Discretization in Semi-Infinite Programming: The rate of convergence

G. Still*

Abstract

The discretization approach for solving semi-infinite optimization problems is considered. We are interested in the convergence rate of the error between the solution of the semi-infinite problem and the solution of the discretized program depending on the discretization mesh-size. It will be shown how this rate depends on whether the minimizer is strict of order one or two and on whether the discretization includes boundary points of the index set in a specific way. This is done for ordinary and for generalized semi-infinite problems.

Keywords: Semi-infinite programming, discretization methods, rate of convergence.

Mathematical Subject Classification 1991: 90C34, 65K05, 90C30.

1 Introduction

A *semi-infinite optimization problem* is a program with infinitely many constraints,

$$\text{SIP:} \quad \min \ f(x) \qquad \text{s.t.} \ \ x \in F = \{x \in \mathbb{R}^n \mid g(x,y) \leq 0, \ y \in Y\}$$

where $f, \ g$ are given real-valued functions and $Y \subset \mathbb{R}^m$ is a given compact infinite index set.

A common approach for solving SIP are *discretization methods* (see e.g. [8], [13], [15] and the references there). In such a method one chooses a finite grid Y_d of Y, $Y_d \subset Y$, and solves the finite program

$$\text{SIP}(Y_d): \quad \min \ f(x) \qquad \text{s.t.} \ \ x \in F(Y_d) = \{x \in \mathbb{R}^n \mid g(x,y) \leq 0, \ y \in Y_d\} \ .$$

*Faculty of Mathematical Sciences, University of Twente, 7500 AE Enschede, The Netherlands

Throughout the paper, d denotes the Hausdorff distance between Y and Y_d,

$$d := \text{dist}\,(Y_d, Y)\ , \quad \text{where} \;\; \text{dist}\,(Y_d, Y) := \max_{y \in Y} \; \min_{\hat{y} \in Y_d} \; ||\hat{y} - y||\ ,$$

with $||\cdot||$, the Euclidean norm. So, the distance d is a measure for the fineness of the grid Y_d.

Now, let $\overline{x}$ be a local solution of SIP and, for each d, let x_d be a local solution of SIP(Y_d) such that $x_d \to \overline{x}$ for $d \to 0$. In this paper we are interested in the rate of convergence, i.e. in the (largest) exponent $\tau > 0$ such that (with $||\cdot||$, the Euclidean norm in $\mathbb{R}^n$),

$$||x_d - \overline{x}|| \leq O(d^\tau) \qquad\qquad \text{for} \;\; d \to 0\ .$$

This question has been studied in the context of linear Chebyshev approximation (see e.g. [1]). Here, a function f defined on a compact set $Y \subset \mathbb{R}^m$ has to be approximated by a linear combination $p_n(x,y) = \sum_{j=1}^n x_j g_j(y)$ of functions g_j. So, the following problem has to be solved

$$\text{A}(Y): \quad \min_{x \in \mathbb{R}^n} \; e(x) \quad \text{with} \;\; e(x) := \max_{y \in Y} \; |f(y) - p_n(x,y)|\ . \tag{1}$$

Note that this problem is a special instance of a semi-infinite program. In fact, by introducing an extra real variable e the problem A(Y) is equivalent with the linear SIP,

$$\text{SIP}_A\text{:} \quad \min_{x,e} \; e \quad \text{s.t.} \quad \pm\,(f(y) - p_n(x,y)) \leq e \quad \text{for all} \;\; y \in Y\ .$$

The following has been shown in Cheney [1]: Let the functions f, g_j be Lipschitz continuous on Y and let the space span $\{g_1, \ldots, g_n\}$ satisfy the so-called Haar condition. Then the solution $\overline{x}$ of A(Y) and, for each grid Y_d of Y with $|Y_d| \geq n$, the solutions x_d of A(Y_d) are unique, and with some $c > 0$ the following inequality holds,

$$||x_d - \overline{x}|| \leq c\; d \quad \text{for } d \text{ small enough}\ . \tag{2}$$

For linear Chebyshev approximation on an interval $Y = [-1, 1]$, Dunham observed in [2] that the better estimate

$$||x_d - \overline{x}|| \leq \mathcal{O}(d^2)$$

holds if the boundary points ± 1 are contained in the grid Y_d. A similar result is proven for complex approximation on a Jordan curve (cf. [3]).

In the present paper we will generalize these results from linear approximation theory to general (nonlinear) semi-infinite problems.

It is well-known in approximation theory that the Haar condition implies that the solution $\overline{x}$ of $\mathrm{A}(Y)$ is strongly unique (strict of order one), i.e. with some $q > 0$ the inequality

$$e(x) - e(\overline{x}) \geq q \; ||x - \overline{x}|| \qquad \text{for all} \quad x \in \mathbb{R}^n$$

holds (cf. (1)). More generally for SIP, a feasible point $\overline{x} \in F$ is said to be a strict local minimizer of order $p = 1$ or $p = 2$ if there is a neighborhood U of $\overline{x}$ and a constant $q > 0$ such that

$$f(x) - f(\overline{x}) \geq q \; ||x - \overline{x}||^p \qquad \text{for all } x \in F \cap U \; . \tag{3}$$

In the following we will show that the rate of convergence of $||x_d - \overline{x}||$ firstly depends on the order of the minimizer and secondly on the choice of the gridpoints (not merely on the density d). Under weak conditions on SIP, without any assumptions on Y, the following rate holds:

$$||x_d - \overline{x}|| = \mathcal{O}(d^{1/p}) \qquad \text{for } d \to 0 \; .$$

For the minimal values the asymptotic behavior is $f(\overline{x}) - f(x_d) = \mathcal{O}(d)$.

Under additional regularity conditions on Y and if we choose the grid Y_d such that also points on the boundary of Y are included in Y_d in a consistent way (to be defined later on) then the rate behaves as

$$||x_d - \overline{x}|| = \mathcal{O}(d^{2/p}) \qquad \text{for } d \to 0 \; ,$$

and $f(\overline{x}) - f(x_d) = \mathcal{O}(d^2)$.

Such relations have been proven in [17] for a vector-valued approximation problem, transformed to a semi-infinite problem with index sets $Y = S^m$, the unit sphere in $\mathbb{R}^m$. A related result has been given earlier by Opfer in [11] where a complex approximation problem led to a SIP with Y, the unit sphere in the complex plane.

Remark : In approximation theory and semi-infinite optimization, for certain restricted classes of problems, the solution can be shown to be strictly unique of order one (cf. [5]).

