in Example 3.6. LSC is trivially satisfied when the constraints of the program (P, θ^*) satisfy Slater's condition.)

The geometry of GSC is shown in Figure 2. Note that the two "layers" $F_*^=(\theta)$ and $F^=(\theta^*)$ coincide at θ^*.

The following example shows that the notions of a globally cooperative point and of the global sandwich condition are essentially different.

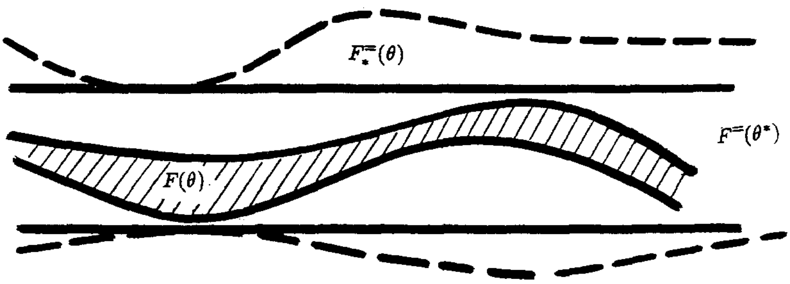

Fig. 2. The global sandwich condition.

EXAMPLE 4.2. Consider a PCP with the single constraint

$$f^1 = x\theta^2 \leqslant 0.$$

Here

$$\mathcal{P}^=(\theta) = \begin{cases} \{1\}, & \text{if } \theta = 0 \\ \emptyset & \text{otherwise.} \end{cases}$$

At $\theta^* = 1$, we find that $F^=(\theta^*) = F^=(\theta) = R$ for every θ. Hence GSC is satisfied. But, since $\mathcal{P}^=(0) \not\subset \mathcal{P}^=(1)$, $z^* = (x^*, 1)$, $x^* \leqslant 0$, cannot be globally cooperative.

On the other hand, any $\bar{z} = (x, 0)$, $x \in R$ is globally cooperative. But, since $F^=(0) = R$ and $F_*^=(\theta) = (-\infty, 0]$ for $\theta \neq 0$, $\bar{\theta} = 0$ does not enjoy GSC. ■

However, LSC implies not only local cooperation of feasible points but also other important information summarized in the theorem below. First, for every $\theta \in \mathcal{F}$ denote by U the point-to-set mapping $U : \theta \to U(\theta) = \{\tilde{u}(\theta)\}$, where $\tilde{u}(\theta)$ are the multipliers in (3.1).

THEOREM 4.3. *Consider a PCP in the form (P, θ) and some $z^* = (x^*, \theta^*) \in Z$. Suppose that the local sandwich condition is satisfied at θ^*. Then both point-to-set mappings $F : \theta \to F(\theta)$ and $F_*^= : \theta \to F_*^=(\theta)$ are lower semicontinuous at θ^*. Moreover, if z^* is a global (or local) minimum, then the sets $U(\theta)$ are nonempty and uniformly bounded for every $\theta \in \mathcal{F}$ sufficiently close to θ^* and the point-to-set mapping U is closed at θ^*.*

Proof. Lower semicontinuity of F follows from the observation that $F(\theta^*) \subset F^=(\theta^*)$ implying $F(\theta^*) \subset F_*^=(\theta)$ in some $\mathcal{F} \cap N(\theta^*)$, where $N(\theta^*)$ is a neighbourhood of θ^*, by LSC. Now use [25, Theorem 3.1]. To prove lower semicontinuity of $F_*^=$ it is enough to observe that $F_*^=(\theta^*) \subset F_*^=(\theta)$ in some feasible neighbourhood $\mathcal{F} \cap N(\theta^*)$, by LSC. Hence

$$\mathcal{A} \cap F_*^=(\theta^*) \neq \emptyset \Rightarrow \mathcal{A} \cap F_*^=(\theta) \neq \emptyset$$

for every open set $\mathcal{A}$, proving the claim. The results about U follow by lower semicontinuity of F and $F_*^=$; see [15, Section 4]. ■

Our motivation for introducing the sandwich conditions is to strengthen the characterizations of optimality. Let us demonstrate how this is possible for a global optimum.

THEOREM 4.4. *Consider a PCP in the form* (P, θ) *and its feasible point* $z^* = (x^*, \theta^*)$ *where* x^* *is an optimal solution of the convex program* (P, θ^*). *Assume that* z^* *is globally cooperative and that the constraints enjoy the global sandwich property. Then* z^* *is a global minimum if, and only if, there exists a vector function*

$$\tilde{u} : \mathcal{F} \to \tilde{u}(\theta) \in R_+^c$$

such that the saddle point inequalities (3.1) hold for every

$$z = (x, \theta) \in \{F^=(\theta^*), \mathcal{F}\}$$

and every $u \in R_+^c$.

Proof. As in the proof of Theorem 3.1 we assume that the first c indices are $\mathcal{P}\backslash\mathcal{P}^=(\theta^*)$. (Necessity:) Let $z^* = (x^*, \theta^*)$ be a global minimum. For every $\theta \in \mathcal{F}$ construct the convex sets

$$K_1(\theta) = \left\{ y : y \geqslant \begin{bmatrix} f(x,\theta) \\ f^1(x,\theta) \\ \cdots\cdots \\ f^c(x,\theta) \end{bmatrix} \quad \text{for at least one } x \in F^=(\theta^*) \right\}$$

and K_2 as in the proof of Theorem 3.1. First we claim that the system

$$\begin{aligned} &f(x,\theta) < f(z^*) \\ &f^i(x,\theta) < 0, \quad i \in \mathcal{P}\backslash\mathcal{P}^=(\theta^*) \\ &x \in F^=(\theta^*) \end{aligned}$$

is inconsistent. (Consistency would imply

$$\begin{aligned} f(x,\theta) &< f(z^*) \\ f^i(x,\theta) &< 0, \quad i \in \mathcal{P}\backslash\mathcal{P}^=(\theta^*) \end{aligned}$$

for some

$$x \in F_*^=(\theta)$$

by the GSC property. Hence it would follow that $x \in F(\theta)$, contradicting global optimality of z^*.) After the separation of the two convex sets we conclude that (3.2) holds for every $x \in F^=(\theta^*)$. If $u_0 = u_0(\bar{\theta}) = 0$ for some $\bar{\theta} \in \mathcal{F}$, then (3.3) would hold for every $x \in F^=(\theta^*)$. But there exists $\bar{x} \in F(\bar{\theta}) \subset F^=(\theta^*)$, again by the GSC, such that (3.4) holds. Since z^* is also assumed to be globally cooperative, the rest of the necessity proof of Theorem 3.1 applies.

