Generalized Semi-Infinite Programming: Theory and Methods

G. Still, University of Twente

Abstract

Generalized semi-infinite optimization problems (GSIP) are considered. The difference between GSIP and standard semi-infinite problems (SIP) is illustrated by examples. By applying the 'Reduction Ansatz', optimality conditions for GSIP are derived. Numerical methods for solving GSIP are considered in comparison with methods for SIP. From a theoretical and a practical point of view it is investigated, under which assumptions a GSIP can be transformed into a SIP.

Keywords: Semi-infinite programming, optimality conditions, numerical methods

Mathematical Subject Classification 1991: 90C34, 90C30, 90C31

1 Introduction

We are concerned with *generalized semi-infinite optimization problems* GSIP of the following form:

$$\text{GSIP:} \quad \min \ f(x) \qquad \text{subject to } x \in M = \{x \in I\!R^n \mid g(x,y) \geq 0, \ y \in Y(x)\}$$
$$\text{with} \ \ Y(x) = \{y \in I\!R^r \mid v_l(x,y) \geq 0, \ l \in L\}$$

and L, a finite index set. If not stated otherwise, we assume, that the functions f, g, v_l are C^2-functions and that the set valued mapping Y satisfies

$$Y \ : \ I\!R^n \to 2^{I\!R^r}, \quad Y(x) \subset C_0, \text{ for all } x \in I\!R^n \ \text{ with } \ C_0 \subset I\!R^r \text{ compact.} \tag{1}$$

For the special case that the set $Y = Y(x)$ does not depend on x, i.e. $v_l(x,y) = v_l(y), \ l \in L$, the problem GSIP is a common semi-infinite problem and will be abbreviated by SIP. If moreover Y is a finite set then GSIP reduces to a finite optimization problem.

For a function $f(x)$ the derivative will be denoted by $Df(x)$ and for a function $h(x,y,t)$ by $D_x h, \ D_y h, \ D_t h$ (row vectors) we denote the partial derivatives w.r.t. the variables $x, \ y, \ t$.

For brevity we omit equality constraints in M and $Y(x)$. The paper is organized as follows. In Section 2 we give some examples of GSIP and try to illustrate the difference between GSIP and SIP. Optimality conditions for GSIP are derived in Section 3 by reducing GSIP to a finite problem. Section 4 treats numerical methods. We show that the numerical solution of GSIP can be much more difficult than the solution of SIP. This leads to the question, under which conditions a GSIP can be transformed into a SIP. We answer this question from a theoretical and a practical viewpoint. It is shown that a transformation of GSIP to SIP is possible if in all points of $Y(x)$ the Mangasarian Fromovitz Constraint Qualification (MFCQ) is satisfied.

2 Examples

In this section we give some examples of GSIP. Chebyshev approximation problems lead to semi-infinite problems (cf. e.g.[4]) but also to GSIP. We give an illustrative example.

Example 1 (Chebyshev approximation and reverse Chebyshev approximation) Let be given $f(y) \in C^2(\mathbb{R}^2, \mathbb{R})$ and a space of approximating functions $p(x, y)$, $p \in C^2(\mathbb{R}^n \times \mathbb{R}^2, \mathbb{R})$, parameterized by $x \in \mathbb{R}^n$. We want to approximate f by functions $p(x, \cdot)$ in the max-norm (Chebyshev-norm) on a compact set $Y \subset \mathbb{R}^2$. To minimize the approximation error ϵ, leads to the problem:

$$\min_{x,\epsilon} \ \epsilon \quad \text{s. t.} \quad g^{\pm}(x, y) := \pm\big(f(y) - p(x, y)\big) \leq \epsilon \text{ for all } \ y \in Y \ . \tag{2}$$

This is a SIP, since Y does not depend on (x, ϵ). The so-called reverse Chebyshev problem consists of fixing the approximation error ϵ and making the region Y as large as possible (see [8] for such problems). Suppose, the set $Y = Y(d)$ is parameterized by $d \in \mathbb{R}^2$ and $v(d)$ denotes the volume of $Y(d)$ (e.g. $Y(d) = [-d_1, d_1] \times [-d_2, d_2]$). The reverse Chebyshev problem then leads to the GSIP (ϵ fixed).

$$\max_{d,x} \ v(d) \quad \text{s. t.} \quad g^{\pm}(x, y) := \pm\big(f(y) - p(x, y)\big) \leq \epsilon \text{ for all } \ y \in Y(d) \ . \tag{3}$$

Many control problems in robotics lead to semi-infinite problems (cf. [5]). We give an example.

Example 2 *(Maneuverability problem)*
Let $\Theta = \Theta(t) \in \mathbb{R}^m$ denote the position of the so-called tool center point of the robot (in robot coordinates). Let $\dot\Theta, \ddot\Theta$ be the corresponding velocities, accelerations (derivatives w.r.t. t). The dynamical equation has (often) the form

$$g(\Theta, \dot\Theta, \ddot\Theta) := A(\Theta)\ddot\Theta + F(\dot\Theta, \ddot\Theta) = K,$$

with (external) forces $K \in \mathbb{R}^m$. Here, $A(\Theta)$ is the inertia matrix and F describes the friction, gravity, centrifugal forces, etc. The forces K are bounded by

$$K^- \leq K \leq K^+ \ .$$

For fixed $\Theta, \dot\Theta$, the set of feasible (possible) accelerations is given by

$$Z(\Theta, \dot\Theta) = \{\ddot\Theta \mid K^- \leq g(\Theta, \dot\Theta, \ddot\Theta) \leq K^+\}.$$

Note that, since g is linear in $\ddot\Theta$, for fixed $(\Theta, \dot\Theta)$, the set $Z(\Theta, \dot\Theta)$ is convex (intersection of half-spaces). Let now be given an 'operating region' Q, e.g.