For most problem classes however minimizers are of order two. In these cases, in discretization methods, it is important to choose the grid Y_d in such a way that the rate

$$||x_d - \overline{x}|| = \mathcal{O}(d) \qquad \text{for } d \to 0$$

is obtained and not the slower rate $||x_d - \overline{x}|| = \mathcal{O}(\sqrt{d})$.

Note that it can be shown, that generically in smooth semi-infinite optimization, i.e. all problem functions are smooth enough, a minimizer is either strict of order one or strict of order two.

The paper is organized as follows. In Section 2 we present a motivating example. In Section 3 we give the main results on the rate of the error $||x_d - \overline{x}||$ for problem SIP. In Section 4 the analysis is extended to a semi-infinite problem with variable index set, also called generalized SIP.

2 Illustrative examples

The aim of this section is to illustrate the main idea behind the analysis in the next section with simple examples.

We recall that discretization methods roughly speaking are methods for solving SIP which generate a sequence (x_k) of (approximate local) solutions x_k of discretized problems SIP(Y_{d_k}) for a sequence of grids $Y_{d_k} \subset Y$ with $d_k \to 0$ for $k \to \infty$. For theoretical and implementational details of discretization methods we refer the reader to [5], [6], [12], [7], [8], [14], [13], [15], [10], [20].

We emphasize that in the present paper we do not investigate under which conditions solutions x_k of SIP(Y_{d_k}) converge to a solution $\overline{x}$ of SIP (see e.g. [5], [8], [13], [15] for such a question). Instead we will always assume that the following situation is true.

A1 : Let discrete subsets Y_d of Y be given such that $d :=$ dist $(Y, Y_d) \to 0$ holds. The point $\overline{x}$ is a local minimizer of SIP and, for each $d > 0$, x_d is a local solution of SIP(Y_d) such that

$$||x_d - \overline{x}|| \to 0 \qquad \text{for } \; d \to 0 \; .$$

Then we are interested in the rate of convergence of $||x_d - \overline{x}||$ for $d \to 0$. To obtain the improved rate

$$||x_d - \overline{x}|| = \mathcal{O}(d^{2/p}) \quad \text{for } d \text{ small} \quad ,$$

the boundary of Y will play an important role.

To illustrate the main idea behind the general analysis in the next section we work out an example.

For $\overline{x}$ feasible for SIP we define the active index set

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

Let us consider the semi-infinite program

$$\min \; f(x) := -x_1 + x_2 \quad \text{s.t.} \quad x_1(1-y)(1+y) + y^2(y^2 - x_2) \leq 0, \; y \in Y \; , \quad (4)$$

where $Y = [-1, 1]$. We directly conclude that $x_1 \leq 0$ and $x_2 \geq 1$ and that for any feasible x with $\overline{x} := (0, 1)$,

$$f(x) - f(\overline{x}) = -x_1 + x_2 - 1 = |x_1| + |x_2 - 1| \geq ||x - \overline{x}|| \; .$$

Consequently, the point $\overline{x} = (0, 1)$ is the strict minimizer of order one. The active index set is

$$Y_0(\overline{x}) = \{0, -1, 1\} .$$

Take now the discretized problem SIP(Y_d) on the grid $Y_d = \{\pm d,\ \pm 2d, \ldots, \pm(n-1)d\}$, with $d := \frac{1}{n}$. It is seen by setting $g(x, y) = 0$ for $y = \pm d, y = \pm(1-d)$ that the solution $\widetilde{x}_d$ is given by the relation

$$\begin{aligned} x_1(1-d^2) + d^4 - d^2 x_2 &= 0 \\ x_1 d(2-d) + (1-d)^4 - (1-d)^2 x_2 &= 0 . \end{aligned}$$

Consequently, the problem has the strict minimizer $\widetilde{x}_d = (\widetilde{x}_1, \widetilde{x}_2)$ with

$$\widetilde{x}_2 = \frac{(1-d)^4 - \frac{d^5(2-d)}{(1-d)^2}}{(1-d)^2 - \frac{d^3(2-d)}{1-d^2}} = 1 - 2d + o(d),\ \widetilde{x}_1 = \frac{d^2}{1-d^2}\widetilde{x}_2 - \frac{d^4}{(1-d)^2} = d^2 + o(d^2)$$

which is of order one:

$$||\widetilde{x}_d - \overline{x}|| = 2d + o(d) .$$

Now, we add the boundary points ± 1 of Y to the grid, $\hat{Y}_d = Y_d \cup \{-1, 1\}$. Recall that $y = \pm 1$ are index points, active at the minimizer $\overline{x}$. Again by setting $g(x, y) = 0$ for $y = \pm d, y = \pm 1$ it follows that the solution of SIP($\hat{Y}_d$) is defined by the relation

$$\begin{aligned} x_1(1-d^2) + d^4 - d^2 x_2 &= 0 \\ 1 - x_2 &= 0 . \end{aligned}$$

Now, we obtain the solution $\hat{x}_d = (\hat{x}_1, \hat{x}_2)$ with $\hat{x}_2 = 1,\ \hat{x}_1 = \frac{d^2 - d^4}{1-d^2} = d^2 + o(d^2)$ and the quadratic rate

$$||\hat{x}_d - \overline{x}|| = d^2 + o(d^2) .$$

This example suggests that the active index points on the boundary of Y play a more crucial role than the inner active points.

To understand this phenomenon, let us fix d and assume that x_d is not feasible for SIP. Consider the problem

$$\max_y\ g(x_d, y) \qquad \text{s.t.} \quad y \in Y = [-1, 1] . \tag{5}$$

Suppose the solution y_d lies in the interior of Y. Then

$$D_y g(x_d, y_d) = 0 . \tag{6}$$

Choose now a gridpoint $\hat{y}_d \in Y_d$ such that $|\hat{y}_d - y_d| \le d$, which is possible by the definition of d. Then, since $g(x_d, \hat{y}_d) \le 0$ we find (for $g \in C^2$)

$$\begin{aligned} g(x_d, y_d) &\le g(x_d, y_d) - g(x_d, \hat{y}_d) \\ &= D_y g(x_d, y_d)(y_d - \hat{y}_d) + \frac{1}{2} D_y^2 g(x_d, y_d)(y_d - \hat{y}_d)^2 + o\big((y_d - \hat{y}_d)^2\big) \\ &= \mathcal{O}\big((y_d - \hat{y}_d)^2\big) = \mathcal{O}(d^2) . \end{aligned} \tag{7}$$

If the solution y_d of (5) is a boundary point $y = \pm 1$ and y_d is not included in Y_d (i.e. possibly $|\hat{y}_d - y_d| = d$) then instead of (6) (in general) we have $D_y g(x_d, y_d) \neq 0$ and instead of (7) we only obtain

$$g(x_d, y_d) \leq g(x_d, y_d) - g(x_d, \hat{y}_d) = D_y g(x_d, y_d)(y_d - \hat{y}_d) + o(y_d - \hat{y}_d) = \mathcal{O}(d) \ .$$

So, it is crucial to add ‘boundary points’ to the discretization of Y.