(Sufficiency:) The right-hand inequality in (3.1) holds for every $z = (x, \theta) \in \{F^{=}(\theta^*), \mathcal{F}\}$ and hence for every $z = (x, \theta)$, where $x \in F(\theta)$, by the GSC. So $f(z^*) \leqslant f(z)$ for every $z \in F$. ∎

Note the following special case: If $f^i(\cdot, \theta) : R^n \to R$, $i \in \mathcal{P}$ are faithfully convex functions (see, e.g., [4]), then the set $F^{=}(\theta^*)$ is a linear subspace. In this (rather common) situation the characterization of optimality reduces to, essentially, unconstrained optimization in the x space.

A special class of PCPs are convex programs. They are recovered in the next section.

5. Convex Programming

Every convex program (P) is a PCP for every splitting $z = (x, \theta)$. Using the fact that local and global optima coincide, the remarks (i) and (ii) from Section 3, and the observation that convexity is not required in the proofs of saddle-point theorems for the sufficiency parts, we immediately have the following characterization of optimality for the convex case.

THEOREM 5.1. *Consider a convex program (P). For an arbitrary splitting of the variable z, let $z^* = (x^*, \theta^*)$ be its feasible point, where x^* is an optimal solution of the convex program (P, θ^*). Then z^* is an optimal solution of (P) if, and only if, there exists a vector function*

$$\tilde{u} : \mathcal{F} \cap N(\theta^*) \to R^c_+$$

such that the inequalities (3.1) hold for every $z = (x, \theta)$, given by (3.9), and every $u \in R^c_+$.

EXAMPLE 5.2. Consider the convex program

$$\begin{aligned} \underset{\text{s.t.}}{\text{Min}} \quad & f = z_1 + z_3 \\ & f^1 = z_1^2 + z_2^2 - 2 \leqslant 0 \\ & f^2 = (z_1 - 2)^2 + (z_2 - 2)^2 - 2 \leqslant 0 \\ & f^3 = e^{-z_3} - 1 \leqslant 0. \end{aligned} \tag{5.1}$$

The KKT conditions are not satisfied at the feasible solution $z_1^* = z_2^* = 1$, $z_3^* = 0$. Hence these conditions cannot establish optimality of $z^* = (z^*)$.

However, (5.1) is a PCP. Using the splitting, say, $z_1 = \theta_1$, $z_2 = \theta_2$ and $z_3 = x$, we note that $\mathcal{P}^{=}(\theta^*) = \{1, 2\}$, $\mathcal{F} = \{\theta^*\}$, and $F^{=}_*(\theta) = R$, where $\theta^* = (\theta_i^*) \in R^2$, $\theta_1^* = \theta_2^* = 1$. The point $z^* = (\theta^*, x^*)$, where $x^* = 0$, is optimal if, and only if, there exists a multiplier function $\tilde{u}_3 \geqslant 0$ such that

$$x + \tilde{u}_3(e^{-x} - 1) \geqslant 0$$

for every x, by Theorem 5.1. Such a multiplier is $\tilde{u}_3 = 1$. Optimality of z^* can also be verified using other splittings.

Our next objective is to formulate two kinds of methods for calculating global optima of PCPs.

6. An Exact Penalty Function

Global optimality can also be characterized without a cooperation of the feasible points or the sandwich condition. In their absence we will require the existence of the KKT points for every $\theta \in \mathcal{F}$. (For a recently introduced class of convex functions $f^i(\cdot, \theta) : R^n \to R$, $i \in \mathcal{P}$ with the property that every optimal solution is a KKT point, see [24].)

We denote by $f^i_+(z) = \max\{0, f^i(z)\}$, $i \in \mathcal{P}$ and, for a vector function $r : R^p \to R^m$, with non-negative components $r_i(\theta), i \in \mathcal{P}$, the penalty function

$$P(x, \theta) = f(x, \theta) + \sum_{i \in \mathcal{P}} r_i(\theta) f^i_+(x, \theta).$$

Under a condition on the function r, $P(x, \theta)$ is an exact penalty function for the program (P):

THEOREM 6.1. *Consider a partly convex program (P) in the form (P, θ) and a feasible point $z^* = (x^*, \theta^*)$, where x^* is an optimal solution of the convex program (P, θ^*). Assume that for every $\theta \in \mathcal{F}$ the convex program (P, θ) has an optimal solution where the Karush–Kuhn–Tucker conditions are satisfied with some multipliers $\bar{u}_i = \bar{u}_i(\theta)$, $i \in \mathcal{P}$. Let $r : R^p \to R^m$ be a vector function with components satisfying*

$$r_i(\theta) > \bar{u}_i(\theta), \quad i \in \mathcal{P}, \ \theta \in \mathcal{F}. \tag{6.1}$$

Then z^* is a global minimum of (P) if, and only if, z^* is a global minimum of the unconstrained program in the x variable

$$\operatorname*{Min}_{\substack{x \in R^n \\ \theta \in \mathcal{F}}} \ P(x, \theta) \tag{6.2}$$

Proof. First we note that under the assumptions of the theorem, but for a fixed $\theta \in \mathcal{F}$, the optimal solutions and the optimal values of the programs (P, θ) and (6.2) coincide. (The standard proof, e.g., the one from [6, Theorem 25.1], carries through if the "R-regularity" assumption is replaced by the existence of a KKT point.) Obviously, $z^* = (x^*, \theta^*)$ is a global optimal solution of (P) if, and only if, θ^* is a global optimum of the optimal value function

$$\tilde{f}(\theta) = \operatorname*{Min}_{(x)}\{f(x, \theta) : f^i(x, \theta) \leqslant 0, \ i \in \mathcal{P}\}$$

over $\mathcal{F}$. This proves the theorem since $\tilde{f}(\theta)$ coincides with the optimal value function

$$\chi(\theta) = \operatorname*{Min}_{x \in R^n} P(x, \theta). \qquad \blacksquare$$

Note that the relations (6.1) require that r_i be strictly bigger than at *least one* KKT multiplier $\bar{u}_i(\theta)$, not bigger than the set of *all* multipliers, at every $\theta \in \mathcal{F}, i \in \mathcal{P}$.

Two unpleasant features of the exact penalty function $P(x, \theta)$ are nonsmoothness (of the functions f_+^i, $i \in \mathcal{P}$) and a possible loss of continuity. The latter is illustrated with the next example.