$$Q = \{(\Theta, \dot\Theta) \in \mathbb{R}^{2m} \mid (\Theta^-, \dot\Theta^-) \leq (\Theta, \dot\Theta) \leq (\Theta^+, \dot\Theta^+)\}$$

with bounds $(\Theta^-, \dot\Theta^-)$ and $(\Theta^+, \dot\Theta^+)$. Then, the set of feasible accelerations $\ddot\Theta$ (accelerations which can be realized in every point $(\Theta, \dot\Theta) \in Q$) becomes

$$Z_0 = \bigcap_{(\Theta,\dot\Theta)\in Q} Z(\Theta, \dot\Theta) = \{\ddot\Theta \mid K^- \leq g(\Theta, \dot\Theta, \ddot\Theta) \leq K^+, \text{ for all } (\Theta, \dot\Theta) \in Q\}.$$

The set Z_0 is convex (as an intersection of the convex sets Z). For the steering of the robot one has to check whether a desired acceleration $\ddot\Theta$ is possible, i.e. whether $\ddot\Theta \in Z_0$. Often, this check takes to much time due to the complicated description of Z_0. Then, one is interested in a simple body Y (e.g. a ball) as large as possible, which is contained in Z_0. Instead of the test $\ddot{\Theta} \in Z_0$ one performs the (quicker) check $\ddot{\Theta} \in Y$. Suppose the body $Y(d)$ depends on the parameter $d \in I\!R^q$ and $v(d)$ is the volume of $Y(d)$. Then, to maximize the volume of the body gives the following GSIP called the *maneuverability problem*,

$$\max_d \ v(d) \quad \text{s.t.} \ \ K^- \leq g(\Theta, \dot{\Theta}, \ddot{\Theta}) \leq K^+, \ \text{for all } (\Theta, \dot{\Theta}) \in Q, \ \ddot{\Theta} \in Y(d). \quad (4)$$

Both examples are problems of the following type. The geometrical interpretation is:

Given a family of sets $S(x) \subset I\!R^p$ depending on $x \in I\!R^n$, find a 'body' Y of a given form and a value $\overline{x}$ such that Y is contained in $S(\overline{x})$ and Y is as large as possible.

The mathematical formulation is as follows. Suppose $S(x)$ is defined by

$$S(x) = \{y \in I\!R^p \mid g(x, y, t) \geq 0, \quad \text{for all } t \in Q\}$$

where Q is a given compact set in $I\!R^s$ and $g \in C^2(I\!R^n \times I\!R^p \times I\!R^s, I\!R)$. Let the body $Y(d) \subset I\!R^p$ be parameterized by $d \in I\!R^q$ (by finitely many inequalities). Let $v(d)$ be a measure of the size of $Y(d)$ (e.g. the volume). To maximize $v(d)$ for $Y(d) \subset S(x)$ becomes:

$$\max_{d,x} \ v(d) \quad \text{s.t.} \ \ g(x, y, t) \geq 0 \quad \text{for all } y \in Y(d), \ t \in Q \ . \quad (5)$$

This problem 'contains' the maneuverability problem (4) (choose $n = 0$, i.e. no variable x, $t = (\Theta, \dot{\Theta})$, $y = \ddot{\Theta}$) and the reverse Chebyshev problem (no t-variable).

We give some illustrative theoretical examples to point out the difference between GSIP and SIP.

The feasible set $M = \{x \in I\!R^n \mid g(x, y) \geq 0, \ y \in Y\}$ of SIP is always closed. This need not be the case for GSIP. Consider the problem with $x, y \in I\!R$,

$$\min_x \ x^2 \quad \text{s.t.} \ \ x \leq y \ \ \text{for all} \ \ y \in Y(x) = \{y \mid v(x, y) = (y + 1)^2 + x^2 \leq 0\} \ .$$

Then

$$Y(x) = \begin{cases} \emptyset & \text{for } x \neq 0 \\ -1 & \text{for } x = 0 \end{cases} \qquad \text{and} \quad M = I\!R \setminus \{0\} \ .$$

Note that since $f(x) = x^2$ is minimal at $x = 0$, this GSIP doesn't have a solution. This behavior, that M is not closed, can only occur if the MFCQ (cf. Section 5) is not satisfied for some point $\overline{y} \in Y(\overline{x})$ (the mapping Y is not lower semi-continuous). In our example, this happens for $\overline{x} = 0$, $\overline{y} = -1$, where we have $v(\overline{x}, \overline{y}) = 0$ and $D_y v(\overline{x}, \overline{y}) = 2(\overline{y} + 1) = 0$. See also [9, Section 2] for this phenomenon.

Another difference between SIP and GSIP is, roughly speaking, that for GSIP the feasible set may have re-entrant corners. For SIP this is excluded in the general (generic) case (see [10] for a discussion of this phenomenon called disjunctive problems). We give an example of a GSIP with M having re-entrant corners: $(x \in I\!R^2, \ y \in I\!R)$

$$M = \{x \mid g(x, y) := y - x_2 \geq 0, \ \ y \in Y(x)\}, \quad Y(x) = \{y \mid y \geq x_1, \ y \geq -x_1\} \quad (6)$$

The feasible set becomes

$$M = \{x \in I\!R^2 \mid |x_1| \geq x_2\} \ .$$

Here at the re-entrant corner point $\overline{x} = (0, 0)$ the MFCQ is fulfilled for the (active) point $\overline{y} = 0$ of $Y(\overline{x})$. Such a re-entrant corner is excluded, if Linear Independency

Constraint Qualification (LICQ) is satisfied on $Y(\overline{x})$. (This follows from Theorem 3a below; under LICQ on $Y(x)$ GSIP is equivalent to a 'smooth' SIP).

We finally point out that, in contrast to SIP, even if all problem functions of GSIP are linear, the feasible set need not be convex. The following is well-known for SIP (e.g. [4]):

If for any fixed y, the function $-g(x,y)$ is convex in x, then the feasible set M of SIP is convex.

This follows directly from the fact: Given $x_1,\ x_2 \in M,\ \alpha \in [0,1]$ then

$$g(\alpha x_1 + (1-\alpha)x_2, y) \geq \alpha g(x_1, y) + (1-\alpha) g(x_2, y) \geq 0 \ \text{ for all } \ y \in Y,$$

i.e. $\alpha x_1 + (1-\alpha)x_2 \in M$.

For GSIP the situation is more complicated. Consider for example the feasible set M (of a GSIP) in (6). M is not convex although all functions involved are linear.