In the next section, also for index sets Y in higher dimensions, it is shown how the convergence rate for $||x_d - \overline{x}||$ can be doubled if gridpoints on the boundary of Y are included in the grid Y_d in a certain way.

3 Rate of convergence

Recall that throughout the paper the assumption A1 holds. We are now going to analyze how fast the error $||x_d - \overline{x}||$ tends to zero with the mesh-size d.

The function g is said to be Lipschitz continuous near $\overline{x}$ w.r.t. Y if there exist $\varepsilon > 0$ and $L > 0$ such that

$$|g(x, y) - g(x, \hat{y})| \leq L||y - \hat{y}|| \quad \text{for all } y, \hat{y} \in Y, \ ||x - \overline{x}|| < \varepsilon \ .$$

Note that this condition is fulfilled if g is continuously differentiable on $\mathbb{R}^n \times \mathbb{R}^m$.

Lemma 1 *Let g be Lipschitz continuous near $\overline{x}$ w.r.t. Y with Lipschitz constant $L > 0$. Then there exists $\delta > 0$ such that for all d, $0 < d < \delta$, the following relation holds,*

$$g(x_d, y) \leq L \ d \qquad \textit{for all } \ y \in Y \ .$$

PROOF : Let y_d be a solution of $\max_{y \in Y} \ g(x_d, y)$ and let $\hat{y}_d$ be a point in Y_d such that $||\hat{y}_d - y_d|| \leq d$. By Lipschitz continuity of g and using $g(x_d, \hat{y}_d) \leq 0$ we find

$$g(x_d, y_d) \leq g(x_d, y_d) - g(x_d, \hat{y}_d) \leq \ L||\hat{y}_d - y_d|| \leq L \ d$$

for $||x_d - \overline{x}||$ small enough. $\square$

Under additional regularity conditions on Y and if, roughly speaking, the boundary of Y is also taken in consideration, the inequality in Lemma 1 can be sharpened to

$$g(x_d, y) \leq L \ d^2 \qquad \text{for all } \ y \in Y \ ,$$

(cf. the illustrative example of Section 2). To do so we have to introduce assumptions on g, Y, Y_d.

A2 : Let the following hold:

(α) There is a neighborhood $\overline{U}$ of $\overline{x}$ such that the function $D_y^2 g(x, y)$ is continuous on $\overline{U} \times Y$. (This implies that g is Lipschitz continuous near $\overline{x}$ w.r.t. Y.)

(β) The index set $Y \subset \mathbb{R}^m$ is compact, non-empty and explicitly given as the solution set of inequalities,

$$Y = \{ y \in \mathbb{R}^m \mid v_i(y) \le 0,\ i \in I \} ,$$

where I is a finite index set and $v_i \in C^2(Y)$.

(γ) The *Mangasarian Fromovitz Constraint Qualification* (MFCQ) holds for Y: For any $\overline{y} \in Y$ with the *active index set* $I(\overline{y}) := \{ i \in I \mid v_i(\overline{y}) = 0 \}$ there exists a vector $\overline{\eta} = \eta(\overline{y})$ such hat

$$Dv_i(\overline{y})\overline{\eta} < 0, \quad i \in I(\overline{y}) . \tag{8}$$

The *Linear Independence Constraint Qualification* (LICQ) is said to hold for Y, if for any $\overline{y} \in Y$ the vectors

$$Dv_i(\overline{y}), \quad i \in I(\overline{y}) , \qquad \text{are linearly independent .} \tag{9}$$

Note that the condition LICQ is stronger than MFCQ.

Consider for fixed x under assumption A2 the so-called lower level problem

$$Q(x) \ : \qquad \max_y \ g(x,y) \qquad \text{s.t.} \ \ v_i(y) \le 0,\ i \in I \ .$$

By MFCQ (see (8)), at a solution $\overline{y}$ of $Q(x)$ necessarily the Kuhn-Tucker condition

$$D_y g(x, \overline{y}) - \sum_{i \in I(\overline{y})} \mu_i Dv_i(\overline{y}) = 0$$

holds with Lagrange multipliers $\mu_i \ge 0$. We need a result on the uniformly boundedness of the Lagrange multipliers.

Lemma 2 *Let the assumption A2 hold for Y, g, v_i and let $C \subset \mathbb{R}^n$ be compact. Then, there exists $\alpha > 0$ such that for any $x \in C,\ y \in Y$ and $\mu \ge 0$ satisfying*

$$D_y g(x, y) - \sum_{i \in I(y)} \mu_i Dv_i(y) = 0$$

the bound $||\mu|| \le \alpha$ holds.

PROOF : Assume that on the contrary there are sequences $\{x_k\}, \{y_k\}, \{\mu_k\}$, $x_k \in C,\ y_k \in Y$ and $\mu_k \ge 0$ such that

$$D_y g(x_k, y_k) - \sum_{i \in I(y_k)} (\mu_k)_i Dv_i(y_k) = 0 \tag{10}$$

and $||\mu_k|| \to \infty$ for $k \to \infty$. We divide (10) by $||\mu_k||$ and choose an appropriate subsequence such that $x_{k_\nu} \to \hat{x} \in C$, $y_{k_\nu} \to \hat{y} \in Y$, $I_0 = I(y_{k_\nu})$ (same index set) and $(\mu_{k_\nu})_i/||\mu_{k_\nu}|| \to \hat{\mu}_i$, $i \in I_0$. By continuity it follows $I_0 \subset I(\hat{y})$ and we find from (10) for $k_\nu \to \infty$,

$$-\sum_{i \in I_0} \hat{\mu}_i Dv_i(\hat{y}) = 0$$

with $\hat{\mu} \neq 0$. By multiplying this relation with the MFCQ-vector $\hat{\eta}$ in A2(γ), using $\hat{\mu} \geq 0$, we are led to the contradiction

$$0 = -\sum_{i \in I_0} \hat{\mu}_i Dv_i(\hat{y})\hat{\eta} > 0 \ .$$

□

Motivated by the example in Section 2 we now pay attention to the boundary of the index set Y and consider all boundary parts of all 'dimensions' in Y.

