SIAM REVIEW
Vol. 35, No. 3, pp. 380–429, September 1993

SEMI-INFINITE PROGRAMMING: THEORY, METHODS, AND APPLICATIONS*

R. HETTICH† AND K. O. KORTANEK‡

Abstract. Starting from a number of motivating and abundant applications in §2, including control of robots, eigenvalue computations, mechanical stress of materials, and statistical design, the authors describe a class of optimization problems which are referred to as semi-infinite, because their constraints bound functions of a finite number of variables on a whole region. In §§3–5, first- and second-order optimality conditions are derived for general nonlinear problems as well as a procedure for reducing the problem locally to one with only finitely many constraints. Another main effort for achieving simplification is through duality in §6. There, algebraic properties of finite linear programming are brought to bear on duality theory in semi-infinite programming. Section 7 treats numerical methods based on either discretization or local reduction with the emphasis on the design of superlinearly convergent (SQP-type) methods. Taking this differentiable point of view, this paper can be considered to be complementary to the review given by Polak [*SIAM Rev.*, 29 (1987), pp. 21–89] on the nondifferentiable approach. The last, short section briefly reviews some work done on parametric problems.

Key words. semi-infinite programming, mathematical programming and optimization techniques, optimality conditions, duality theory, nonlinear programming, parametric analysis, algorithms for functional approximation

AMS subject classifications. 90C, 49D39, 65K05, 90C30, 65D15

1. Introduction. In its most common statement the optimization problem we study in this paper seeks the maximum of functional values $F(z)$ subject to a system of constraints on z, expressed as $g(z, t) \leq 0$ for all t in some set B. The theoretical and practical manifestations of this mathematical model are abundant and suggest various mathematical properties for the functions F and g, as well as for the set B. And, of course, g itself could arise in the form of a finite list of functions.

The term "semi-infinite programming" (SIP) derives from the property that z denotes finitely many variables while B is an infinite set, for example, in many applications closed and bounded. In any case, finitely many variables appear in infinitely many constraints.

Considering the prototype problem above, we believe that a brief review of first-order optimality conditions is the most natural way to begin the technical part of the survey. We do this in §3.

In §4 we discuss possibilities for reducing the intrinsically infinite problem (with infinitely many constraints) to a finitely constrained one, so that widespread results and algorithms of common mathematical programming can be applied. The most obvious way to do this is to replace B with a finite subset or—in a more sophisticated way—a sequence of successively refined grids. When proceeding this way, in general, only finite approximations of the original infinite problem are available. Much more productive is a second procedure, where the infinitely many constraints $g(z, t) \leq 0, t \in B$, are replaced by constraints $g(z, t^l(z)) \leq 0$, with $t^l(z)$ denoting local solutions of the parametric problem in parameter z, $\max\{g(z, t) | t \in B\}$. Using results from parametric programming, assumptions can be given which ensure that (locally) the problem is equivalent to a common finite nonlinear problem with smooth constraints. Both approaches lead to their own classes of algorithms (cf. §7).

After having made these connections between the infinite problem and an equivalent finite problem in §§3 and 4, we can then effectively extend various fruitful second-order sufficiency conditions for finite nonlinear programming to the semi-infinite programming case. This is

*Received by the editors April 8, 1992; accepted for publication (in revised form) October14, 1992.

†University of Trier, FB IV (Mathematics), P.O. Box 3825, 5500 Trier, Germany.

‡Department of Management Sciences, College of Business Administration, University of Iowa, Iowa City, Iowa 52242.

done in §5, where a new sufficient optimality result is proved, which is a more direct approach than the usual ones and requires less restrictive regularity assumptions.

We have not forgotten the purpose of our first nonintroductory section, namely, §2. To the contrary, in some ways it is the most important section because of its motivation toward real problems. We say a few words about this next.

The reader will note immediately the attractive and highly nonlinear semi-infinite problems presented first in §2, namely, path planning problems in robotics as well as a class of vibrating membrane problems. These problems have been treated numerically by us and our colleagues over the last five to eight years, as well as other researchers. We hope to provide convincing results that the numerical approaches reviewed in this paper are promising, and that they provide state of the art numerical results.

But there is another emphasis provided by additional problems presented in §2. If the F and g functions of our prototype problem above are relaxed in another direction, namely, algebraic rather than topological, we encounter one of the earliest formulations of this entire class of optimization problems, namely, linear semi-infinite programming, whose early treatment by means of generalized finite sequences dates back to 1962. Roughly speaking, such sequences are merely functions from one set to another, having only finitely many nonzero images, i.e., a special linear space over the real numbers. On the one hand, generalized finite sequences have played a key role in the duality development of linear SIP, which is studied in §6. But on the other hand, one of the main points of some examples in §2 is that modeling can certainly begin with their use in a particular physical or social science situation. We present example problems that emphasize this point from the fields of air pollution abatement and variational principles in continuum mechanics, principally involving plane strain and plane stress problems. In these modeling tasks generalized finite sequences have their own meaningful interpretations, and are not simply a collection of formal Lagrange multipliers for an infinite-dimensional problem.