3 Reduction Ansatz and optimality conditions

In this section we briefly review the 'Reduction Ansatz' to obtain optimality conditions for GSIP. For $\overline{x} \in M$ we define the *set of active points*

$$Y_0(\overline{x}) = \{\overline{y} \in Y(\overline{x}) \mid g(\overline{x}, \overline{y}) = 0\} .$$

Obviously, for feasible $\overline{x} \in M$, any point $\overline{y} \in Y_0(\overline{x})$ is a (global) minimum of the following parametric optimization problem (the so-called *lower level problem*):

$$Q(\overline{x}): \qquad \min_y g(\overline{x}, y) \quad \text{s.t.} \ \ y \in Y(\overline{x}) . \tag{7}$$

Given $\overline{x} \in M$, for $\overline{y} \in Y(\overline{x})$ we define the active index set $L_0(\overline{x}, \overline{y})$ w.r.t. $Q(\overline{x})$,

$$L_0(\overline{x}, \overline{y}) = \{l \in L \mid v_l(\overline{x}, \overline{y}) = 0\}.$$

and the Lagrange function with $\gamma \in I\!R^{|L_0(\overline{x}, \overline{y})|}$,

$$\mathcal{L}^{\overline{y}}(x, y, \gamma) = g(x, y) - \sum_{l \in L_0(\overline{x}, \overline{y})} \gamma_l v_l(x, y). \tag{8}$$

By assumption (1), the sets $Y(\overline{x})$ are compact. Thus, for any $\overline{x}$, a global minimizer of $Q(\overline{x})$ exists. We will assume, that the following conditions are satisfied for the lower level problem.

A$_{\text{red}}$**:** *We have for any* $\overline{y} \in Y_0(\overline{x})$*:*

1. *LICQ:* $D_y v_l(\overline{x}, \overline{y}), \quad l \in L_0(\overline{x}, \overline{y})$ *are linearly independent.*
2. *Kuhn-Tucker condition: There exists a multiplier* $\overline{\gamma} \in I\!R^{|L_0(\overline{x}, \overline{y})|}$ *such that*

$$D_y \mathcal{L}^{\overline{y}}(\overline{x}, \overline{y}, \overline{\gamma}) = 0$$

and $\overline{\gamma}_l > 0, \quad l \in L_0(\overline{x}, \overline{y})$ *(strict complementary slackness).*

3. *The second order condition (SOC): With* $\overline{\gamma}$ *in 2.,*

$$\eta^T D_y^2 \mathcal{L}^{\overline{y}}(\overline{x}, \overline{y}, \overline{\gamma}) \eta > 0, \quad \text{for all} \ \ \eta \in T(\overline{x}, \overline{y}) \setminus \{0\}$$

where $T(\overline{x}, \overline{y}) = \{\eta \in I\!R^r \mid D_y v_l(\overline{x}, \overline{y}) \eta = 0, \quad l \in L_0(\overline{x}, \overline{y})\}$.

We obtain the following stability result.

Theorem 1 *Suppose, for $\overline{x} \in M$, that the assumption A_{red} is satisfied. Then, the set $Y_0(\overline{x})$ (possibly empty) contains only finitely many points,*

$$Y_0(\overline{x}) = \{\overline{y}^1, \cdots, \overline{y}^p\}$$

and for any $\overline{y}^j \in Y_0(\overline{x})$ (i.e. $\overline{y}^j$ is a minimizer of $Q(\overline{x})$), the following holds: There exist a neighborhood U of $\overline{x}$ and C^1-functions $y^j : U \to I\!R^r$, $y^j(\overline{x}) = \overline{y}^j$, $\gamma_l^j : U \to I\!R$, $\gamma_l^j(\overline{x}) = \overline{\gamma}_l^j$, $l \in L_0(\overline{x}, \overline{y}), j = 1, \cdots, p$, such that for any $x \in U$ the value $y^j(x)$ is a local minimizer of $Q(x)$ (locally unique near $\overline{y}^j$) with corresponding multipliers $\gamma_l^j(x)$. The value functions $g_j(x) = g(x, y^j(x))$ are C^2-functions satisfying for $x \in U$ with the Lagrange functions $\mathcal{L}^j := \mathcal{L}^{\overline{y}^j}$ in (8) the relations,

$$\begin{aligned} Dg_j(x) &= D_x\mathcal{L}^j(x, y^j(x), \gamma^j(x)) \\ D^2g_j(x) &= D_x^2\mathcal{L}^j(x, y^j(x), \gamma^j(x)) - D^T y^j(x) D_y^2 \mathcal{L}^j(x, y^j(x), \gamma^j(x)) Dy^j(x) \\ &\quad - \sum_{l \in L_0(\overline{x}, \overline{y}^j)} D^T\gamma_l^j(x) D_x v_l(x, y^j(x)) + D_x^T v_l(x, y^j(x)) D\gamma_l^j(x). \end{aligned}$$

Proof. The proof is done by applying the implicit function theorem to the following Karush-Kuhn-Tucker equations for $Q(x)$, near $(\overline{x}, \overline{y}^j, \overline{\gamma}^j)$, with Lagrange functions $\mathcal{L}^j = \mathcal{L}^{\overline{y}^j}$ (cf. (8)):

$$F(x, y, \gamma) := \begin{array}{lcl} D_y\mathcal{L}^j(x, y, \gamma) & = & 0 \\ v_l(x, y) & = & 0, \qquad l \in L_0(\overline{x}, \overline{y}^j) \ . \end{array}$$

Under assumption A_{red} the Jacobian $D_{(y,\gamma)}F(\overline{x}, \overline{y}^j, \overline{\gamma}^j)$ is regular and the formula for $Dg_j(x)$, $D^2g_j(x)$ can be obtained by implicitly differentiating the equation $F(x, y^j(x), \gamma^j(x)) = 0$. For more details we refer to [7]. □

Let the assumptions of Theorem 1 hold. Then, in a neighborhood U of $\overline{x}$ the feasible set M of GSIP can be described by finitely many constraints: For any $x \in U$ we have

$$x \in M \iff g_j(x) := g(x, y^j(x)) \geq 0, \quad j = 1, \cdots, p. \tag{9}$$

Consequently, $\overline{x}$ is a local minimizer of GSIP if and only if $\overline{x}$ is a solution of the following reduced problem

$$GSIP_{red}(\overline{x}) : \qquad \min_x f(x) \quad \text{s.t.} \quad g_j(x) := g(x, y^j(x)) \geq 0, \quad j = 1, \cdots, p. \tag{10}$$