To that end let us fix a subset $I_* \subset I$ such that there exists a point $y_* \in Y$ with $I_* = I(y_*)$. The set

$$S_* := \{y \in Y \mid v_i(y) = 0, \ i \in I_*\}$$

defines a subset of Y. There are finitely many such index sets I_*, say $I_1, I_2, \ldots, I_k$ with corresponding subsets of Y,

$$S_l := \{y \in Y \mid v_i(y) = 0, \ i \in I_l\} \ , \quad l = 1, \ldots, k \ , \tag{11}$$

satisfying $S_l \cap Y \neq \emptyset$. For $I_1 := \emptyset$, for example, S_1 defines the whole set Y. If $I_l \neq \emptyset$, then S_l defines a boundary-part of Y. Note that since $Y \neq \emptyset$, for $\overline{y} \in Y$, by MFCQ the system $Dv_i(\overline{y})\eta \leq -1, \ i \in I(\overline{y})$, has a solution η. Thus, in view of

$$v_i(\overline{y} + t\eta) = v_i(\overline{y}) + tDv_i(\overline{y})\eta + o(t) \leq -t + o(t) < 0 \ , \quad i \in I(\overline{y}) \ ,$$

for all sufficiently small $t > 0$, (every component of) the set Y has interior points.

Remark : Suppose that instead of the assumption MFCQ in A2(γ) the stronger LICQ holds for Y. Then, consider any point $\overline{y} \in S_l$, $l \in \{1, \ldots, k\}$ with $I_l = I(\overline{y})$. By LICQ, the gradients $Dv_i(\overline{y}), \ i \in I_l$, are linearly independent and by using the implicit function theorem, locally near $\overline{y}$, the set S_l appears to be a manifold in $\mathbb{R}^m$ of dimension $m - |I_l|$. With regard to the geometry of polyhedra we can call these sets S_l faces of Y of dimension $m - |I_l|$.

Under the weaker MFCQ the sets S_l may have a more complicated structure. We however will use the term 'faces' for the sets S_l also in this case.

We introduce an assumption on the grid Y_d.

A3 : Suppose Y satisfies A2 and the sets $S_l,\ S_l \cap Y \neq \emptyset,\ l = 1, \ldots, k$, (see (11)) define the subsets of Y as constructed above. Let the grids Y_d be chosen such that

$$\max_{y \in S_l} \min_{\hat{y} \in Y_d \cap S_l} ||\hat{y} - y|| \leq d \quad \text{for all } l = 1, \ldots, k, \text{ and all } d \, .$$

This means that on all 'faces' of Y of all 'dimensions', gridpoints are chosen with mesh-size d.

Remark : As an example let us consider the index set $Y = \{y \in \mathbb{R}^2 \mid |y_i| \leq 1,\ i = 1, 2\}$. Here, a grid Y_d as defined in A3 means that points on the faces '$y_i = \pm 1$' have to be included and the 4 vertices $y = (\pm 1, \pm 1)$ must be added to Y_d.

We are now able to sharpen the result of Lemma 1.

Theorem 1 *Let A2 and A3 hold. Then, with the local solutions x_d of SIP(Y_d) for some $L > 0$ the following inequality is valid for small d:*

$$\max_{y \in Y} g(x_d, y) \leq L\, d^2 \, .$$

PROOF : Given d small, let y_d be a solution of

$$\max_{y} g(x_d, y) \qquad \text{s.t.} \quad v_i(y) \leq 0,\ i \in I \, .$$

Then y_d is contained in some face S_l, l = l(d), with $I_l = I(y_d)$ and y_d must also be a local solution of

$$\max_{y} g(x_d, y) \qquad \text{s.t.} \quad v_i(y) \leq 0,\ i \in I_l \, .$$

By MFCQ (see A2(γ)), y_d satisfies the Kuhn-Tucker condition

$$D_y g(x_d, y_d) - \sum_{i \in I_l} (\mu_d)_i D v_i(y_d) = 0$$

with $\mu_d \geq 0$. By Lemma 2 the sequence μ_d is bounded, $||\mu_d|| \leq K$ for $d \to 0$. In view of A3 (for any d) there is a gridpoint $\hat{y}_d \in S_l \cap Y_d$ such that $||\hat{y}_d - y_d|| \leq d$. Thus using $v_i(y_d) = v_i(\hat{y}_d) = 0,\ i \in I_l$, for d small we find

$$\begin{aligned}
g(x_d, y_d) &\leq g(x_d, y_d) - g(x_d, \hat{y}_d) - \sum_{i \in I_l} (\mu_d)_i \Big(v_i(y_d) - v_i(\hat{y}_d) \Big) \\
&= \Big(D_y g(x_d, y_d) - \sum_{i \in I_l} (\mu_d)_i D v_i(y_d) \Big) (y_d - \hat{y}_d) \\
&\quad + \frac{1}{2} (y_d - \hat{y}_d)^T \Big(D_y^2 g(x_d, y_d) - \sum_{i \in I_l} (\mu_d)_i D^2 v_i(y_d) \Big) (y_d - \hat{y}_d) \\
&\quad + o\left(||y_d - \hat{y}_d||^2 \right) \\
&\leq L\, ||y_d - \hat{y}_d||^2 \leq L d^2
\end{aligned}$$

with some $L > 0$. $\square$

In the next step, from the solution x_d of $\mathrm{SIP}(Y_d)$ we construct a nearby point $\hat{x}_d$ which is feasible for $\mathrm{SIP}(Y)$. To achieve this we assume that the *Mangasarian Fromovitz Constraint Qualification* (MFCQ) is valid for $\overline{x}$: There exists a vector $\xi \in \mathbb{R}^m$ such that

$$D_x g(\overline{x}, \overline{y})\, \xi \leq -1, \quad \text{for all } \ \overline{y} \in Y_0(\overline{x}) \ , \tag{12}$$

with $Y_0(\overline{x})$, the active index set at $\overline{x}$. Note that for semi-infinite formulations of Chebyshev approximation problems the condition MFCQ is automatically fulfilled at any feasible point $\overline{x}$.