EXAMPLE 6.2. Consider the program from Example 3.6. The two KKT multiplier functions are

$$\bar{u}_1(\theta) = \begin{cases} \frac{1}{\theta}, & \text{if } \theta > 0 \\ \alpha, & \text{if } \theta = 0 \end{cases}$$

where α is some arbitrary fixed non-negative number (e.g., $\alpha = 0$) and

$$\bar{u}_2(\theta) = \begin{cases} 0, & \text{if } \theta > 0 \\ 1, & \text{if } \theta = 0. \end{cases}$$

Since $\bar{u}_1(\theta)$ is not uniformly bounded on $\mathcal{F} = [0, \infty)$, the function $r(\theta)$, satisfying (6.1), is necessarily discontinuous. A good choice is

$$r_1(\theta) = \begin{cases} \frac{2}{\theta}, & \text{if } \theta > 0 \\ 1, & \text{if } \theta = 0 \end{cases}, \quad r_2(\theta) = 2, \quad \text{for } \theta \geqslant 0.$$

Hence follows the discontinuity of the exact penalty function

$$P(x, \theta) = \begin{cases} x + \frac{2}{\theta} \max\{0, -x\theta\} + 2\max\{0, -x - \theta\}, & \text{if } \theta \neq 0 \\ x + 2\max\{0, -x\}, \quad \text{at } \theta = 0. \end{cases}$$

A differentiable version of the penalty function $P(x, \theta)$ is

$$\psi(x, \theta) = f(x, \theta) + \sum_{i \in \mathcal{P}} r_i [f_+^i(x, \theta)]^2$$

where r_i, $i \in \mathcal{P}$ are some scalars. While ψ can be used in the same way as $P(x, \theta)$ to solve a PCP, it is not generally an exact penalty function. However, one can still obtain an estimate for the speed of weak convergence using KKT multipliers. (In what follows we denote by $\mathcal{P}(x^*)$ and $\mathcal{P}(\tilde{x}(\theta))$ the sets of active constraints of the programs (P, θ^*) and (P, θ), at the optimal solutions x^* and $\tilde{x}(\theta)$, respectively.)

THEOREM 6.3. *Consider a partly convex program* (P) *in the form* (P, θ) *and its global optimal solution* $z^* = (x^*, \theta^*)$, *where* x^* *is an optimal solution of the* *convex program (P, θ^*) where the Karush–Kuhn–Tucker condition holds with some multipliers $\bar{u}_i \geqslant 0$, $i \in \mathcal{P}(z^*)$. Then for every set of scalars*

$$r_i > \bar{u}_i, \quad i \in \mathcal{P}(z^*)$$

we have

$$\underset{x \in Z}{Min}\, f(z) - \sum_{i \in \mathcal{P}(z^*)} \frac{\bar{u}_i^2}{4r_i} \leqslant \underset{z \in R^N}{Inf}\, \psi(z) \leqslant \underset{z \in Z}{Min}\, f(z).$$

Proof. First note that the proof from [6, Theorem 25.2] carries through if the assumption of "R-regularity" is replaced by the assumption that an optimal solution $\tilde{x}(\theta)$ of the convex program (P, θ), for a fixed θ, is a KKT point. Hence, for every $\theta \in \mathcal{F}$:

$$\underset{x \in F(\theta)}{\mathrm{Min}}\, f(x, \theta) - \sum_{i \in \mathcal{P}(\tilde{x}(\theta))} \frac{[u_i(\theta)]^2}{4r_i(\theta)} \leqslant \underset{x \in R^n}{\mathrm{Inf}}\, \psi(x, \theta) \leqslant \underset{x \in F(\theta)}{\mathrm{Min}}\, f(x, \theta).$$

The result now follows after specifying $\theta = \theta^*$ and using the fact that $z^* = (x^*, \theta^*)$, with $x^* = \tilde{x}(\theta^*)$, is a global minimum. ■

Estimates for the distance between optimal solutions of (P) and ψ, for a PCP, are not yet available.

Using Theorems 6.1 and 6.2 one can formulate a penalty function method for solving the PCP (P):

For a strictly increasing sequence of m-tuples

$$0 < r^k < r^{k+1}, \ldots$$

solve the program (6.2), i.e.,

$$\underset{\substack{x \in R^n \\ \theta \in R^p}}{\mathrm{Min}}\, f(x, \theta) + \sum_{i \in \mathcal{P}} r_i^k f_+^i(x, \theta)$$

to obtain a sequence of unconstrained global optimal solutions $z^k = (x^k, \theta^k)$, $k = 0, 1, 2, \ldots$ It is well known that, assuming compactness of the perturbed feasible set

$$\{z : f^i(z) \leqslant \varepsilon, \ i \in \mathcal{P}\}$$

for some $\varepsilon > 0$, every convergent subsequence of $\{z^k\}$ converges to a global optimum of (P). The novelty here is that, under the additional assumption on the existence of the KKT points for every $\theta \in \mathcal{F}$, it follows from Theorem 6.1 that, for some sufficiently large k, *every* global optimum can be obtained as the limit of such a sequence. (The latter is not generally true for an arbitrary non-PCP.) Similarly, one can use the function ψ to calculate a global optimum. Computational experience is reported in Section 8.

7. A Parametric Optimization Method

Partly convex programs can be solved by the penalty function method described in Section 6. For the sake of comparison, we will also describe a two-level method that employs specific features of PCP programs. The method is adapted from input optimization (see [26]) and it is based on a marginal value formula.

First we need more notation. The formula is applied to a sequence of feasible points.

$$z^k = (x^k, \theta^k), \quad k = 0, 1, 2, \ldots$$

where $\theta^k \in \mathcal{F}$ and $x^k = x^k(\theta^k)$ is an optimal solution of the convex program (P, θ^k). At every iteration we use the Lagrangian function

$$L_k^<(z, u) = f(z) + \sum_{i \in \mathcal{P} \setminus \mathcal{P}^=(\theta^k)} u_i f^i(z), \quad k = 0, 1, 2, \ldots$$

(introduced earlier for testing global optimality). We recall (see [25, Theorem 2.2]) that under the assumption that the set of optimal solutions of $P(\theta^k)$ is nonempty and bounded and the mapping F is lower semicontinuous at θ^k relative to $\mathcal{F}$, there exists $u^k = u^k(\theta) \geqslant 0$ in $R_+^{c(k)}$, where $c(k) = \text{card } \mathcal{P} \setminus \mathcal{P}^=(\theta^k)$, such that

$$L_k^<(z^k, u) \leqslant L_k^<(z^k, u^k) \leqslant L_k^<(x, \theta^k, u^k)$$

for every $x \in F^=(\theta^k) = \{x : f^i(x, \theta^k) \leqslant 0, k \in \mathcal{P}^=(\theta^k)\}$ and every $u \in R_+^{c(k)}$. Hence with every z^k one can associate u^k, $k = 0, 1, 2, \ldots$.