Finally, Example 2.6 exhibits the interplay between systems of inequalities and optimization over spaces of probability measures. For other interesting relations to statistics, see [2.67].

In §6 we continue with a technical treatment of duality theory for linear SIP. The reader will observe that actually a linear SIP is a convex programming problem with the feasible region defined by supporting halfspaces. This raises the question whether a separate duality for SIP is not simply derived from convex programming theory as developed extensively (see [1.10], [1.11], for instance). However, keeping computational applications in mind, the powerful Lagrangian duality requires knowledge of finitely many smooth, convex functions describing the region which are not available from the given SIP data. We observe that only in regular cases the locally reduced problem (Problem 4.9) gives such a description and only with implicitly defined convex functions; therefore, the need of a duality theory in terms of the SIP data was evident from the outset.

In §6.1 the dual (D) in the space of generalized finite sequences is derived and related to other formulations exhibiting the relation to moment problems, which appear to be equivalent to our dual.

In §6.2 we discuss the stable (superconsistent) case where solvability and strong duality is preserved under perturbations of the data.

When certain properties of the problem function data are present, then SIP duality is closely related to the duality theory of ordinary finite linear programming. To distinguish this case from what in general might be encountered, the term "perfect duality" was introduced for convenience. This occurred in 1972 in one of Richard J. Duffin's papers as well as in conversation with him, although already in the 1960s many recognized these useful properties of perfect duality. We present this case in §6.3.

In contrast to this case one must address "nonperfect" duality, and so we next describe in §6.4 some of its properties, in particular the existence of gaps between primal and dual program values. We review an approach for resolving duality gaps by what has been termed as a fundamental theorem in linear SIP: if an infinite system of linear inequalities over the reals is inconsistent but every finite system is consistent, then the original infinite system has a solution in an extended number system. We briefly review some of the consequences of this result for duality.

In §7 we review those classes of methods that are based on reformulations or approximations of the problem in terms of problems with finitely many differentiable constraints. For methods based on a nondifferentiable formulation the reader is referred to the paper of Polak [7.43]. In §7.1 we briefly describe exchange methods and sketch the connection to the well-known cutting plane methods from convex programming and to the Remes algorithms in Chebyshev approximation. In §7.2 discretization methods are considered that consist not simply in the replacement of the infinitely many constraints by a finite subset of uncorrelated constraints, but "remember" the continuous heritage in grid selection strategies. Finally in §7.3 we consider superlinearly convergent methods based on the locally reduced problem. The special emphasis here is on Sequential Quadratic Programing (SQP) methods to be applied to this reduced problem.

Our final section, §8, is a brief review of parametric problems with a focus not so much on breadth but on procedures which extend ideas and techniques from finite programming to SIP.

The list of references (which is far from complete) shows a growing interest in SIP during the last decade. This certainly is due to the various interesting applications. A monograph on linear programming with special emphasis on SIP is available in English [1.5], whereas a textbook on nonlinear SIP is currently only available in German [1.8]. In addition there are three conference volumes [1.2], [1.4], [1.7] and review papers [1.6], [7.43]. A thorough study of the connections of SIP to approximation theory can also be found in [1.9]. For various theoretical and algorithmic aspects of SIP in a more general framework see [1.3].

We remark that there are often textbooks where one elementary result is found, and then other books where yet other results of a related type are found. As almost all of the elementary material on semi-infinite programming can be found in the (German) textbook [1.8] (but not in an English text), it appeared most convenient to use this as a common reference. For all advanced, less elementary results, however, we have tried to refer to the original work or to more easily accessible English texts.

2. SIP models in applications. Numerous models in the physical and social sciences require the consideration of constraints on the state or the control of the system during a whole period of time or in every point of a geometric region, i.e., functional inequality constraints which are characteristic for semi-infinite programming. For a number of interesting examples from engineering design we refer to the review paper of Polak [7.43]; see also [2.58], [2.60], [2.61], [2.62], [2.83], [2.86]. Later on (§6) we will see that the linear SIP has a dual which is just a standard problem in moment theory. This relation, which gives rise to a number of interesting applications (see [2.10], [2.12], [2.13], [2.15], [2.22]–[2.25], [2.29], [2.42], [2.43], [2.63], [2.79], [2.84], [2.88]), has not yet been exploited to its full extent. For the future we expect to encounter a wide field for further research also with respect to possible generalizations of SIP like, for instance, separably infinite programs [6.15], [6.16].

In the sequel we give a selection of problems treated more deeply by at least one of the authors during recent years, mostly in collaboration with others. This selection already demonstrates a broad range of applications. In the list of references we have included a number of papers on problems which are out of the scope of this review [2.26], [2.27], [2.40], [2.41], [2.44], [2.50], [2.51], [2.56], [2.59], [2.70], [2.71], [2.81], [2.82], [6.28], [6.29].