A4 : We assume that $D_x g(x, y)$ is continuous on $\overline{U} \times Y$, where $\overline{U}$ is a neighborhood of $\overline{x}$. Let moreover MFCQ be valid at the local minimizer $\overline{x}$ of SIP.

Lemma 3 *Let A4 be satisfied and assume that with some $\gamma > 0$, $c_1 > 0$ the following relation holds:*

$$\max_{y \in Y} \ g(x_d, y) \leq c_1 d^\gamma \ , \qquad \textit{for } d \textit{ small} \, .$$

Then there exists $\rho > 0$ such that with the MFCQ-vector ξ in (12), for small d, the points $\hat{x}_d := x_d + \rho d^\gamma \xi$ are feasible for SIP, i.e.

$$g(\hat{x}_d, y) \leq 0 \ , \quad y \in Y \ .$$

PROOF : We proceed in two steps. Firstly, for $\varepsilon > 0$ we consider the relative open set $Y_0^\varepsilon(\overline{x}) := \{y \in Y \mid ||y - \overline{y}|| < \varepsilon \text{ for some } \overline{y} \in Y_0(\overline{x})\}$. By the MFCQ and the continuity of $D_x g$ there is some $\varepsilon > 0$ such that

$$D_x g(x, y)\xi \leq -\frac{1}{2} \quad \text{ for } \ y \in Y_0^\varepsilon(\overline{x}), \ ||x - \overline{x}|| < \varepsilon \ .$$

Thus if $||\rho d^\gamma \xi|| < \varepsilon$, for all $y \in Y_0^\varepsilon(\overline{x})$ and small d we find,

$$\begin{aligned} g(\hat{x}_d, y) &= g(x_d, y) + \rho d^\gamma D_x g(x_d, y)\xi + o(\rho d^\gamma) \\ &\leq c_1 d^\gamma - \rho \frac{1}{2} d^\gamma + o(\rho d^\gamma) \\ &= d^\gamma (c_1 - \rho \frac{1}{2}) + o(\rho d^\gamma) \leq 0 \end{aligned} \tag{13}$$

if we choose $\rho > 2\, c_1$.

Secondly, we consider the compact set $Y \setminus Y_0^\varepsilon(\overline{x})$. By continuity of g, for given $\varepsilon > 0$, there exists $\varepsilon_1 > 0$ such that

$$g(x, y) < 0 \quad \text{ for all } \ y \in Y \setminus Y_0^\varepsilon(\overline{x}), \ ||x - \overline{x}|| < \varepsilon_1 \ .$$

Together with (13) we find $g(\hat{x}_d, y) \le 0,\ y \in Y$, for $\hat{x}_d$ if

$$||\hat{x}_d - \overline{x}|| = ||x_d - \overline{x} + \rho d^\gamma \xi|| \le ||x_d - \overline{x}|| + \rho d^\gamma ||\xi|| < \min\{\varepsilon, \varepsilon_1\} \ .$$

$\square$

Note that since Y_d is contained in Y, for the values of the objective function at $\overline{x}$ and x_d we trivially have

$$f(\overline{x}) \ge f(x_d) \ .$$

We also will assume that the objective function f is Lipschitz continuous near $\overline{x}$: There exist $\varepsilon > 0$ and $L_f > 0$ such that

$$|f(\hat{x}) - f(x)| \le L_f \ ||\hat{x} - x|| \quad \text{for all } \ ||\hat{x} - \overline{x}|| < \varepsilon, \ ||x - \overline{x}|| < \varepsilon \ .$$

By combining the preceding lemmas we obtain a result on the error $f(\overline{x}) - f(x_d)$.

Corollary 1 *Let A4 hold and let f be Lipschitz continuous near $\overline{x}$. Then the following is true.*

(a) There is some $c_2 > 0$ such that for all d small enough,

$$0 \le f(\overline{x}) - f(x_d) \le c_2 d \ .$$

(b) If in addition, A2 and A3 are satisfied for Y, Y_d, then with some $c_3 > 0$ the inequality

$$0 \le f(\overline{x}) - f(x_d) \le c_3 d^2$$

is valid for small d.

PROOF : By MFCQ, using Lemma 1, Lemma 3 we find with $\gamma = 1$ that for some $\rho > 0$ the points $x_d + \rho d^\gamma \xi$ are feasible,

$$g(x_d + \rho d^\gamma \xi, y) \le 0, \quad y \in Y,$$

if d is small. This implies by Lipschitz continuity of f,

$$f(\overline{x}) - f(x_d) \le f(x_d + \rho d^\gamma \xi) - f(x_d) = \mathcal{O}(d^\gamma) \ .$$

Under the assumptions A2, A3, by using Theorem 1 instead of Lemma 1, we obtain the same relation with $\gamma = 2$. $\square$

The behavior of the error $||x_d - \overline{x}||$ depends on the type of the local minimizer $\overline{x}$, namely on the fact whether $\overline{x}$ is a strict local minimizer of order $p = 1$ or of order $p = 2$.

Theorem 2 *Let the assumptions of Corollary 1 hold. Then if $\overline{x}$ is a strict local minimizer of order $p = 1$ or $p = 2$ the following is true.*

(a) There is some $\sigma_2 > 0$ such that for all d small enough,

$$||x_d - \overline{x}|| \leq \sigma_1 \; d^{1/p} \; .$$

(b) If in addition A2 and A3 is satisfied for Y, Y_d, then with some $\sigma_2 > 0$ the inequality

$$||x_d - \overline{x}|| \leq \sigma_2 \; d^{2/p}$$

is valid.

PROOF : By Corollary 1, with $\gamma = 1$ in (a) or $\gamma = 2$ in (b) the vectors $x_d + \rho d^\gamma \xi$ are feasible for SIP and with the factor q in (3) we obtain

$$\begin{aligned} ||x_d + \rho d^\gamma \xi - \overline{x}||^p &\leq \frac{f(x_d + \rho d^\gamma \xi) - f(\overline{x})}{q} \\ &= \frac{f(x_d + \rho d^\gamma \xi) - f(x_d)}{q} - \frac{f(\overline{x}) - f(x_d)}{q} = \mathcal{O}(d^\gamma) \; . \end{aligned}$$

Finally, we find in all cases $p, \gamma \in \{1, 2\}$ the relation

$$||x_d - \overline{x}|| \leq ||x_d + \rho d^\gamma \xi - \overline{x}|| + ||\rho d^\gamma \xi|| = \mathcal{O}(d^{\gamma/p}) + \mathcal{O}(d^\gamma) = \mathcal{O}(d^{\gamma/p}) \; .$$

$\square$

Remark : In our analysis, for brevity, we did not consider equality constraints for the definition of Y (and/or F). If for example the index set Y is defined by

$$Y = \{y \in \mathbb{R}^m \mid u_j(y) = 0, \; j \in J, \quad v_i(y) \leq 0, \; i \in I\} \; ,$$

then, all results of this section remain valid if the condition MFCQ in (8) is modified as follows: For any $\overline{y} \in Y$ the vectors $Du_j(\overline{y}), \quad j \in J$ are linearly independent and there exists a vector $\overline{\eta} = \eta(\overline{y})$ such hat

$$Du_j(\overline{y}) \, \overline{\eta} = 0, \quad j \in J, \qquad Dv_i(\overline{y}) \, \overline{\eta} < 0, \quad i \in I(\overline{y}) \; .$$

A generalization of the results in this paper to the case of vector-valued constraints g is straightforward as well.