In what follows we will consider a path γ emanating from θ^k and on that path a sequence $\theta^{kl} \to \theta^k$, as $l \to \infty$. Furthermore, we will consider the two limits

$$l^k = \lim_{l \to \infty} \frac{\theta^{kl} - \theta^k}{\|\theta^{kl} - \theta^k\|} \quad \text{and} \quad s^k = \lim_{l \to \infty} \frac{x^{kl} - x^k}{\|\theta^{kl} - \theta^k\|}$$

where $x^{kl} = x^{kl}(\theta^{kl})$ is an optimal solution of the program $P(\theta^{kl})$. (For conditions that guarantee the existence of these limits see, e.g., [16]. Also we denote $z^{kl} = (x^{kl}, \theta^{kl})$. Finally, at each θ^k, we denote by $F_k^=$ the mapping

$$F_k^= : \theta \to F_k^=(\theta) = \{x : f^i(x, \theta) = 0, \ i \in \mathcal{P}^=(\theta^k)\}.$$

The PCP version of the marginal value formula (adapted from [16, 26]), at the k-th iteration, follows.

THEOREM 7.1. *Consider the partly convex program $P(\theta)$ around its arbitrary feasible point $z^k = (x^k, \theta^k)$, where x^k is an optimal solution of the convex program $P(\theta^k)$. Suppose that the point-to-set mappings F and $F_k^=$ are lower semicontinuous at θ^k relative to $\mathcal{F}$. Also assume that the saddle point (x^k, u^k) is unique, and that the functions*

$$\nabla f^i : z \to \nabla f^i(z), \ i \in \{0\} \cup \{\mathcal{P} \setminus \mathcal{P}^=(\theta^k)\}$$

are continuous at z^k. *If, for some sequence* $\theta^{kl} \to \theta^k$, *the limits* l^k *and* s^k *exist, then*

$$\lim_{l\to\infty} \frac{f(z^{kl}) - f(z^k)}{\|\theta^{kl} - \theta^k\|} = \nabla_z L_k^{<}(z^k, u^k) \begin{bmatrix} s^k \\ l^k \end{bmatrix}. \quad \blacksquare \tag{7.1}$$

If the saddle point is not unique then (7.1) can be replaced by two inequalities of the minimax and maximin type. (For details over a "region of stability" see [16, Theorem 4.3].) For the sake of simplicity, we will look only at a simple case of the marginal value formula, describe a simple numerical method, and then show how the method can solve Zermelo's problems.

If we assume that the constraints of the convex program (P, θ^k) satisfy Slater's condition, i.e.,

$$\text{There exists } \hat{x} \text{ such that } f^i(\hat{x}, \theta^k) < 0, \quad i \in \mathcal{P}$$

then $\mathcal{P}^{=}(\theta^k) = \phi$ and $F_k^{=} : \theta \to R^n$ is trivially lower semicontinuous. Also (see, e.g., [25]) F is lower semicontinuous at θ^k. Moreover, the KKT conditions are satisfied at x^k, the Lagrangian becomes

$$L^{<}(z, u) = L(z, u) = f(z) + \sum_{i\in\mathcal{P}} u_i f^i(z)$$

and the x component of the gradient in the right-hand side of (7.1) becomes zero. The marginal value formula now assumes a simpler form:

COROLLARY 7.2. *Consider the partly convex program* $P(\theta)$ *around its arbitrary feasible point* $z^k = (x^k, \theta^k)$, *where* x^k *is an optimal solution of the convex program* $P(\theta^k)$. *Assume that the constraints of* (P, θ^k) *satisfy Slater's condition, that the saddle point* (x^k, u^k) *is unique, and that the functions*

$$\nabla f^i : z \to f^i(z), \quad i \in \{0\} \cup \mathcal{P}$$

are continuous at z^k. *If, for some sequence* $\theta^{kl} \to \theta^k$, *the limit* l^k *exists, then*

$$\lim_{l\to\infty} \frac{f(z^{kl}) - f(z^k)}{\|\theta^{kl} - \theta^k\|} = \nabla_\theta L(z^k, u^k) l^k. \tag{7.2}$$

From now on we will discuss only the formula (7.2). The key observation for numerical methods is that on a path $\theta \to \theta^k$, along which l^k exists and the right-hand side in (7.2) is negative, the value of the objective function must locally decrease, i.e., $f(z^{kl}) < f(z^k)$ for every θ^{kl} sufficiently close to θ^k. The simple choice of an improvable path can be sought in the form $\theta = \theta^k + \alpha d^k, \alpha \geqslant 0$ for some fixed direction d^k. Since now

$$l^k = \frac{d^k}{\|d^k\|}$$

the elements of $d^k = (d_i^k)$ can be chosen by the following simple rule:

$$d_i^k = \begin{cases} 1, & \text{if } [\nabla_\theta L(z^k, u^k)]_i < 0 \\ -1, & \text{otherwise,} \end{cases} \qquad i \in \mathcal{P}. \tag{7.3}$$

This particular choice of d^k guarantees negative marginal value and hence it is a direction of feasible improvement. The corresponding optimal step size $\alpha = \alpha_k$ can be determined by solving the program

$$\begin{aligned} \underset{\text{s.t.}}{\text{Min}} \quad & f(x(\alpha), \theta^k + \alpha d^k) \\ & \theta^k + \alpha d^k \in \mathcal{F} \end{aligned} \tag{7.4}$$

where $x(\alpha)$ is an optimal solution of the convex program $P(\theta^k + \alpha d^k)$ for some $\alpha \geqslant 0$. Since the function $x = x(\alpha)$ is not generally known explicitly, the program (7.4) is solved approximately after testing several values of α. (The Golden Rule Method has been used in [5, 28].) If $\alpha_k > 0$ is a satisfactory step size, one specifies $\theta^{k+1} = \theta^k + \alpha_k d^k$ and calculates x^{k+1}, an optimal solution of the program $P(\theta^{k+1})$. We have thus obtained a new feasible point $z^{k+1} = (x^{k+1}, \theta^{k+1})$ with $f(z^{k+1}) < f(z^k)$.

The above method minimizes the optimal value function (as a function of θ) and it has a tendency to ignore local minima. If z^* is a local minimum of (P) and the kth iteration z^k sufficiently close to z^*, then, under rather weak assumptions (for an input optimization case see [25]) one can derive the speed of convergence

$$|f(z^k) - f(z^*)| < \frac{C}{k}, \quad k = 1, 2, \ldots$$

where C is a constant that depends on the global behaviour of the Lagrangian around the optimal component θ^*. The method is typically slow and the present version is time consuming. However, the experience suggests that it is reliable for the class of piece-wise linear paths (perturbations) if the feasible set Z is "close" to being convex. A practical implementation of the method over nonlinear paths is the subject of current research.

8. Zermelo's Problems

In this section we will identify the classic navigation problems of Zermelo as PCPs. We will then solve the problems using the parametric method suggested in the preceding section and then verify global optimality by the saddle-point characterization from Section 3.