Example 2.1 (*robot trajectory planning*). The control of robots provides a wide field of illustrative examples:

—to avoid obstacles, each point of a robot arm has to stay out of some region all the time (cf. [7.43]);

—at every moment the acceleration supplied by a driving motor is restricted.

Let us describe a model given in [2.52] which will also be used later on (§7). Suppose we are given a robot, the position of which is described by controllable coordinates $\Theta_1, \ldots, \Theta_R$ (angles in joints, lengths of links, etc.). Assume that a path $\Theta(\tau)$, $\tau \in [0, 1](\Theta = (\Theta_1, \ldots, \Theta_R)^T)$ is given. Then a reparametrization $t = h(\tau)$, $\tau = h^{-1}(t)$ to time t is sought such that $\underset{\sim}{\Theta}(t) := \Theta(h^{-1}(t))$ is an executable movement and such that the performance time $T = h(1)$ is minimized.

Marin [2.52] restricts functions h to those given by

$$h(\tau) = \int_0^\tau g(s)ds,$$

where

$$g(s) = \sum_{j=1}^{n} z_j B_j(s)$$

is a cubic spline on an equidistant set of $n - 4$ knots in $(0, 1)(\{B_j\}$, for instance, a B-spline basis) and $g(s) > 0$, $s \in [0, 1]$.

In the most simple case $\underset{\sim}{\Theta}(t)$ is considered to be an executable movement if for given constants c_{ij}, for joints $i = 1, \ldots, R$, we have

$$\left| \frac{d^j \underset{\sim}{\Theta}_i(t)}{dt^j} \right| \leq c_{ij} \quad \text{for } j = 1, 2, 3,$$

i.e., velocities, accelerations, and jerks are bounded. More realistically one could consider the torque constraints derived from the dynamic equations. In this setting the reparametrization problem is of the following type:

Minimize a linear function $c^T z$ subject to constraints

$$g_l^{i,j}(z, \tau) \leq 0, \quad \tau \in [0, 1], \quad i = 1, \ldots, R; \quad j = 1, 2, 3; \quad l = 1, 2,$$

where $g_l^{i,j}$ depend nonlinearly on z for $j = 2, 3$. Thus, we obtain a nonlinear semi-infinite problem with $6R$ functional constraints on [0, 1] and $4R$ of these nonlinear. See also [2.14], [2.30], [2.32], [2.36] for this problem, where different numerical approaches are considered. In §7 we will give a numerical example.

Another problem from robotics (the so-called maneuverability problem) leads to a generalized type of SIP in that the region B also depends on z (cf. [2.21]).

Example 2.2 (*vibrating membrane problem*). Chebyshev approximation and its applications (see, for instance, [1.8], [1.9], [2.8], [2.9], [2.18], [2.54], [2.55], [2.76], [2.77], [2.78], [2.80], [2.87], [2.89]) provide another well-known and broad class of semi-infinite programming problems. Examples include defect minimization methods for solving operator equations. In [2.1], [2.11], [2.17], [2.39], [2.45], [2.64], [2.65], [2.66], [2.69] (free) boundary value problems, and especially the Stefan problem and its inverse are treated successfully in this way. Other examples include eigenvalue problems for elliptic differential operators

[2.31], [2.33]–[2.35], [2.72], [2.73], [8.17]. As the latter problem in a natural way leads to a parametric SIP, we give the idea of this approach.

Consider the problem of finding $\lambda \in \mathbb{R}$, $u \in C^2(B) \cap C(B \cup \Gamma)$ such that

$$\Delta u + \lambda u = 0 \quad \text{in } B, \tag{2.1}$$

$$u = 0 \quad \text{on } \Gamma, \tag{2.2}$$

with $B \subset \mathbb{R}^2$ an open connected region with boundary Γ. For given λ, let $V(\lambda) := \{v(\lambda; z, \bullet) = \sum_{i=1}^n z_i\varphi_i(\lambda, \bullet)\}$ be a linear space of dimension n, all elements of which satisfy the differential equation.

Minimizing the defect on the boundary amounts to solving the linear semi-infinite problem

$$\text{SIP}(\lambda) \qquad \text{Minimize } \mu \text{ subject to } |v(\lambda; z, x)| \le \mu, \quad x \in B, \quad \rho(v(\lambda; z, \bullet)) = 1$$

with some (linear) normalizing constraint $\rho(v) = 1$ to exclude trivial solutions. From a solution $\mu(\lambda)$, $z(\lambda)$ of SIP(λ) an error bound $\varepsilon(\lambda)$ can be computed such that the existence of an eigenvalue λ_k of (2.1), (2.2) is guaranteed to satisfy

$$\frac{|\lambda_k - \lambda|}{|\lambda|} < \varepsilon(\lambda).$$

Now, approximations $\tilde{\lambda}_k$ to λ_k are computed as local minima of $\varepsilon(\lambda)$. So, this method requires the numerical treatment of a one-parameter family SIP(λ) of linear semi-infinite problems.