$GSIP_{red}(\overline{x})$ is a common finite optimization problem. Thus, the standard optimality conditions of finite optimization can be applied to obtain optimality conditions for GSIP. To that end we define the cone

$$C(\overline{x}) = \{\xi \in I\!R^n \mid Df(\overline{x})\xi \leq 0, \ Dg_j(\overline{x})\xi \geq 0, \ j = 1, \cdots, p\}$$

and the Lagrange function (of the *upper level*)

$$\widetilde{\mathcal{L}}(x, \mu) = \mu_0 f(x) - \sum_{j=1}^{p} \mu_j g_j(x) \ .$$

The following Theorem gives necessary and sufficient optimality conditions of F. John type for GSIP. For more details, necessary conditions and sufficient conditions under weaker assumptions, see [7] but also [9].

Theorem 2 *Suppose, for $\overline{x} \in M$, that the assumption $A_{\rm red}$ is satisfied such that by Theorem 1, in a neighborhood U of $\overline{x}$, GSIP can be locally reduced to $GSIP_{red}(\overline{x})$ according to (10). Then the following holds.*

a. *Suppose, $\overline{x}$ is a local minimizer of GSIP. Then, to any $\xi \in C(\overline{x})$ there exists a multiplier $\overline{\mu} \geq 0$ such that (with $D_x\widetilde{\mathcal{L}}$, $D_x^2\widetilde{\mathcal{L}}$ given below)*

$$D_x\widetilde{\mathcal{L}}(\overline{x},\overline{\mu}) = 0 \quad \text{and} \quad \xi^T D_x^2\widetilde{\mathcal{L}}(\overline{x},\overline{\mu})\xi \geq 0 \; .$$

b. *Suppose, for any $\xi \in C(\overline{x}) \setminus \{0\}$, that there exists a multiplier $\overline{\mu} \geq 0$ such that*

$$D_x\widetilde{\mathcal{L}}(\overline{x},\overline{\mu}) = 0 \quad \text{and} \quad \xi^T D_x^2\widetilde{\mathcal{L}}(\overline{x},\overline{\mu})\xi > 0 \; .$$

Then $\overline{x}$ is a (strict) local minimizer of GSIP.

The expressions for $D_x\widetilde{\mathcal{L}}(\overline{x},\overline{\mu})$ and $D_x^2\widetilde{\mathcal{L}}(\overline{x},\overline{\mu})$ read:

$$D_x\widetilde{\mathcal{L}}(\overline{x},\overline{\mu}) = \mu_0 Df(\overline{x}) - \sum_{j=1}^{p} \overline{\mu}_j D_x g(\overline{x},\overline{y}^j) + \sum_{j=1}^{p} \overline{\mu}_j \Bigg(\sum_{l \in L_0(\overline{x},\overline{y}^j)} \overline{\gamma}_l^j D_x v_l(\overline{x},\overline{y}^j) \Bigg)$$

$$D_x^2\widetilde{\mathcal{L}}(\overline{x},\overline{\mu}) = \mu_0 D^2 f(\overline{x}) - \sum_{j=1}^{p} \overline{\mu}_j D_x^2 g(\overline{x},\overline{y}^j) + \sum_{j=1}^{p} \overline{\mu}_j D^T y^j(\overline{x}) D_y^2 \mathcal{L}^j(\overline{x},\overline{y}^j,\overline{\gamma}^j) D y^j(\overline{x})$$
$$+ \sum_{j=1}^{p} \overline{\mu}_j \sum_{l \in L_0(\overline{x},\overline{y}^j)} \Big(\overline{\gamma}_l^j D_x^2 v_l(\overline{x},\overline{y}^j) + D^T \gamma_l^j(\overline{x}) D_x v_l(\overline{x},\overline{y}^j) + D_x^T v_l(\overline{x},\overline{y}^j) D\gamma_l^j(\overline{x}) \Big)$$

(the first and second terms in $D_x\widetilde{\mathcal{L}}(\overline{x},\overline{\mu})$ and $D_x^2\widetilde{\mathcal{L}}(\overline{x},\overline{\mu})$ are present in finite optimization, the third term in $D_x^2\widetilde{\mathcal{L}}(\overline{x},\overline{\mu})$ is the additional term for SIP and the third term in $D_x\widetilde{\mathcal{L}}(\overline{x},\overline{\mu})$ and the fourth term in $D_x^2\widetilde{\mathcal{L}}(\overline{x},\overline{\mu})$ are typical for GSIP, containing the dependence of v_l (and Y) on x.)

Proof. The formulas for $D_x\widetilde{\mathcal{L}}$ and $D_x^2\widetilde{\mathcal{L}}$ follow immediately by using the formulas for Dg_j, $D^2 g_j$ in Theorem 1. □

4 Numerical methods

In this section we briefly discuss the question of how to compute a solution of GSIP numerically. For a review of methods for SIP we refer to [6], [2] (see also [1]). Below, it will be shown that the numerical solution of GSIP might be much more difficult than the solution of SIP.

Note that with the lower level problem $Q(x)$ (cf. (7)) the GSIP can equivalently be stated as:

$$\text{GSIP}: \quad \min f(x) \;\; \text{s.t.} \;\; g(x,y(x)) \geq 0, \quad \text{where } y(x) \text{ is a global solution of } Q(x).$$

In this form, GSIP has the form of a so-called bi-level problem.

We firstly turn to a method based on local reduction as described in Theorems 1 and 2. This method can directly be generalized from SIP to GSIP. We give a conceptual description (see [6, Section 7.3]).

Algorithm

Step k: Given x^k (not necessarily feasible)

1. Determine the local minima $y^1, \cdots, y^{p_k}$ of $Q(x^k)$.

2. Apply N_k steps (of a finite programming algorithm) to the locally reduced problem (cf. (10)) with $y^j(x)$ the local solutions of $Q(x)$ (cf. (7)),

$$GSIP_{red}(x^k): \qquad \min_x f(x) \quad \text{s.t.} \;\; g_j(x) := g(x, y^j(x)) \geq 0, \;\; j = 1, \cdots, p_k \;,$$

leading to iterates $x^{k,i}, \; i = 1, \cdots, N_k$.