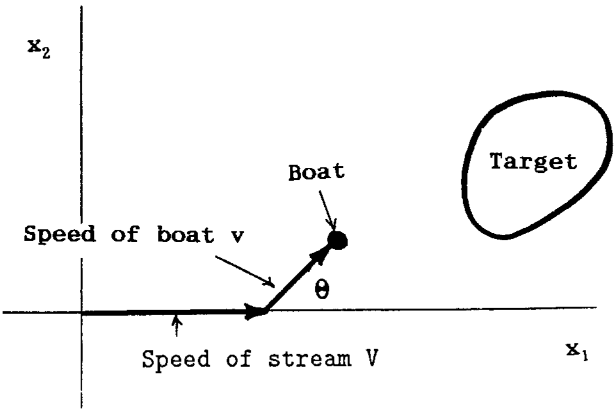

Fig. 3. Zermelo's navigation problem.

8.1. Zermelo's PCP

A boat, situated at the origin, is moving with a velocity v of unit magnitude relative to a stream of constant speed, say, $V = 2$. The problem is to determine a constant steering angle θ that will minimize the time t required to reach a target, say

$$T = \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} : (x_1 - 5)^2 + (x_2 - 1)^2 \leqslant 1 \right\}.$$

(For the sake of comparison we borrow the data from [20].) The dynamics of the system is described by the system of differential equations

$$\begin{aligned} \frac{\mathrm{d}x_1}{\mathrm{d}t} &= 2 + \cos\theta \\ \frac{\mathrm{d}x_2}{\mathrm{d}t} &= \sin\theta. \end{aligned}$$

Its solution (with the initial condition $x_1(0) = x_2(0) = 0$) is

$$\begin{aligned} x_1 &= t(2 + \cos\theta) \\ x_2 &= t\sin\theta. \end{aligned}$$

Zermelo's problem can be formulated as

$$\underset{\text{s.t.}}{\mathrm{Min}} \quad t$$

$$(2t + t\cos\theta - 5)^2 + (t\sin\theta - 1)^2 \leqslant 1. \tag{8.1}$$

Clearly, this is a PCP. (For every θ, the program is convex in t.)

8.2. Solution

Let us start the parametric method from, say,

$$\begin{aligned} z^0 &= (t^0, \theta^0) \\ &= (1.411, 30^0). \end{aligned}$$

(Note that $t^0 = 1.411$ is the optimal solution for $\theta = 30^0$. Since Slater's condition holds at θ^0 we use the classic Lagrangian

$$L(z, u) = t + u[(2t + t\cos\theta - 5)^2 + (t\sin\theta - 1)^2 - 1].$$

The corresponding Lagrange multiplier is $u^0 = 0.173$ (calculated from the KKT condition using the derivatives). Hence

$$\nabla_\theta L(z, u) = 2ut[-(2t + t\cos\theta - 5)\sin\theta + (t\sin\theta - 1)\cos\theta]$$

and

$$\nabla_\theta L(z^0, u^0) = 0.109.$$

Since the derivative is positive, we know that the objective function decreases for the choice $d^0 = -1$.

The step-size problem along d^0 is now just

$$\begin{aligned} &\underset{\text{s.t.}}{\text{Min}} \quad t(\alpha) \\ &\qquad 30^0 - \alpha \in \mathcal{F}. \end{aligned}$$

(Here $\mathcal{F}$ is the segment between zero and, roughly, 72°.)

The optimal step size is calculated by the Golden Rule Method to be around $\alpha_0 = 5.44°$. Hence $\theta^1 = \theta^0 + \alpha_0 d^0 = 24.56^0$, with $t^1 = 1.406$, is a better point. In fact, we have found an optimal point!

8.3. Verification

Let us verify, using Theorem 3.1 and a continuity argument, that $t^* = 1.406$, $\theta^* = 24.56^0$ is a global minimum for the above problem of Zermelo. Slater's condition holds at θ^*; hence the point $z^* = (t^*, \theta^*)$ is locally cooperative. However, since Slater's condition holds for every θ in the interior of $\mathcal{F}$, the cooperation of z^* extends to all feasible points except the extreme points corresponding to the boundary of $\mathcal{F}$. But the mapping F is continuous at these extreme points (the feasible set Z being locally convex), so is the optimal value function. Therefore we conclude that optimality of z^*, once established for the set

$$\Phi(z^*) = \left\{ \begin{bmatrix} x \\ \theta \end{bmatrix} : x \in F(\theta), \ \theta \in \text{int}\mathcal{F} \right\} \subset Z$$

extends to the closure of $\Phi(z^*)$, i.e., to the entire feasible set.

Theorem 3.1 now claims that z^* is a global minimum if, and only if, there exists a function $\tilde{u} : \mathrm{int}\mathcal{F} \to \tilde{u}(\theta) \geqslant 0$ such that

$$t^* \leqslant t + \tilde{u}(\theta)[(2t + t\cos\theta - 5)^2 + (t\sin\theta - 1)^2 - 1]$$

for every $\theta \in \mathrm{int}\mathcal{F}$ and every t. Such a function is

$$\tilde{u}(\theta) = \frac{1}{2}(20\sin\theta + 5\sin 2\theta - 24\sin^2\theta)^{-\frac{1}{2}}. \tag{8.2}$$

This function has a U-shaped graph with $\tilde{u}(\theta) \to +\infty$ as θ approaches the two boundary points of $\mathcal{F}$. (The function is constructed from the KKT condition at θ, after substituting for t the smaller of the two roots of $t = t(\theta)$ in the equality constraint.) This confirms global optimality of z^*.

8.4. Disjoint Feasible Set

The feasible set of the program (8.1) in the (θ, t)-plane relative to the interval $0 \leqslant \theta \leqslant 90°$ is an oval-shaped convex set. (The steering angles close to 90°, or negative close to 0°, result in the boat missing the target.)

A graphing of the feasible set on this particular interval, by the computer, has produced a "tail", i.e., a broken line emanating from the oval, into the area where the feasible set should have been empty. The appearance of the tail warrants another look at the problem. First we note that the inequality constraint can be replaced by the equation

$$(2t + 5\cos\theta - 5)^2 + (t\sin\theta - 1)^2 = 1 \tag{8.3}$$

without changing locally optimal solutions. Now the classic result of Lagrange tells us that, at a local extremum, the gradients of the objective function and the constraint are linearly dependent, i.e.,

$$\mu_0 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 2\mu_1 \begin{bmatrix} (2t + t\cos\theta - 5)(2 + \cos\theta) + (t\sin\theta - 1)\sin\theta \\ -(2t + t\cos\theta - 5)t\sin\theta + (t\sin\theta - 1)t\cos\theta \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

with $(\mu_0, \mu_1) \neq 0$. Clearly, $\mu_1 \neq 0$ and hence

$$-(2t + t\cos\theta - 5)t\sin\theta + (t\sin\theta - 1)t\cos\theta = 0. \tag{8.4}$$

The solutions of (8.3) and (8.4) are candidates for local extrema. In order to find these, let us note that the equations can be simplified:

$$\cot\theta = 5 - 2t$$
$$2t\sin\theta - (4t^2 - 10t)\cos\theta = 5t^2 - 20t + 25.$$