Remark. Much computational work via the SIP approach has been done on this problem, principally based in Trier. Computational results are also reported in [2.49], performed in collaboration with the first author during his visit to Carnegie Mellon University in 1986.

In [1.8], [8.17], [2.37] another example for a parametric SIP may be found: a method for general rational Chebyshev approximation which is closely related to the Differential Correction Algorithm, leading to the problem of determining λ^* such that the value $v(\lambda)$ of a linear parametric problem SIP(λ) becomes zero.

Example 2.3 (*minimal cost control of air pollution*). Assume that there are n sources of a chemically inert pollutant in a region, denoted by B, whose emission rates must be controlled so that the annual mean ground level pollution concentration at each point in the region satisfies some standard. A control policy is sought whose total cost is minimal; see [2.19], [2.20], [2.22], [2.28].

Let $a_j(t)$ denote the annual mean concentration at point t in B due to source j before control, and let $\psi(t)$ denote the maximum concentration of pollutant permitted at t. Let z_j denote the fraction of reduction for source j, bounded above by a constant $u_j \le 1$. Due to the assumption of chemical inertia, the concentration after control is $\sum_{j=1}^n (1 - z_j)a_j(t)$ and it must not exceed the standard $\psi(t)$. The cost of achieving z_j is denoted $C^j(z_j)$. An optimal control policy program in case of linear cost $C_j(z_j) = c_j z_j$ is the following program (P):

$$v(P) = \max \left\{ -\sum_{j=1}^n c_j z_j | z \in Z^P \right\} \tag{P}$$

with the feasible set $Z^P \subset \mathbb{R}^n$ defined by the inequalities

$$-\sum_{j=1}^n z_j a_j(t) \le \psi(t) - \sum_{j=1}^n a_j(t) =: a_{n+1}(t) \quad \text{for all } t \in B \tag{2.3}$$

and

$$0 \le z_j \le u_j, \qquad j = 1, \dots, n. \tag{2.4}$$

Assuming B is compact, a_j, ψ continuously differentiable on B, and the existence of a Slater point z^0 (for which all inequalities (2.3), (2.4) hold strictly), (P) obviously has a solution z^*. It follows from duality theory (of Theorem 6.9 below) that the value $v(P)$ equals the value $v(D)$ of the following dual problem (D) which also has a solution, (μ^*, y^*).

$$\text{(D)} \qquad v(D) = \min \left\{ \sum_{t \in B} a_{n+1}(t)\mu(t) + \sum_{j=1}^{n} y_j u_j \,|(\mu, y) \in Z^D \right\}$$

where Z^D is the set of pairs (μ, y), μ a nonnegative function with finite support supp(μ) on B denoted also as a generalized finite sequence (cf. [6.10]), $y \in \mathbb{R}$, $y \ge 0$, such that

$$\sum_{t \in B} a_j(t)\mu(t) - y_j \le c_j, \qquad j = 1, \dots, n.$$

Similar to linear programming theory $\mu^*(t)$ may be interpreted as the marginal value per unit reduction at point t in B given that the emission reductions z^* have been implemented. The dual program thereby provides objective means to evaluate air quality standards derived from other considerations. Analogously, if $y_j^* > 0$, one could argue that considerations other than economic criteria were used to set the maximum limits on admissible source reductions.

In [2.22] this problem is treated (with nonlinear $C_j(z_j)$) in terms of nondifferentiable optimization.

Example 2.4 (*a problem in the stress of materials*). This example illustrates an approach based on SIP for formulating problems involving the stress and deformation of materials (cf. [2.2]–[2.5], [2.48], [2.53], [2.74], [2.75]).

The problem involves a hollow circular disc

$$D = \left\{ \begin{pmatrix} r\cos\Theta \\ r\sin\Theta \end{pmatrix} | a \le r \le b, \Theta \in [0, 2\pi] \right\} \subset \mathbb{R}^2,$$

which rotates about the origin with constant angular velocity. This generates radial and angular stresses σ_{rad} and σ_{ang}, which may cause the material to rupture. Because of symmetry the stresses are functions only of r, $\sigma_{\text{rad}}(r)$, and $\sigma_{\text{ang}}(r)$. We may also assume that the radial displacement u is only a function of r and that the inner boundary is fixed, i.e., $u(a) = 0$.