3. Put $x^{k+1} = x^{k,N_k}$ and $k = k + 1$.

The iteration in sub-step 2 can be done by performing 'sequential quadratic programming Newton' steps applied to the Karush-Kuhn-Tucker system of GSIP$_{red}(x^k)$. For a discussion of such a method combining globally convergence and locally superlinear convergence we refer to [6].

Unfortunately, there arise serious difficulties when trying to generalize the so-called exchange or discretization methods from SIP to GSIP. A detailed description of these methods for SIP can be found in [6, Sections 7.1,7.2]. For brevity we will only point out the difficulty. Both methods make use of a discretization of the set Y.

For SIP, this results into a finite problem

$$SIP_d: \qquad \min_x \;\; f(x) \quad \text{s.t.} \;\; g(x, y) \geq 0 \quad \text{for all} \;\; y \in Y_d$$

where $Y_d \subset Y$ is a finite discretization of the compact set Y. For $GSIP$ we would have to choose finite discretizations $Y_d(x)$ of $Y(x)$ and to solve

$$GSIP_d: \quad \min_x \;\; f(x) \quad \text{s.t.} \; x \in M_d := \{x \in I\!R^n \mid g(x, y) \geq 0 \text{ for all } \; y \in Y_d(x)\}$$

This problem represents a (finite) optimization problem of which the number (and quality) of the constraints may change with x. There are no standard procedures for solving such problems GSIP$_d$.

Problems GSIP$_d$ may have all undesirable properties of a GSIP. Even if for the corresponding GSIP the feasible set M is closed, this need not to be the case for the set M_d of GSIP$_d$. Consider an illustrative example of a set M_d :

$$M_d = \{x \in I\!R \mid g(x, y) = x - y \geq 0, \;\; y \in Y_d(x)\}, \quad Y_d(x) = \begin{cases} \{-1, 1\} & \text{if} \;\; x \geq 0 \\ \{-1\} & \text{if} \;\; x < 0 \end{cases}$$

We find

$$M_d = [-1, 0) \cup [1, \infty) \;.$$

In view of these difficulties it is important to investigate which type of GSIP can be transformed to a problem of simpler structure. In [12] a class of GSIP is investigated which can be solved approximately by solving a finite number of convex problems.

5 Transformation of GSIP into SIP

In this section we ask under which conditions a GSIP can be transformed into a SIP. In [14] it has been pointed out that this transformation can be done (at least theoretically) under appropriate compactness assumptions and the assumption that LICQ is satisfied on $Y(x)$. On the other hand, in [3] it has been shown that if the MFCQ is satisfied on $Y(\overline{x})$, for x near $\overline{x}$, the feasible sets $Y(x)$ of the lower level problem are homeomorphic to $Y(\overline{x})$. Thus, a transformation of GSIP into SIP should be possible under MFCQ. Before stating the results we introduce two assumptions.

A$_{\text{MFCQ}}$: *Suppose,* $g, v_l \in C^1(I\!R^n \times I\!R^r, I\!R),\ l \in L$. *The following is valid with a compact set* $K \subset I\!R^n$, *such that* $M \cap K \neq \emptyset$.

1. *The mapping* Y *satisfies condition (1).*
2. *For all* $x,\ y,\ x \in K,\ y \in Y(x)$ *the Mangasarian Fromovitz Constraint Qualification holds:*

$$\text{there exists } \xi\ (= \xi(x,y)) \ \text{ such that } \ D_y v_l(x,y)\xi > 0, \quad l \in L_0(x,y) \qquad (11)$$

A$_{\text{LICQ}}$: *The following is valid with a compact set* $K \subset I\!R^n$, *such that the condition* A_{MFCQ} *holds with MFCQ (cf. (11)) replaced by the stronger Linear Independency Constraint Qualification:*

$$\text{the vectors } D_y v_l(x,y), \quad l \in L_0(x,y) \quad \text{are linearly independent.}$$

Let in the sequel S^r denote the unit sphere, $S^r = \{b \in I\!R^r \mid ||b|| = 1\}$, and $B_\kappa(\overline{y}),\ \kappa > 0$, the ball $B_\kappa(\overline{y}) = \{y \in I\!R^r \mid ||y - \overline{y}|| < \kappa\}$.

Theorem 3 *(Transformation of GSIP into SIP)*

a. *Let be given GSIP such that* A_{LICQ} *is satisfied. Then, there are finitely many* C^1*-functions* $G^j(x,z)$ *and sets* $Z^j = [a_1^j, b_1^j] \times \cdots \times [a_r^j, b_r^j]$ *in* $I\!R^r$, $j = 1, \cdots, P$, *such that*

$$x \in M \cap K \quad \Longleftrightarrow \quad G^j(x,z) \geq 0 \quad \text{for all } \ z \in Z^j,\ j = 1, \cdots, P.$$

b. *Let be given GSIP such that* A_{MFCQ} *is satisfied. Then, there are finitely many Lipschitz continuous functions* $g^j(x,b,\tau)$, $j = 1, \cdots, N$, *such that*

$$x \in M \cap K \quad \Longleftrightarrow \quad g^j(x,b,\tau) \geq 0 \quad \text{for all } \ b \in S^r,\ \tau \in [0,1],\ j = 1, \cdots, N.$$

Proof.

a. A detailed proof can be found in [14]. The proof is based on so-called standard-diffeomorphism which by using coordinate-transformations, locally near a given point $(\overline{x}, \overline{y}),\ \overline{x} \in K,\ \overline{x} \in Y(\overline{x})$ transforms the set $Y(x)$ to canonical form.