Using the substitution

$$\sin\theta = \frac{1}{\sqrt{1 + \cot^2\theta}} = \frac{1}{\sqrt{1 + (5 - 2t)^2}}$$

the second equation becomes

$$9t^4 - 120t^3 + 546t^2 - 1000t + 625 = 0.$$

Using the method of Weierstrass (see [8]) we can find all four roots simultaneously. First we normalize the equation (division by the leading coefficient 9) and then start iterating from, say, $t_1^0 = 4.0$, $t_2^0 = 1.5$, $t_3^0 = 2.1$ and $t_4^0 = 2.9$. After only five iterations the Weierstrass method has produced the four roots correct to four decimal places:

k	t_1	t_2	t_3	t_4
0	4.0	1.5	2.1	2.9
1	4.3190	1.858	2.5175	5.3110
2	3.4210	1.3449	2.2721	6.2953
3	3.7816	1.4057	2.1710	5.9751
4	3.8155	1.4059	2.1836	5.9284
5	3.8165	1.4059	2.1835	5.9274
6	3.8165	1.4059	2.1835	5.9274

The corresponding steering angles for the roots

$$t_1^* = 3.8165, \quad t_2^* = 1.4059, \quad t_3^* = 2.1835, \quad t_4^* = 5.9274$$

are

$$\theta_1^* = -20.997^0, \quad \theta_2^* = 24.56^0, \quad \theta_3^* = 57.666^0, \quad \theta_4^* = -8.3^0.$$

The solution (θ_2^*, t_2^*) confirms the result obtained by the parametric method, while (θ_3^*, t_3^*) turns out to be a local maximum. The other two solutions (θ_1^*, t_1^*) and (θ_4^*, t_4^*) are physically unacceptable because with these steering angles the boat is definitely going to miss the target. The crucial observation, that explains the existence of the tail in the graph, is that $\sin\theta$ is a periodic function and hence there are also other steering angles that solve (8.4) for t_1^* and t_4^*. These are, e.g., $\bar{\theta}_1 = 171°42'$ and $\bar{\theta}_4 = 159°01'$. The former yields a locally maximal and the latter a locally minimal time for the boat to reach the target. Both new steering angles are directed *against* the flow of the river! Using Theorem 3.1 one can establish that $\bar{z} = (\bar{\theta}_1, t_1^*)$ is actually a globally maximal solution. The global minimality of $z^* = (\theta_2^*, t_2^*)$ extends over the newly "discovered" feasible island with the same U-shaped multiplier function $\tilde{u}(\theta)$ given by (8.2). It appears that the oval's tail has been the computer's way of expressing the fact that there is a disjoint part of the feasible set to the right of the oval. (This part could not be plotted by the computer because the prescribed interval in θ was too short). Hence, using an alternative approach to optimality, we have reconfirmed that the shape, and actually

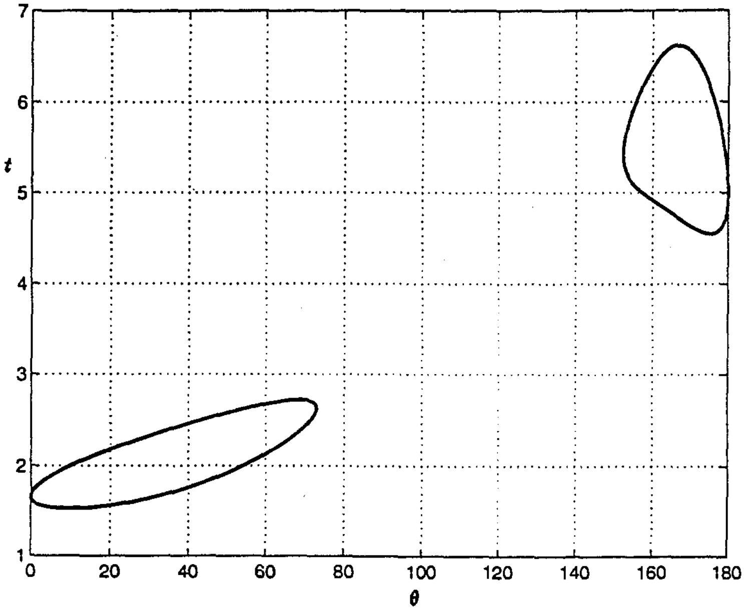

Fig. 4. Feasible set of Zermelo's PCP.

the number of disjoint feasible subsets, depends on the velocity V of the stream. The inter-dependencies can be studied by parametric optimization (e.g., using the results on "regions of stability"). Figure 4 depicts a typical two-island feasible set situation for a Zermelo problem with a convex target.

The geometry and the verification of global optima for the three-dimensional Zermelo PCPs is analogous to the two-dimensional case but more complicated. We will study the three-dimensional case here only briefly, primarily to compare the two types of numerical methods introduced in Sections 6 and 7.

8.5. Three-dimensional Zermelo's Problem

The dynamics of an object (say, a torpedo) moving with a velocity of unit magnitude relative to a three-dimensional medium is described by the system of differential equations

$$\begin{aligned}
\frac{\mathrm{d}x_1}{\mathrm{d}t} &= u + \cos\phi \cdot \cos\psi \\
\frac{\mathrm{d}x_2}{\mathrm{d}t} &= v + \cos\phi \cdot \sin\psi \\
\frac{\mathrm{d}x_3}{\mathrm{d}t} &= w + \sin\phi
\end{aligned}$$

where u, v and w are the components of the velocity vector of the medium and ϕ, ψ are the corresponding angles with the coordinate axes. The problem of finding the constant steering angles that minimize the time required to reach a convex target can be formulated as PCP (with $\theta = (\phi, \psi)$ and $x = t$). Let us illustrate and solve a typical three-dimensional Zermelo's problem.

Assuming that the object is at time $t = 0$ at the origin, that the components of the velocity vector of the medium are $u = 2$, $v = 0$, $w = 0$ and that the target is the unit sphere

$$T = \left\{ \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} : (z_1 - 10)^2 + (z_2 - 1)^2 + (z_3 - 2)^2 \leqslant 1 \right\}$$

the problem becomes

$$\begin{aligned} &\operatorname*{Min}_{\text{s.t.}} \quad t \\ &\qquad (2t + t\cos\phi\cos\psi - 10)^2 + (t\cos\phi\sin\psi - 1)^2 + (t\sin\phi - 2)^2 \leqslant 1. \end{aligned}$$

This is a PCP with $x = t$ and $\theta = (\phi, \psi)$.