Remark : Suppose we use an exchange method for solving SIP (see e.g. [8, Sect.7.1]). In such an approach the active point set $Y_0(\overline{x})$ is approximately computed in each step. So, in this method it is sufficient to choose the gridpoints only in the neighborhood of the approximation of $Y_0(\overline{x})$ in each step.

Remark : As a special case, the following result is covered by the analysis above. Consider for $\varepsilon > 0$ the problem

$$\text{SIP}_\varepsilon: \quad \min\ f(x) \qquad \text{s.t.}\ \ x \in F_\varepsilon = \{x \in \mathbb{R}^n \mid g(x,y) \le \varepsilon,\ y \in Y\}\ .$$

This means, that the constraints only have to be satisfied approximately. Assume that $\overline{x}$ is a local minimizer of SIP of order $p = 1$ or $p = 2$ and that, for each $\varepsilon > 0$, x_ε is a local solution of SIP(Y_ε) such that $x_\varepsilon \to \overline{x}$ for $\varepsilon \to 0$. Then with some $\sigma > 0$ the following rate holds:

$$||x_\varepsilon - \overline{x}|| \le \sigma\ \varepsilon^{1/p}$$

4 Semi-infinite problems with variable index sets

In this section we extend the previous results to the case of semi-infinite problems of the form

$$\text{GSIP:} \quad \min\ f(x) \qquad \text{s.t.}\ \ x \in F = \{x \in \mathbb{R}^n \mid g(x,y) \le 0,\ y \in Y(x)\}$$

with index sets $Y(x) \subset \mathbb{R}^m$ depending on x. Such programs are also called *generalized semi-infinite problems.*

These semi-infinite problems with variable index sets may have a much more complicated structure than ordinary SIP. For example, the feasible set F of GSIP does not need to be closed if the set-valued mapping $Y(x)$ is not continuous. So, in the following we will assume: The mapping Y is continuous on $\mathbb{R}^n$ and satisfies with some compact set $C_0 \subset \mathbb{R}^m$:

$$Y\ :\ \mathbb{R}^n \to 2^{\mathbb{R}^r}, \quad Y(x) \subset C_0,\ \text{for all}\ x \in \mathbb{R}^n\ . \tag{14}$$

For a discussion of GSIP we refer e.g. to [16] and [18].

Discretization methods for GSIP have been considered in [19]. Similar as for SIP let, for each $d > 0$, be chosen a finite 'grid' $Y_d(x) \subset Y(x),\ x \in \mathbb{R}^n$, and let GSIP$(Y_d)$ denote the finite problem with the index-set $Y_d(x)$ and corresponding feasible set $F(Y_d)$. To obtain a closed feasible set $F(Y_d)$ also the grid $Y_d(x)$ must be continuous in x (see [19] for details). So, we assume, that the grid is given by

$$Y_d(x) = \{y_d^j(x),\ j = 1, \ldots, j_d\}, \quad x \in \mathbb{R}^n\ ,$$

with continuous functions $y_d^j : \mathbb{R}^n \to C_0$, $y_d^j(x) \in Y(x)$.

As before we assume that the following situation always holds true.

A1$_G$: Let $\overline{x}$ be a local minimizer of GSIP and let, for each $d > 0$, x_d be a local solution of GSIP(Y_d) such that

$$||x_d - \overline{x}|| \to 0 \qquad \text{for}\ \ d \to 0\ .$$

Here we assume that $N_{\overline{x}} \subset \mathbb{R}^n$ is a (fixed) compact set with $\overline{x}$ as inner point and d is defined by

$$d = \max_{x \in N_{\overline{x}}} \operatorname{dist}(Y_d(x), Y(x)) \ .$$

Again interested in the rate of convergence of the error $||x_d - \overline{x}||$, we are going to generalize the results of Section 3 to semi-infinite problems with variable index sets. Firstly we extend Lemma 1.

Lemma 4 *Let g be Lipschitz continuous near $\overline{x}$ w.r.t. C_0 with Lipschitz constant $L > 0$. Then there exists $\delta > 0$ such that for all d, $0 < d < \delta$:*

$$g(x_d, y) \leq L\, d \qquad \textit{for all} \ \ y \in Y(x_d) \ .$$

PROOF : Let y_d be a solution of $\max_{y \in Y(x_d)} \ g(x_d, y)$ and let $\hat{y}_d$ be a point in $Y_d(x_d)$ such that $||\hat{y}_d - y_d|| \leq d$. By Lipschitz continuity of g we find $g(x_d, y_d) \leq L\, d$ as in the proof of Lemma 1. □

The assumption A2 has to be modified.

A2$_G$: Let the following hold:

(α) The condition as in A2(α) is satisfied with Y replaced by C_0.

(β) The index set $Y(x)$ is given by

$$Y(x) = \{y \in \mathbb{R}^m \mid v_i(x, y) \leq 0, \ i \in I\} \ ,$$

where I is a finite index set and $v_i \in C^2(\mathbb{R}^n \times C_0)$.

(γ) MFCQ holds for $Y(\overline{x})$: For any $\overline{y} \in Y(\overline{x})$ there exists a vector $\overline{\eta} = \eta(\overline{y})$ such that

$$D_y v_i(\overline{x}, \overline{y})\, \overline{\eta} < 0, \quad i \in I(\overline{x}, \overline{y}) := \{i \in I \mid v_i(\overline{x}, \overline{y}) = 0\} \ . \tag{15}$$

LICQ is modified accordingly.

It is not difficult to show that MFCQ (and LICQ) are stable in the following sense (we omit the proof).