Now the material experiences radial and angular strains, given, respectively, by

$$\varepsilon_{\text{rad}} = \frac{du}{dr}, \qquad \varepsilon_{\text{ang}} = \frac{u}{r}.$$

The internal work done is given by

$$\begin{aligned} W(\sigma, u) &= \iint_D (\sigma_{\text{rad}}\varepsilon_{\text{rad}} + \sigma_{\text{ang}}\varepsilon_{\text{ang}})dD \\ &= 2\pi \int_a^b [\sigma_{\text{rad}}\, r\, du + \sigma_{\text{ang}} u\, dr]. \end{aligned}$$

Body forces $f(r)$ pull toward the center of the disc, while there is an outgoing surface force g at the boundary $r = b$. The external work F is defined to be a function of u according to the sum of two integrals, the latter being a surface integral:

$$F(u) = \iint fu\, dA + \int gu\, ds = 2\pi \left(\int_a^b rf(r)u(r)dr + bgu(b) \right).$$

The static principle of limit analysis seeks the largest multiplier λ^* which may be placed on the external work so that there will be an equilibrium stress vector σ, as characterized by the following infinite system of equations:

$$W(\sigma, u) = \lambda F(u) \quad \text{for all } u.$$

Usually the stress vector σ is further constrained, and we shall require

$$\sigma(r) = \begin{bmatrix} \sigma_{\text{rad}}(r) \\ \sigma_{\text{ang}}(r) \end{bmatrix}$$

to lie in the convex Tresca yield set

$$T = \left\{ \sigma = \begin{pmatrix} \sigma_1 \\ \sigma_2 \end{pmatrix} \middle| \; |\sigma_i| \le k\sqrt{3}, i = 1, 2, |\sigma_1 - \sigma_2| \le k\sqrt{3} \right\},$$

where k is a parameter depending on the yield strength of the material; see [2.5]. It is known that the static principle is equivalent to the following biextremal problem:

$$\lambda_* = \inf_u \left\{ \sup_{\sigma \varepsilon T} W(\sigma, u) | \text{ subject to } F(u) = 1 \right\}.$$

This obviously may also be written as a semi-infinite problem:

$$\text{(P)} \qquad \begin{aligned} v(P) = \min \delta \quad & \text{subject to} \\ W(\sigma, u) \le \delta \quad & \text{for all } \sigma \in T, F(u) = 1. \end{aligned}$$

In [2.47], [2.48] a generalized finite sequence space is introduced for the displacements u and generalized stress coordinates for σ_r and σ_Θ. This transforms the biextremal problem into one where the generalized finite sequences and the stress coordinates are the new variables. Then, using a construction borrowed from constrained two-person game theory, the biextremal problem is decomposed into an equivalent dual pair of uniextremal problems. In one problem, the stress variables are the principle variables, but they are augmented by dual variables associated with constraints involving the displacements. In the dual, uniextremal problem (which employs variables which are completely different) the displacement variables are augmented with dual variables associated with the convexity constraints of the stress variables.

Remark 2.4. Recent work in plane strain plastic limit analysis yields a dual pair of linear semi-infinite programs, each having a specific physical interpretation. Computations so far have been done for finite element approximations and "strategic" discretizations of the unbounded convex set of stresses. Numerical experiments on the associated finite dual LP's are reported in [2.6], [2.7].

Example 2.5 (*splines and generalized finite sequences*). The next example reviews a concept for modeling problems in terms of SIP dual variables, i.e., generalized finite sequences (cf. §6), rather than an explicit model.

Let η be a generalized finite sequence on the interval $[a, b]$, i.e., a function defined on $[a, b]$ with finite support. With η we may associate a step function

$$u(t) = \sum_{t' \in [a,t]} \eta(t').$$

(Note that the summation involves only finitely many nonzero numbers.)

Integration over $[a, b]$ gives

$$U(t) := \int_0^t u(t')dt' = \sum_{t' \in [a,t]} (t - t')\eta(t'),$$

as is seen by applying Riemann–Stieltjes integration by parts. Obviously this is an alternative description of the linear space of piecewise linear (continuous) splines with an arbitrary but finite number of free knots. This connection is rather straightforward [2.46]. The higher-dimensional Riemann–Stieltjes Integral and its properties are developed in [2.38].

Example 2.6 (*optimization over probability measures*). Let f_j, $j = 1, \ldots, n$ be continuous functions over a compact set B. Consider the vector of n^2 functions defined by

$$a(t) = \left(f_1^2(t), f_1(t)f_2(t), \ldots, f_i(t)f_j(t), \ldots, f_n^2(t)\right).$$

Later in §6 we shall review the role of the important convex cones, called "moment cones," from which we introduce a special case here, namely,

$$M_{n^2} = \text{co}\,\{a(t) | t \in B\}.$$

For applications in probability and statistics one is interested in the convex hull of $\{a(t), t \in B\}$, which can be expressed using generalized finite sequences on B, namely,

$$\Omega = \left\{ w = \sum_{t \in B} a(t)\mu(t) | \mu \in \mathbb{R}_+^B, \sum_{t \in B} \mu(t) = 1 \right\}.$$

It has been well known that Ω is a compact convex set; see [2.16], [2.57], [2.85]. Analogous to a result of Rogosinsky [6.50], which we shall mention in §6.1, Ω can be written in an equivalent way with nonnegative regular Borel measures on Ω. The reader may also recognize that the generators of Ω can be conveniently taken to be the rank 1 matrices, $f(t)f(t)^T$.