Applying the primitive parametric method from the initial point, say, $\theta^0 = (26°, 12°)$, we obtain the following improvements:

k θ^k:	ϕ^k	ψ^k	$\tilde{t}(\theta^k)$	d_1^k	d_2^k	α_k
0	26	12	3.2276	1	1	2.91
1	28.91	14.91	3.2146	−1	1	0.64
2	28.27	15.55	3.2141	1	1	0.07
3	28.34	15.62	3.21406	−1	1	0.01
4	28.33	15.63	3.21406			

Hence the optimal steering angles are $\phi^* = 28.33^0$ and $\psi^* = 15.63°$ and the corresponding shortest time to reach the target is $t^* = 3.214$. Global optimality is verified by the arguments analogous to those used in the two-dimensional case.

8.6. Numerical Experience

A variety of Zermelo's problems has been solved by a group of 20 graduate students in mechanical and electrical engineering, economics and applied mathematics. Using noncommercial programs the success rate has been 75% and 50%, respectively, for 2- and 3-dimensional problems using various penalty function methods and 100%, in each case, using the parametric method. Both methods worked find if the initial approximations were chosen from the feasible islands containing the optimal solutions. However, while the parametric method has taken an advantage of a rather simple oval-shaped feasible set, the penalty function methods struggled with a rather complicated analytic form of the constraints. The commercial versions of the penalty function methods (such as GINO) worked out fine on 2-

and 3-dimensional problems for all values of the parameter r above certain limits (studied in Section 6). Our numerical experience suggests that a further work on the parametric method for solving PCP is warranted and that the method should be competitive in situations where the feasible set is relatively simple, the space of "parameters" θ low dimensional, and the constraints are highly nonconvex.

Zermelo [23] was also studying the situations where the steering angles depend on time. He studied this optimal control problem using a variational approach. (This is one of the directions where the results of this paper could be generalized.) His problems have also been studied by Vincent and Grantham [20] using mathematical programming and the classical optimality conditions.

9. Characterizing Local Optimality

Local optimality of a feasible point $z^* = (x^*, \theta^*)$ of a PCP can be characterized by the existence of a saddle point over the set (3.9), but additional assumptions are required. (An advantage of using (3.9) is that $F_*^=(\theta)$ is generally a larger set than $F_*^=(\theta) \cap N(x^*)$ and hence the necessary condition is more selective. The program from Example 3.4 shows that the existence of a saddle point over the set (3.9) is not a characterization of local optimality!) These assumptions are, e.g., uniqueness of the optimal solution x^* of the convex program (P, θ^*) and lower semicontinuity of the feasible set mapping $F : \theta \to F(\theta)$ at θ^*, relative to $\mathcal{F}$. Recall that the latter is a stronger condition than z^* assumed to be a locally cooperative.

The following result has been recently proved in [28]; also see [27].

THEOREM 9.1. *Consider a PCP and its feasible point $z^* = (x^*, \theta^*)$, where x^* is a unique optimal solution of the convex program (P, θ^*). Also assume that the point-to-set mapping F is lower semicontinuous at θ^*, relative to $\mathcal{F}$. Then z^* is a local minimum if, and only if, there exists a vector function*

$$\tilde{u} : \mathcal{F} \cap N(\theta^*) \to \tilde{u}(\theta) \in R_+^c$$

such that

$$L_*^<(z^*, u) \leqslant L_*^<(z^*, \tilde{u}(\theta^*)) \leqslant L_*^<(z, \tilde{u}(\theta)) \tag{3.1}$$

for every

$$z = (x, \theta) \in \{F_*^=(\theta), \mathcal{F} \cap N(\theta^*)\} \tag{3.9}$$

where $N(\theta^*)$ is some neighbourhood of θ^*, and every $u \in R_+^c$.

Let us make an interesting observation: If z^* is a local minimum, but the saddle-point condition is not satisfied over the strip (3.9), then the optimal solution x^* of the program (P, θ^*), and hence the optimal solution z^* of the program (P), cannot be unique. An illustration follows.

EXAMPLE 9.2. Consider the PCP program from Example 3.5. The point $z_1^* = 1$, $z_2^* = 0$ is a local minimum, F is a constant mapping, but $L_*^{<}(z, u)$ has no saddle point over the set $\{F_*^{=}(\theta), \mathcal{F} \cap N(\theta^*)\} = R \times N(\theta^*)$ for any $N(\theta^*)$. Hence we conclude that the local minimum is not isolated.

The claims of Theorem 4.3 extend to local optimality and Theorem 9.1 can be strengthened under LSC:

COROLLARY 9.3. *Consider a PCP and its feasible point $z^* = (x^*, \theta^*)$, where x^* is a unique optimal solution of the convex program (P, θ^*). Assume that the local sandwich condition is satisfied at θ^*. Then z^* is a local minimum if, and only if, there exists a vector function*

$$\tilde{u} : F \cap N(\theta^*) \to \tilde{u}(\theta) \in R_+^n$$

such that the inequalities (3.1) hold for every $z = (x, \theta)$ on the set (4.2) and every $u \in R_+^c$.

10. Partly Linear Constraints

An important class of PCPs are programs with partly convex objective function and partly linear constraints, i.e., $f^i(\cdot, \theta) : R^n \to R$, $i \in \mathcal{P}$ are linear functions. The results on optimality can now be applied also in the situations where the constraints are equations. The only essential difference is that the multiplier functions, corresponding to the equality constraints, may assume also negative values. Every feasible point is now globally cooperative and local optimality still requires lower semicontinuity of the feasible set mapping and uniqueness of the optimal solution in the complement of "frozen" variables. It is interesting to note that our results are applicable in the situations where the classic method of Lagrange fails. One such situation is described in the following example (augmented from [13,14]).

EXAMPLE 10.1. Consider the program

$$\begin{aligned}
&\text{Min} -z_1 + z_2 \quad -z_4 \\
&\text{s.t.} \\
&\quad z_1^2 = 0 \\
&\quad z_2 - z_3^2 = 0 \\
&\quad -z_4^3 + z_5 - z_6^2 = 0 \\
&\quad z_4^2 - z_5 \quad -z_7^2 = 0.
\end{aligned}$$

We want to check if, say, $z_1^* = z_2^* = z_3^* = z_6^* = z_7^* = 0$, $z_4^* = z_5^* = 1$ is a global minimum. A simple calculation shows that the Lagrange system

$$\nabla f(z^*) + \sum_{i=1}^{4} \lambda_i \nabla f^i(z^*) = 0, \quad \lambda_i \in R, \; i = 1, \ldots, 4$$

is inconsistent.

However, we note, after the identification: $z_1 = \theta_1$, $z_2 = x_1$, $z_3 = \theta_2$, $z_4 = \theta_3$, $z_5 = x_2$, $z_6 = \theta_4$, $z_7 = \theta_5$, that the program is partly linear.