Lemma 5 *Let MFCQ hold for $Y(\overline{x})$. Then there exists $\varepsilon_0 > 0$ such that for any $\hat{x}$ with $||\hat{x} - \overline{x}|| \leq \varepsilon_0$ and any $\hat{y} \in Y(\hat{x})$ there is a vector $\hat{\eta}$ satisfying*

$$D_y v_i(\hat{x}, \hat{y})\, \hat{\eta} < 0, \quad i \in I(\hat{x}, \hat{y}) \ . \tag{16}$$

Note, that the condition MFCQ in $\mathrm{A2}_G$ (γ) is a natural condition implying that locally near $\overline{x}$ the index set $Y(x)$ is continuous.

Again, we consider for fixed x under assumptions $\mathrm{A2}_G$ the parametric problem

$$Q(x) \;:\quad \max_y \; g(x,y) \qquad \text{s.t.} \;\; v_i(x,y) \le 0, \; i \in I \; . \tag{17}$$

Under MFCQ at a solution $\overline{y}$ of $Q(x)$ the Kuhn-Tucker condition must hold:

$$D_y g(x,\overline{y}) - \sum_{i \in I(x,\overline{y})} \mu_i D_y v_i(x,\overline{y}) = 0 \quad \text{ with } \quad \mu_i \ge 0 \; .$$

Lemma 6 *Let the assumption $A2_G$ hold. Let $\varepsilon_0 > 0$ be as in Lemma 5. Then, there exists $\alpha > 0$ such that for any x, $||x - \overline{x}|| \le \varepsilon_0$, $y \in Y(x)$ and $\mu \ge 0$ satisfying*

$$D_y g(x,y) - \sum_{i \in I(x,y)} \mu_i D_y v_i(x,y) = 0$$

the bound $||\mu|| \le \alpha$ is true.

PROOF : Suppose on the contrary, that there are sequences $\{x_k\}$, $\{y_k\}$, $\{\mu_k\}$, $||x_k - \overline{x}|| \le \varepsilon_0$, $y_k \in Y(x_k)$ and $\mu_k \ge 0$ such that

$$D_y g(x_k,y_k) - \sum_{i \in I(x_k,y_k)} (\mu_k)_i D_y v_i(x_k,y_k) = 0$$

and $||\mu_k|| \to \infty$ for $k \to \infty$. Then as in the proof of Lemma 2, for accumulation points $\hat{x}$, $||\hat{x} - \overline{x}|| \le \varepsilon_0$, $\hat{y} \in Y(\hat{x})$, $\hat{\mu}_i \ge 0$, $i \in I_0$ and $I_0 \subset I(\hat{x},\hat{y})$ we find

$$-\sum_{i \in I_0} \hat{\mu}_i D_y v_i(\hat{x},\hat{y}) = 0$$

with $\hat{\mu} \ne 0$ and by multiplying this relation with the MFCQ-vector $\hat{\eta}$ in (16) we obtain a contradiction. $\square$

For $I_l \subset I$ we again can define the subsets of $Y(x)$,

$$S_l(x) := \{y \in Y(x) \mid v_i(x,y) = 0, \; i \in I_l\}, \quad x \in \mathbb{R}^n \;, \quad l = 1,\ldots,k \;,$$

such that $S_l(x) \cap Y(x) \ne \emptyset$, which now are depending on x.

$\mathbf{A3}_G$ **:** Let the sets $S_l(x)$, $S_l(x) \cap Y(x) \ne \emptyset$, $l = 1,\ldots,k$, be as defined above. Let the grids Y_d be chosen such that for all $l = 1,\ldots,k$, and all d with the set $N_{\overline{x}}$ in $\mathrm{A1}_G$ the following relation holds:

$$\max_{y \in S_l(x)} \; \min_{\hat{y} \in Y_d(x) \cap S_l(x)} ||\hat{y} - y|| \le d \qquad \text{for all } \; x \in N_{\overline{x}} \; .$$

Theorem 3 *Let $A2_G$ and $A3_G$ hold. Then with the local solutions x_d of GSIP(Y_d) for some $L > 0$ the relation*

$$\max_{y \in Y(x_d)} g(x_d, y) \leq L\, d^2 .$$

is satisfied.

PROOF : The proof is a straightforward modification of the proof of Theorem 1. □

Concerning a result as in Lemma 3, the situation is more complicated for semi-infinite problems with variable index sets. To prove such a result we have to strengthen the assumptions. We will assume that the so-called *reduction approach* is possible near $\overline{x}$ (see e.g. [9], [4], [18]).

Note that by definition of the active index set $Y_0(\overline{x}) = \{\overline{y} \in Y(\overline{x}) \mid g(\overline{x}, \overline{y}) = 0\}$, for feasible $\overline{x}$, any point $\overline{y} \in Y_0(\overline{x})$ is a global solution of the finite optimization problem $Q(\overline{x})$ in (17).

We make the following assumptions.

$\mathbf{A}_{\text{red}}$**:** Let the functions g, v_i be twice continuously differentiable on $\overline{U} \times C_0$, where $\overline{U}$ is a neighborhood of $\overline{x}$. Suppose for any $\overline{y} \in Y_0(\overline{x})$ (i.e. a solution $\overline{y}$ of $Q(\overline{x})$) the following is fulfilled:

(α) LICQ: $D_y v_i(\overline{x}, \overline{y}),\quad i \in I(\overline{x}, \overline{y})$ are linearly independent.

(β) Kuhn-Tucker condition

$$D_y g(\overline{x}, \overline{y}) - \sum_{i \in I(\overline{x}, \overline{y})} \overline{\mu}_i D_y v_i(\overline{x}, \overline{y}).$$

with $\overline{\mu}_i > 0,\ i \in I(\overline{x}, \overline{y})$, (*strict complementary slackness*).

(δ) A certain second order sufficiency condition on the Hessian of the Lagrange function (see e.g. [18]).