Analogous to a geometrical program which we introduce later in Lemma 6.1, an important class of problems in statistical regression experimental design can be formulated as follows:

$$\inf\{\sigma(w) | w \in \Omega\},$$

where σ is a convex function on Ω, which satisfies certain differential conditions, but where σ may be infinite on a subset of the boundary of Ω. For illustration, we assume that σ is continuously differentiable on an open set in $\mathbb{R}^{n^2}$ (or $\mathbb{R}^{n \times n}$ if matrix notation is used), which contains the subset of Ω over which σ is finite.

We illustrate a general procedure (see [2.23]) for deriving a linear semi-infinite inequality system which characterizes an optimal solution w^*. The procedure applies to the general case where σ may not be defined at w^*, as our simple example shows. We mention first that an intuitive way to characterize an optimal $w^* \in \Omega$ is to formulate a linear inequality system according to the following observation: No $w \in \Omega$ is made worse by moving slightly in the direction of w^*; see [2.23]. The merit of this idea is that $\nabla\sigma(w^*)$ need not be evaluated. In terms of directional derivatives, this optimality condition becomes

$$\frac{\partial}{\partial t}\sigma((1-t)w + tw^*)|_{t=0} \leq 0 \quad \text{for all } w \text{ for which } \sigma(w) \text{ exists, i.e.,} \tag{2.5}$$

$$\sum_{j=1}^{n}\sum_{i=1}^{n} \frac{\partial \sigma}{\partial w_{ij}}(w)\left(w_{ij}^* - w_{ij}\right) \leq 0 \quad \text{for all such } w, \tag{2.5'}$$

which is a linear system in w^*.

For the example, let $B = \{0\} \cup \{1\}$, $f_1(t) = 1$, $f_2(t) = t$. Then any w in Ω has the form

$$w = (1-p)\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + p\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & p \\ p & p \end{bmatrix}$$

for $0 \le p \le 1$.

Using the trace operator of a square matrix, the following σ is convex on the subset of nonsingular matrices of Ω:

$$\sigma(w) = \operatorname{tr}\left[\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix}^{-1}\right] = w_{22}\left(w_{11}w_{22} - w_{12}w_{21}\right)^{-1}.$$

Let $w = \begin{bmatrix} 1 & p \\ p & p \end{bmatrix}$. Then a simple differentiation yields

$$\nabla\sigma(w) = \left[\frac{-1}{(1-p)^2}, \frac{1}{(1-p)^2}, \frac{1}{(1-p)^2}, \frac{-1}{(1-p)^2}\right].$$

The optimal w^* is of the form

$$\begin{pmatrix} 1 & p^* \\ p^* & p^* \end{pmatrix}$$

and is determined by (2.5), i.e.,

$$(p^* - p)/(1-p)^2 \le 0 \quad \text{for all } p, 0 < p < 1.$$

Hence, $p^* = 0$, giving the optimal design

$$w^* = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},$$

for which σ is not defined.

3. First-order optimality conditions. In this section some definitions and theorems are reviewed which are basic for all subsequent sections. To most of the readers all of these will be rather familiar. We assume for convenience that in our problem

$$\text{(P)} \qquad \begin{aligned} v(P) &= \max\{F(z) | z \in Z^P\} \quad \text{with} \\ Z^P &= \{z | g(z,t) \le 0, t \in B\} \subset \mathbb{R}^n, \end{aligned}$$

B is a compact index set, and the functions F, g are continuously differentiable with respect to z everywhere on $\mathbb{R}^n$ and $\mathbb{R}^n \times \mathbb{R}^m$, respectively. Let $z \in Z^P$. Then we define

$$\bar{E} = E(\bar{z}) := \{t \in B | g(\bar{z}, t) = 0\} = \{\bar{t}^l | l \in L\}, \tag{3.1}$$

the set of t for which our constraint is active. It is an easy exercise to show that if $\bar{z}$ happens to be an optimal point of (P), then there cannot be a $\xi \in \mathbb{R}^n$ such that

$$\xi^T F_z > 0, \quad \xi^T g_z^l < 0, \qquad l \in L, \tag{3.2}$$

where $F_z := F_z(\bar{z})$, $g_z^l := g_z(\bar{z}, \bar{t}^l)$ for brevity. Condition (3.2) can be interpreted to be a sufficient condition for ξ being a feasible ascent direction. We will call a direction ξ feasible, if there is a feasible smooth arc emanating from $\bar{z}$ with tangent ξ. Obviously, if ξ is a feasible direction, we must have

$$\xi^T g_z^l \leq 0, \qquad l \in L. \tag{3.3}$$

Imposing a "constraint qualification," which ensures that every ξ with (3.3) is a feasible direction, we can then strengthen the above requirement to the following.

LEMMA 3.1. *Let $\bar{z} \in Z^P$ be optimal for (P) and suppose that a constraint qualification in the above sense holds. Then there is no $\xi \in \mathbb{R}^n$ such that*

$$\xi^T F_z > 0, \quad \xi^T g_z^l \leq 0, \qquad l \in L. \tag{3.4}$$

In order to obtain the Karush–Kuhn–Tucker (KKT) conditions we use the following version of the Farkas Lemma [1.8].