b. Let be given $x^j \in K,\ y^j \in Y(x^j)$. Let ξ_0 be a MFCQ-vector satisfying for (x^j, y^j) the conditions (11) and $||\xi_0|| = 1$. Define the point

$$y_*^j = y^j + \rho\xi_0 \qquad (12)$$

with $\rho > 0$ (which will be chosen later). Taylor expansion of $v_l(x^j, y^j + \rho\xi_0)$ around (x^j, y^j) shows that for small $\rho > 0$ we have $y_*^j \in \text{int}\, Y(y^j)$. For fixed $l \in L_0(x^j, y^j)$ we define

$$\xi^l = \frac{D_y v_l(x^j, y^j)}{||D_y v_l(x^j, y^j)||}\ .$$

The MFCQ implies $-\xi_0^T \xi^l < 0$. Now, consider a vector $\overline{b} \in S^r$ such that $\overline{b}^T \xi^l < 0$. By choosing $\rho > 0$ small enough (cf. (12)), there exists a minimum value $t = \overline{t}$ such that for $b = \overline{b}$ the ray

$$y_*^j + tb, \qquad t > 0$$

intersects the solution set of $v_l(x^j, y) = 0$ near y^j. We apply the implicit function theorem to the equation

$$F(x,b,t) := v_l(x, y_*^j + tb) = 0$$

(for (x,b,t) near $(x^j,\overline{b},\overline{t})$). This is possible since by $\overline{b}^T\xi^l<0$ we have $D_tF(x^j,\overline{b},\overline{t}) = D_y v_l(x^j,y^j_*+\overline{t}\overline{b})\,\overline{b}\neq 0$ (here we use $y^j_*+\overline{t}\overline{b}\approx y^j$ and $D_y v_l(x^j,y^j)\,\overline{b}\neq 0$). Consequently, there exist neighborhoods $\overline{U}\times\overline{V}$ of $(x^j,\overline{b})$ and $\overline{W}$ of $\overline{t}$ and a C^1-function $t:\overline{U}\times\overline{V}\to\overline{W}$ such that $t(x^j,\overline{b})=\overline{t}$ and the value $t(x,b)$ is the unique solution in $\overline{W}$ of

$$v_l(x,y^j_*+t(x,b)b)=0, \qquad (x,b)\in\overline{U}\times\overline{V}\ .$$

Consider, with $\epsilon>0$ (small) such that $-\xi_0^T\xi^l=:-\epsilon_0<-\epsilon$, the compact set

$$C_\epsilon=\{b\in S^r\mid b^T\xi^l\le-\epsilon\}\ .$$

By standard arguments using the partition of unity we can glue together finitely many of the functions $t(x,b)$ constructed above, which were defined locally near points $(x^j,\overline{b})$, $\overline{b}\in C_\epsilon$ such that the following holds: There exist $\rho>0$ and a neighborhood U^l of x^j and a C^1-function $t^l:U^l\times C_\epsilon\to I\!R$, $t^l(x^j,-\xi_0)=\rho$ such that for $(x,b)\in U^l\times C_\epsilon$,

$$v_l(x,y^j_*+\tau t^l(x,b)b)\ge 0\ (=0) \quad\Longleftrightarrow\quad \tau\in[0,1]\ (\tau=1).$$

Using the formula $-b^T\xi^l=\cos\varphi$, for the angle φ between b and $-\xi^l$, we find for $b\in C_\epsilon$ (small ρ (cf. (12)),

$$t^l(x^j,b)=\frac{t^l(x^j,\xi^l)}{-b^T\xi^l}+\mathcal{O}(\rho)\ .$$

By continuity, in view of $t^l(x^j,-\xi_0)=\rho,\ -\xi_0^T\xi^l=-\epsilon_0<0$ we can choose $\epsilon_1,\ \epsilon_2,\ \epsilon_3$ such that $\epsilon_0>\epsilon_1>\epsilon_2>\epsilon_3>0$ and such that for all $x\in U^l$ it follows:

$$t^l(x,b)\ \begin{cases}\le 2\rho & \text{if}\quad b^T\xi^l\le-\epsilon_1\\ \ge 2\rho & \text{if}\quad -\epsilon_2\le b^T\xi^l\le-\epsilon_3\ .\end{cases}$$

By defining for $x\in U^l$

$$\tilde{t}^l(x,b):=\begin{cases}\min\{2\rho,t^l(x,b)\} & \text{for}\ \ b\in C_{\epsilon_2}\\ 2\rho & \text{for}\ \ b\in S^r\setminus C_{\epsilon_2}\ ,\end{cases}$$

we obtain a Lipschitz function $\tilde{t}^l$. By construction, for (x,y) in a neighborhood $U^l\times B_{2\rho}(y^j_*)$ of (x^j,y^j), we have

$$v_l(x,y)\ge 0 \quad\Longleftrightarrow\quad y=y^j_*+\tau\tilde{t}^l(x,b)b\ \text{ with }\ \tau\in[0,1],\ b\in S^r\ . \tag{13}$$

This construction can be done for any $l\in L_0(x^j,y^j)$ (with a common choice of a small ρ for all l). Then, we put

$$t_j(x,b):=\min_{l\in L_0(x^j,y^j)}\ \tilde{t}^l(x,b)\quad \textit{and}\quad U_j:=\cap_{l\in L_0(x^j,y^j)}U^l. \tag{14}$$

As the 'minimum' of Lipschitz functions, this function t_j is also Lipschitz continuous. Using (13), for x in the neighborhood U_j of x^j, we have

$$Y(x)\cap B_{2\rho}(y^j_*)=\{y=y^j_*+\tau t_j(x,b)b\mid\tau\in[0,1],\ b\in S^r\}\ .$$

This implies that for $x\in U_j$ the following two conditions are equivalent:

$$\begin{array}{ll} g(x,y)\ge 0 & \text{for all } x\in U_j,\ y\in Y(x)\cap B_{2\rho}(y^j_*)\\ g^j(x,b,\tau):=g(x,y^j_*+\tau t_j(x,b)b)\ge 0, & \text{for all } \tau\in[0,1],\ b\in S^r\ .\end{array} \tag{15}$$

By an appropriate partition of the unity in $I\!R^n$ we can define a Lipschitz function (still denoted by g^j) which is zero for $I\!R^n\setminus U_j$ but coincides with the function g^j (cf.