It can be proven, that under these conditions A_{red} the set $Y_0(\overline{x})$ is finite, $Y_0(\overline{x}) = \{\overline{y}_1, \ldots, \overline{y}_r\}$, and that there is a neighborhood $\overline{U}$ of $\overline{x}$ such that with local maximizer functions $y_j(x) \in C^1(\overline{U}),\ y_j(\overline{x}) = \overline{y}_j,\ j = 1, \ldots, r$, of $Q(x)$, locally on $\overline{U}$, GSIP is equivalent with the smooth finite problem (*reduced problem*):

$$\text{GSIP}_{\text{loc}}(\overline{x}) : \quad \min\ f(x) \quad \text{s.t.} \quad g_j(x) := g(x, y_j(x)) \geq 0, \quad j = 1, \ldots, r.$$

For the functions g_j the relations

$$Dg_j(\overline{x}) = D_x g(\overline{x}, \overline{y}_j) - \sum_{i \in I(\overline{x}, \overline{y}_j)} \overline{\mu}_i^j D_x v_i(\overline{x}, \overline{y}_j).$$

hold with the Lagrange-functions corresponding to the maximizers $\overline{y}_j \in Y_0(\overline{x})$, $j = 1, \ldots, r$.

The MFCQ has now to be modified: We say that MFCQ$_G$ is valid for $\overline{x}$ if there exists a vector $\xi \in \mathbb{R}^m$ such that

$$Dg_j(\overline{x})\, \xi \leq -1, \quad \text{for all} \;\; j = 1, \ldots, r \; . \tag{18}$$

Lemma 7 *Let A_{red}, MFCQ$_G$ be valid at the local minimizer $\overline{x}$ of GSIP and let with some $\gamma > 0$, $c_1 > 0$ the following relation hold:*

$$\max_{y \in Y(x_d)} \; g(x_d, y) \leq c_1 d^\gamma \; , \qquad \textit{for sufficiently small } d \; .$$

Then there exists $\rho > 0$ such that with the MFCQ-vector ξ in (18), for small d, the points $\hat{x}_d := x_d + \rho d^\gamma \xi$ are feasible for GSIP, i.e.

$$g(\hat{x}_d, y) \leq 0 \; , \quad y \in Y(\hat{x}_d) \; .$$

PROOF : By the reduction assumption A_{red}, for x near $\overline{x}$, given any $\sigma \in \mathbb{R}$ the following is true:

$$g(x, y) \leq \sigma \; , \quad y \in Y(x) \quad \text{is equivalent with} \quad g_j(x) \leq \sigma, \; j = 1, \ldots, r \; .$$

Using the continuity of Dg_j, by MFCQ$_G$ there exists $\varepsilon > 0$ such that

$$Dg_j(x)\xi \leq -\frac{1}{2} \quad \text{for} \;\; j = 1, \ldots, p \; , \; x, \; ||x - \overline{x}|| < \varepsilon \; .$$

Now we find for all $j = 1, \ldots, r$,

$$\begin{aligned} g_j(\hat{x}_d) &= g_j(x_d) + \rho d^\gamma Dg(x_d)\xi + o(\rho d^\gamma) \\ &\leq c_1 d^\gamma - \rho\frac{1}{2} d^\gamma + o(\rho d^\gamma) \\ &= d^\gamma(c_1 - \rho\frac{1}{2}) + o(\rho d^\gamma) \leq 0 \end{aligned} \tag{19}$$

if we choose $\rho > 2c_1$ and d small enough. This shows that $\hat{x}_d$ is feasible. □

The generalization of Corollary 1 and Theorem 2 is again straightforward. We only give

Theorem 4 *Assume that A_{red}, MFCQ$_G$ hold and that f is Lipschitz continuous at $\overline{x}$. Then if $\overline{x}$ is a strict local minimizer of order $p = 1$ or $p = 2$ the following is true:*

(a) There is some $\sigma_2 > 0$ such that for all d small enough,

$$||x_d - \overline{x}|| \leq \sigma_1 \; d^{1/p} \; .$$

(b) If in addition A2$_G$ and A3$_G$ are satisfied for Y, Y_d, then with some $\sigma_2 > 0$ the inequality

$$||x_d - \overline{x}|| \leq \sigma_2 \; d^{2/p}$$

is valid.

References

[1] Cheney E.W., *Introduction to approximation theory*, Chelsea, New York, 2nd edition (1982).

[2] Dunham C.B., *Efficiency of Chebyshev approximation on finite subsets*, J. Assoc. Comp. Mach. 21, No. 2, 311-313, (1974).

[3] Dunham C.B., Williams J., *Rate of convergence of discretization in Chebyshev approximation*, Math. of Computation 37, No. 155, 135-139, (1981).

[4] Gramlich G., Hettich R., Sachs E.W., *Local convergence of SQP-methods in semi-infinite programming*, SIAM J. Optimiz. 5, No. 3, 641-658, (1995).

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

[6] Hettich R., An implementation of a discretization method for semi-infinite programming, Math. Programming 34, 354-361, (1986).

[7] Hettich R., Gramlich G., A note on an implementation of a method for quadratic semi-infinite programming, Math. Programming 46, (1990), 249-254, (1990).

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

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

[10] Lawrence C.T., Tits A.L., *Feasible sequential quadratic programming for finely discretized problems for SIP*, in 'Semi-infinite Programming', Reemtsen/Rückmann (eds.), Kluwer, Boston, p. 159 -193 (1998).

[11] Opfer G., *Solving complex approximation problems by semi-infinite optimization techniques: A study on convergence*, Num. Math. 39, 411-420, (1982).

[12] Panier E.R., Tits A.L., *A globally convergent algorithm with adaptively refined discretization for semi-infinite optimization problems arising in engineering design*, IEEE Trans. Automat. Contr., AC-34, 903-908, (1989).

[13] Polak E., *Optimization, Algorithms and consistent Approximations*, Springer, Berlin (1997).

[14] Reemtsen R., *Some outer approximation methods for semi-infinite optimization problems*, J. Comp. Appl. Math. 53, 87-108, (1994)

[15] Reemtsen R., Görner S., *Numerical methods for semi-infinite programming: A survey*, in 'Semi-infinite Programming', Reemtsen/Rückmann (eds.), Kluwer, Boston, p. 195 -262 (1998).

[16] Rückmann J.-J., Stein O., *On linear and linearized generalized semi-infinite problems*, Annals of Operations Research, to appear.

[17] Still G. *Vector-valued approximation problems; Numerical and genericity aspects*, Proceedings of the 3rd International Conference on Approximation and Optimization (Puebla, 1995), (16pp. electronic), Benemerita Univ. Auton. Puebla, Puebla, (1997).

[18] Still G., *Generalized semi-infinite programming: Theory and methods*, European Journal of Operational Research 119, 301-313, (1999).

[19] Still G., *Generalized semi-infinite programming: Numerical aspects*, Optimization, to appear

[20] Tanaka Y., *A trust region method for semi-infinite programming problems*, Intern. J. Systems Sci. 30, 199-204, (1999).