LEMMA 3.2. *Let $S \subset \mathbb{R}^n$ be arbitrary and* $\operatorname{co}(S)$ *be the convex cone generated by S. Then for every $v \in \mathbb{R}^n$, exactly one of the following alternatives is true*:

(i) $v \in cl(\operatorname{co}(S))$;

(ii) *there is a solution ξ of*

$$\xi^T v > 0, \quad \xi^T s \leq 0, \qquad s \in S. \tag{3.5}$$

Applying Lemma 3.2 to Lemma 3.1, we find that if $\bar{z}$ is optimal and any constraint qualification holds, we have that

$$F_z \in cl(C_g) \quad \text{with } C_g = \operatorname{co}(\{g_z^l | l \in L\}).$$

If C_g is closed, we get a KKT type theorem (i.e., F_z is a nonnegative linear combination of a finite number of g_z^l). It is well known (cf. [1.9]) that C_g is closed if we use for instance the following constraint qualification.

(CQ) There exists a ξ such that

$$\xi^T g_z^l < 0, \qquad l \in L. \tag{3.6}$$

Remark. This indeed is a constraint qualification in the above sense, i.e., if (CQ) holds, then every ξ satisfying (3.3) is a feasible direction. The proof is rather lengthy. However, assuming (CQ), a direct proof of Lemma 3.1 is obvious (cf. [1.8]).

THEOREM 3.3 (KKT). *Let $\bar{z} \in Z^P$ be optimal for* (P) *and suppose that* (CQ) *holds. Then there exists a $L' \subset L$, $|L'| < \infty$, and $\mu_l > 0, l \in L'$, such that*

$$F_z = \sum_{l \in L'} \mu_l g_z^l \qquad (\text{i.e., } F_z \in C_g;\ \text{cf. } (3.5)). \tag{3.7}$$

We can state this also in a different way.

Let $M^+(B)$ be the space of nonnegative Borel measures on B. Then (3.7) means that there is a $\mu \in M^+(B)$ with finite support $\operatorname{supp}(\mu) \subset E$ such that

$$F_z = \sum_{t \in B} \mu(t) g_z(\bar{z}, t) := \sum_{t \in \operatorname{supp}(\mu)} \mu(t)\, g_z(\bar{z}, t). \tag{3.7'}$$

The subset

$$\mathbb{R}_+^{(B)} := \{\mu \in M^+(B) |\ \operatorname{supp}(\mu) \text{ finite}\} \tag{3.8}$$

is called the set of nonnegative generalized finite sequences (cf. [6.10]).

Finally, it is easy to prove that if there is no $\xi \neq 0$ such that

$$\xi^T F_z \geq 0, \quad \xi^T g_z^l \leq 0, \qquad l \in L, \tag{3.9}$$

then there is no chance to improve $\bar{z}$ locally, stated more precisely as follows.

LEMMA 3.4. *If for $\bar{z} \in Z^P$ there is no $\xi \neq 0$ satisfying* (3.9), *then $\bar{z}$ is a strongly unique local solution of* (P), *i.e., there exists a $k > 0$ such that*

$$F(\bar{z}) - F(z) \geq k\|\bar{z} - z\| \tag{3.10}$$

for all $z \in Z^P \cap \bar{U}$, $\bar{U}$ some neighborhood of $\bar{z}$.

We can restate this lemma using the following variant of the Farkas lemma [1.8].

LEMMA 3.5. *Let $S \subset \mathbb{R}^n$ be arbitrary. Then for every $v \in \mathbb{R}^n$ we have exactly one of the following alternatives*:

(i) $v \in \operatorname{int}(\operatorname{co}(S))$;

(ii) *there is no $\xi \neq 0$ such that*

$$\xi^T v \geq 0, \quad \xi^T s \leq 0, \qquad s \in S. \tag{3.11}$$

This implies the following.

THEOREM 3.6. *If $F_z \in \operatorname{int}(C_g)$* (cf. (3.5)), *then $\bar{z}$ is a strongly unique local solution of* (P).

Remark 3.7. Equivalent to the assumption $F_z \in \operatorname{int}(C_g)$ is a stronger KKT condition (3.7) stating that L' can be chosen such that span $\{g_z^l | l \in L'\} = \mathbb{R}^n$ (cf. [3.3]).

Remark 3.8. Strong unicity has received much attention in Chebyshev approximation, where it is well known [3.1] that in Haar spaces best approximations are always strongly unique (for recent results see, for instance, [3.4], [3.5]). Strong unicity, moreover, ensures the continuous dependence of the best approximation on the function to be approximated (continuity of the Chebyshev operator; cf. [3.1]) and is responsible for the convergence of Remes-type algorithms (cf. [3.2]).