The saddle-point condition for global optimality reduces to

$$\begin{aligned}-1 \leqslant\ & (1+\lambda_2)x_1 + (\lambda_3 - \lambda_4)x_2 - \theta_1 - \theta_3 + \lambda_1\theta_1^2 - \lambda_2\theta_2^2 - \lambda_3(\theta_3^3 + \theta_4^2) \\ & + \lambda_4(\theta_3^2 - \theta_5^2)\end{aligned}$$

for every $x \in R^2$ and every $\theta \in \mathcal{F}$. Hence $\lambda_2 = -1$, $\lambda_3 = \lambda_4$, yielding

$$-1 \leqslant -\theta_1 - \theta_3 + \lambda_1\theta_1^2 + \theta_2^2 - \lambda_3(\theta_3^3 + \theta_4^2 + \theta_5^2 - \theta_3^2) \tag{10.1}$$

for every $\theta \in \mathcal{F}$. The feasibility requirement yields both $\theta_1 = 0$ and $\theta_3^3 + \theta_4^2 + \theta_5^2 - \theta_3^2 = 0$. Hence the test for global optimality reduces to

$$-1 \leqslant -\theta_3 + \theta_2^2.$$

This inequality is certainly satisfied for every $\theta \in \mathcal{F}$. (Otherwise $\theta_3 > 1$, implying $\theta_3^3 > \theta_3^2$. But this violates the λ_3-term in (10.1) being zero.) Global optimality of $z^* = (z_i^*)$ is established.

Remark. Every program with twice continuously differentiable functions can be transformed into a PCP (see [11]). However, this transformation generally destroys some of the useful properties of the original program (e.g., its interior). If an interior-preserving transformation of a nonlinear program into a PCP existed, it would enable us to extend the above results to the general case.

Acknowledgement

The author is indebted to Professor Floudas for drawing his attention to several references. He is also indebted to his students for their cooperation in solving Zermelo's problems. The numerical results in Paragraph 8.5 are obtained by Mr. G. Hepworth and Figure 8.2 is provided by Mr. Najafzadeh-Azghandi.

References

1. R.A. Abrams and L. Kerzner (1978), A simplified test for optimality, *Journal of Optimization Theory and Applications* **25**, 161–170.
2. M. Avriel and A.C. Williams (1971), An extension of geometric programming with applications in engineering optimization, *Journal of Engineering Mathematics* **5**, 187–194.
3. B. Bank, J. Guddat, D. Klatte, B. Kummer, and K. Tammer (1982), *Nonlinear Parametric Optimization*, Akademie-Verlag, Berlin.
4. A. Ben-Israel, A. Ben-Tal, and S. Zlobec (1981), *Optimality in Nonlinear Programming: A Feasible Directions Approach*, Wiley-Interscience, New York.
5. M.P. Brunet (1989), *Numerical Experimentations with Input Optimization*, MSc thesis, McGill University, Department of Mathematics and Statistics, Montreal, Quebec.
6. I.I. Eremin and N.N. Astafiev (1978), *An Introduction to the Theory of Linear and Convex Programming*, Nauka, Moscow. (In Russian.)

7. C.A. Floudas and A. Aggarwal (1990), A decomposition strategy for global optimum search in the pooling problem, *ORSA Journal on Computing* **2**, 119–145.
8. M. Hopkins, B. Marshall, G. Schmidt, and S. Zlobec (1994), On a method of Weierstrass for simultaneous calculation of all roots of an algebraic equation, *Zeitschrift für Angewandte Mathematik und Mechanik* **74**, 295–306.
9. R. Horst and P.M. Pardalos (editors) (1988), *Handbook of Global Optimization*, Kluwer Academic Publishers, Dordrecht.
10. H.Th. Jongen, T. Möbert, and K. Tammer (1988), On iterated minimization in nonconvex optimization, *Mathematics of Operations Research* **11**, 679–691.
11. W.B. Liu and C.A. Floudas (1993), A remark on the GOP algorithm for global optimization, *Journal of Global Optimization* **3**, 519–521.
12. C. Maranas and C.A. Floudas (1992), A global optimization approach for Lennard–Jones microclusters, *Journal of Chemical Physics* **97**, 7667–7678.
13. A. Miele, P.E. Mosaley, A.V. Levy, and G.M. Coggins (1972), On the method of multipliers for mathematical programming problems, *Journal of Optimization Theory and Applications* **10**, 1–33.
14. A. Miele, J.L. Tietze, and A.V. Levy (1972), Comparison of several gradient algorithms for mathematical programming problems, *Aero-Astronautics Report No. 94*, Rice University, Houston, Texas.
15. M. van Rooyen and S. Zlobec (1990), A complete characterization of optimal economic systems with respect to stable perturbations, *Glasnik Matematički* **25**(45), 235–253.
16. M. van Rooyen and S. Zlobec (1993), The marginal value formula on regions of stability, *Optimization* **27**, 17–42.
17. M. van Rooyen, M. Sears, and S. Zlobec (1989), Constraint qualifications in input optimization, *Journal of the Australian Mathematical Society* **30** (Series B), 326–342.
18. F. Sharifi Mokhtarian, Ph.D. Thesis, McGill University (in preparation).
19. V. Visweswaran and C.A. Floudas (1993), New properties and computational improvement of the GOP algorithm for problems with quadratic functions and constraints, *Journal of Global Optimization* **3**, 439–462.
20. T.L. Vincent and W.J. Grantham (1981), *Optimality in Parametric Systems*, Wiley Interscience, New York.
21. I. Zang and M. Avriel (1975), On functions whose local minima are global, *Journal of Optimization Theory and Applications* **16**, 183–190.
22. I. Zang, E.U. Choo, and M. Avriel (1976), A note on functions whose local minima are global, *Journal of Optimization Theory and Applications* **18**, 555–559.
23. E. Zermelo (1931), Über das Navigationsproblem bei Ruhender oder veränderlicher Windverteilung, *Zeitschrift für Angewandte Mathematik und Mechanik* **11**, 114–124.
24. X. Zhou, F. Sharifi Mokhtarian, and S. Zlobec (1993), A simple constraint qualification in convex programming, *Mathematical Programming* **61**, 385–397.
25. S. Zlobec (1988), Characterizing optimality in mathematical programming models, *Acta Applicandae Mathematicae* **12**, 113–180.
26. S. Zlobec (1991), The marginal value formula in input optimization, *Optimization* **22**, 341–386.
27. S. Zlobec (1991), Characterizing optimality in nonconvex optimization, *Yugoslav Journal of Operations Research* **1**, 3–14; Addendum 2 (1991) 69–71.
28. S. Zlobec (1992), Partly convex programming, in V. Bahovec, Lj. Martić, and L. Neralić (eds.), *Zbornik KOI'2* (Proceedings of the Second Conference in Operations Research held in Rovinj, Croatia, October 5–7, 1992), University of Zagreb, Faculty of Economics, pp. 33–50.