(15)) on a smaller neighborhood $\hat{U}_j$ contained in U_j. By the assumption A_{MFCQ}, the set $S = \{(x,y) \mid x \in K,\ y \in Y(x)\}$ is compact. Hence, S can be covered by finitely many neighborhoods $\hat{U}_j \times B_{2\rho}(y_*^j)$ of points $(x^j, y^j) \in S$. The corresponding functions $g^j(x,b,\tau),\ j = 1,\cdots,N$, are Lipschitz continuous on $\mathbb{R}^n \times S^r \times [0,1]$ and in view of the the equivalence of the constraints in (15) the statement follows. □

Remark 1 In the proof of Theorem 3b we have constructed functions $t_j(x,b)$ (cf. (14)). For any point (x^j, y^j) with y^j on the boundary of $Y(x^j)$ these functions $t_j(x,b)$ locally near x^j yield a parameterization of the boundary of $Y(x)$. Hence, this construction implicitly contains the following result: If $Y(\overline{x})$ satisfies MFCQ, then locally near $\overline{x}$ the boundaries $\partial Y(x)$ are (Lipschitz-) homeomorphic to $\partial Y(\overline{x})$. Our proof is similar but more elementary than the construction used in [3, Theorem B] to obtain this result (among others). We also refer to [14], where a related construction has been used to obtain optimality conditions for GSIP.

Under further convexity assumptions, the local transformation used in the proof of Theorem 3 yields a global transformation of a GSIP into SIP.

A_s: *Let be given an open set $K_0 \subset \mathbb{R}^n$, ($K_0 \cap M \neq \emptyset$), such that for all $x \in K_0$, the sets $Y(x)$ are star-shaped in the following sense: There exist continuous functions $c : K_0 \to \mathbb{R}^r$, $r : K_0 \times S^r \to \mathbb{R}$, satisfying for all $x \in K_0,\ b \in S^r$*

$$c(x) + \tau\, r(x,b)b \quad \begin{cases} \in \mathit{int}\, Y(x), & \tau \in [0,1) \\ \notin\ Y(x), & \tau \in (1,\infty)\ . \end{cases} \tag{16}$$

Here the interior of $Y(x)$ is denoted by $\mathit{int}\, Y(x)$. Note that since $Y(x)$ is closed this implies that $c(x) + r(x,b)b,\ b \in S^r$ is a parameterization of the boundary $\partial Y(x)$.

The following lemma gives a sufficient condition for A_s.

Lemma 1 *The condition A_s is fulfilled if the following holds on an open set $K_0 \subset \mathbb{R}^n$ ($K_0 \cap M \neq \emptyset$)*

i. The set-valued map satisfies (1) for all $x \in K_0$.

ii. For any $x \in K_0$ the Slater condition is satisfied, i.e. there exists a point $y \in \mathbb{R}^r$ such that $v_l(x,y) > 0,\ l \in L$. (This in particular implies that the sets $Y(x)$ have inner points.)

iii. For any $x \in K_0$ the functions $-v_l(x,y)$ are convex functions of $y \in \mathbb{R}^r$ and $v_l \in C(K_0 \times \mathbb{R}^r, \mathbb{R}),\ l \in L$.

Proof. By conditions i and iii all sets $Y(x)$ are convex and compact. Let $\overline{x}$ be fixed and $\overline{c}$ be an inner point of $Y(\overline{x})$. We firstly show, that for any (fixed) x in a neighborhood $\overline{U}$ of $\overline{x}$ and $b \in S^r$ there exists a value $r(x,\overline{c},b)$ such that

$$\overline{c} + tb \quad \begin{cases} \in \mathrm{int}\, Y(x), & 0 \le t < r(x,\overline{c},b) \\ \notin\ Y(x), & t > r(x,\overline{c},b)\ . \end{cases} \tag{17}$$

To this aim we choose $\overline{U}$ such that $\overline{c} \in \mathrm{int}\, Y(x)$ for all $x \in \overline{U}$ and define

$$h(x,\overline{c},b,t) := \min_{l \in L} v_l(x, \overline{c} + tb) \tag{18}$$

which is continuous (in $x, \overline{c}, b$ and t). We have $h(x,\overline{c},b,0) > 0$ and for a value $t_m > 0$ (large enough) $h(x,\overline{c},b,t_m) < 0$ for all $b \in S^r$ (since $Y(x)$ is bounded). Note that since $-v_l(x,y)$ are convex in y, the function $-h$ is convex in t. Since h is continuous and the upper-level set $\{t \ge 0 \mid h(x,\overline{c},b,t) \ge 0\}$ ($\subset [0,t_m]$) is compact the following function is well-defined,

$$r(x,\overline{c},b) := \arg\max\{t \ge 0 \mid h(x,\overline{c},b,t) \ge 0\}\ . \tag{19}$$

Using convexity, it follows easily that for the solution function $r(x,\overline{c},b)$ the relations (17) hold. Moreover, by continuity arguments, this function $r(x,\overline{c},b)$ is continuously depending on $x \in \overline{U},\ b$ (and $\overline{c}$). Now, we will show that we can choose a continuous function $c(x)$ such that $c(x) \in \operatorname{int} Y(x)$ for $x \in K$. For $x \in \overline{U}$ we obviously have constructed a parameterization of $Y(x)$ of the form

$$Y(x) = \{y = \overline{c} + tb \mid b \in S^r,\ t \in [0, r(x,\overline{c},b)]\} .$$

Hence, with the transformation $T : [0,1] \times S^r \to I\!R^r$ given by $T(t,b) = \overline{c} + tb$, the volume $v(x)$ and the bary-center $c(x)$ of the convex sets $Y(x),\ x \in \overline{U}$ are given by:

$$\begin{aligned} v(x) &= \int_{Y(x)} dy = \int_{b\in S^r} \int_0^{r(x,\overline{c},b)} |\det DT(t,b)| dt\, db \\ c(x) &= \frac{1}{v(x)} \int_{Y(x)} y\, dy = \int_{b\in S^r} \int_0^{r(x,\overline{c},b)} (\overline{c}+tb)\, |\det DT(t,b)|\, dt\, db . \end{aligned}$$

Obviously, both functions $v(x),\ c(x)$ are continuous. We finally show, that $c(x)$ is an inner point of $Y(x)$. To do so, consider for fixed x the support function (with $d \in S^r$)

$$s(d) := \max_{y\in Y(x)} d^T y .$$

In convexity theory it is well-known that a point c is an inner point of the convex set $Y(x)$ if and only if,

$$s(d) > d^T c \qquad \text{for all} \quad d \in S^r$$

(cf. e.g. [13, Theorem 13.1]). Since $\overline{c}$ is an inner point of $Y(x)$ by choosing $\overline{U}$ sufficiently small we can assume that there exists $\kappa > 0$ such that the ball $B_\kappa(\overline{c})$ lies in the interior of $Y(x)$. Consequently we obtain for any $d \in S^r$

$$s(d) - d^T c(x) = \frac{1}{v(x)} \int_{Y(x)} (s(d) - d^T y)\, dy \geq \frac{1}{v(x)} \int_{B_\kappa(\overline{c})} (s(d) - d^T y)\, dy \ > 0 ,$$

i.e. $c(x) \in \operatorname{int} Y(x)$. Now, we choose $\overline{c} = c(x)$ and put $r(x,b) := r(x,c(x),b)$ (cf. (19)). Then, by substituting $t = \tau\, r(x,b)$, the relation (17) is equivalent with (16) and the functions $c(x)$ and $r(x,b)$ satisfy the conditions in A_s. □

Under the assumption A_s the GSIP can be transformed into SIP.

Theorem 4 *Suppose that the assumption A_s is fulfilled in the open set $K_0 \subset I\!R^n$, $(K_0 \cap M \neq \emptyset)$.*

a. *Then, the problem GSIP restricted to K_0 can be written equivalently in the form of the following SIP:*

$$\begin{aligned} \min\ f(x) \quad & s.t. \quad x \in K_0 \quad and \\ & \hat{g}(x,b,\tau) := g(x, c(x) + \tau\, r(x,b)b) \geq 0 \ \textit{for all}\ b \in S^r,\ \tau \in [0,1] \end{aligned} \tag{20}$$

b. *If moreover, for any fixed x, the function $-g(x,y)$ is convex in $y \in I\!R^r$, then the inequality constraints in (20) can be replaced by*

$$\widetilde{g}(x,b) := \hat{g}(x,b,1) = g(x, c(x) + r(x,b)b) \geq 0 \ \textit{for all}\ \ b \in S^r . \tag{21}$$

Proof.
a. The proof follows immediately by noticing that A_s implies

$$Y(x) = \{y \in I\!R^r \mid y = c(x) + \tau\, r(x,b)b, \quad b \in S^r,\ \tau \in [0,1]\} .$$

b. It suffices to show that we have (for fixed $x \in K_0$)

$$\hat{g}(x, b, 1) \geq 0, \quad b \in S^r \quad \Longrightarrow \quad \hat{g}(x, b, \tau) \geq 0, \quad b \in S^r, \ \tau \in [0, 1] \, . \qquad (22)$$

Obviously, the points $y = c(x) + \tau r(x, -b)(-b)$ and $y = c(x) + \tau r(x, b)b$, $\tau \in [0, 1]$ are points on the line segment between $y^- := c(x) + r(x, -b)(-b)$ and $y^+ := c(x) + r(x, b)b$. Using the convexity of $-g$ w.r.t. y the relation (22) follows. □

Remark 2 Even if in Theorem 4 in addition to A_s we would assume that A_{LICQ} is satisfied, the functions $r(x, b)$ and then $\hat{g}$ would only be Lipschitz continuous but in general not C^1-functions. So, to solve the transformed problem (20) we cannot use the 'Newton-method' but we have to apply a discretization method. We emphasize, that for an application of the transformation in Theorem 4 we need not have these functions $c(x)$, $r(x, b)$ explicitly. We only have to compute the corresponding values on every actual discretization.

Often, additional conditions on the problem functions (such as linearity) can be used to get a simpler transformation of the GSIP into a SIP. When for example the conditions of Lemma 1 and of Theorem 4b are both satisfied, then, the conditions $g(x, y) \geq 0$ only need to hold for all extreme points of $Y(x)$. A special class of variational problems is treated in [11]. For an example in connection with the maneuverability problem we refer to [5].

References

[1] Coope I.D., Price C.J., *Computational numerical experiments in semi-infinite programming*, Comput. Optimiz. Appl. 6,169-189, (1996).

[2] Görner S., Reemtsen R., *Numerical methods for semi-infinite programming: A survey*, in 'Semi-infinite Programming', Reemtsen, Rückmann (eds.), Kluwer, Boston, (1998).

[3] Guddat J., Jongen H.Th. and Rückmann J., *On stability and stationary points in nonlinear optimization*, J. Australian Math. Soc., Series B 28, 36-56, (1986).

[4] Hettich R. and Zencke P., *Numerische Methoden der Approximation und der semi-infiniten Optimierung*, Teubner, Stuttgart (1982).

[5] R. Hettich, G. Still, *Semi-infinite programming models in Robotics*, in 'Parametric Optimization and Related Topics II', Guddat et al. (eds.), Akademie Verlag, Berlin (1991).

[6] Hettich R., Kortanek K., Semi-infinite programming: Theory, methods and applications, SIAM Review, vol 35, No.3, 380-429, (1993).

[7] R. Hettich, G. Still, *Second order optimality conditions for generalized semi-infinite programming problems*, Optimization Vol. 34, pp. 195-211, (1995).

[8] Hoffmann A. and Reinhardt R., *On reverse Chebyshev approximation problems*, Technical University of Illmenau, Preprint No. M08/94, (1994).

[9] H. Th Jongen, J.-J. Rückmann, O. Stein, *Generalized semi-infinite optimization: A first order optimality condition and examples*, Mathematical Programming 83, 145-158, (1998).

[10] H. Th Jongen, J.-J. Rückmann, O. Stein, *Disjunctive optimization: critical point theory*, J. Optim. Theory Appl. 93, no. 2, 321–336, (1997).

[11] Kaplan A. and Tichatschke R., *On a class of terminal variational problems*, in 'Parametric Optimization and Related Topics IV', Guddat et al. (eds.), Peter Lang Verlag, Frankfurt a.M., (1997).

[12] Levitin E. and Tichatschke R., *A branch and bound approach for solving a class of generalized semi-infinite programming problems*, J. of Global Optimization 13, 299-315, (1998).

[13] Rockafellar R.T., Convex Analysis, Princeton University Press, Princeton (1970).

[14] Weber W., *Generalized semi-infinite optimization: on some foundations*, Preprint, University of Darmstadt, (1996).