4. Discretization and (local) reduction to finite problems. In this section we consider the problem of describing or at least approximating the feasible set

$$Z^P = \{z | g(z, t) \leq 0, t \in B\}$$

of our semi-infinite problem (P) by imposing only finitely many constraints. The simplest way is through ordinary discretization.

Choose $\bar{B} \subset B$, $|\bar{B}| < \infty$, and replace Z^P by

$$Z^P(\bar{B}) := \{z | g(z, t) \leq 0, t \in \bar{B}\}, \tag{4.1}$$

and consider the approximate problem

$$(P(\bar{B})) \qquad \max\{F(z) | z \in Z^P(\bar{B})\}. \tag{4.2}$$

Typically $\bar{B}$ is termed a grid. Two questions arise in this context.

(a) Are there subsets $\bar{B}$ of B for which $(P(\bar{B}))$ and (P) have the same set of solutions?

(b) Let

$$d(\bar{B}) = \max_{t \in B} \min_{\bar{t} \in \bar{B}} \|t - \bar{t}\| \tag{4.3}$$

be the usual measure for the density of $\bar{B}$ in B. Then, if $\bar{B}^{(n)}$ is a sequence of finite subsets with $d(\bar{B}^{(n)}) \to 0$, is it true that points of accumulation of solutions $\bar{z}^{(n)}$ of $(P(\bar{B}^{(n)}))$ solve (P)?

In general the answer to both questions is negative. Regarding (a), a negative example is given by the following.

Example 4.1.

$$\max\{-z_2 | 0 \le z_1 \le 1, -(z_1 - t)^2 - z_2 \le 0, t \in [0, 1]\}$$

where the value of $(P(\bar{B}))$ is less than zero for every $\bar{B}$, $|\bar{B}| < \infty$ (see Fig. 4.1).

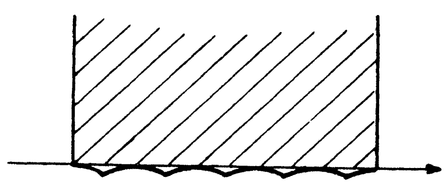

FIG. 4.1. *Feasible set of Example* 4.1.

The following is an immediate consequence of more general results given by Borwein [6.7] and obtained from Helly-type theorems (see also [6.1]).

THEOREM 4.2. *In Program (P) assume that B is compact, F is concave, $g(z, t)$ is convex with respect to z(F and $g(\cdot, t)$ all finite over $\mathbb{R}^n$), and $v(P)$ is finite. Assume further that the following type of Slater condition holds.*

For every set of $n+1$ points $t_0, \ldots, t_n \in B$ a $\tilde{z}$ exists such that $g(\tilde{z}, t_i) < 0$, $i = 0, \ldots, n$. Then there exists $T_n = \{t_1, \ldots, t_n\} \subset B$ such that

(i) $v(P) = v(P(T_n))$;

(ii) *there exist multipliers $\mu_i \ge 0$, $i = 1, \ldots, n$, such that*

$$v(P) = \sup\left\{F(z) - \sum_{i=1}^{n} \mu_i g(z, t_i) | z \in \mathbb{R}^n\right\}.$$

We note that (ii) is basically a convex SIP Lagrangian duality result (cf. §6). Of course a set T_n, even if it is known to exist, will not be known explicitly, but usually is a result of a numerical solution of the problem (cf. §7).

Remark 4.3. For the linear SIP related to linear Chebyshev approximation (see [1.8]) the assumptions in Theorem 4.2 obviously hold. The existence of T_n indeed is a well-known fact in linear Chebyshev approximation and actually the famous Remes algorithm (for approximation with Haar spaces) iteratively determines T_n.

With respect to question (b), a simple answer is possible in the linear case (see also [4.3]).

THEOREM 4.4 (cf. [1.8]). *Consider the linear problem (P) with $F(z) = c^T z$, $g(z, t) = a^T(t)z - b(t)$. Assume that all* P-*level sets (cf. Definition* 6.10 *in* §6) *are bounded (for equivalent assumptions cf.* §6, *Theorem* 6.11). *Then, for every $\varepsilon > 0$ there exists $\delta_\varepsilon > 0$ such* *that for all $\bar{B} \subset B, d(\bar{B}) < \delta_\varepsilon$ it is the case that $(P(\bar{B}))$ is solvable, and for every solution $\bar{z}$ there exists a solution z^* of (P) such that $\|\bar{z} - z^*\| < \varepsilon$.*

Convergence results when $d(\bar{B}) \to 0$ were given in [7.20] for convex objective functions and linearly constrained SIP problems which are in addition regularized so that the level sets are also bounded.

Remark 4.5. A good review on the question of approximation by discretization can be found in [4.2], [4.5]. See also [7.44] in this context.

Now let us turn to the second approach, where certain assumptions in some neighborhood of a point $z \in Z^P$, Z^P is equivalently described by a finite set of restrictions. We term this "an intrinsic SIP approach."

To this end, consider the following parametric optimization problem.

Problem 4.6.

$$(O(z)) \qquad \max\{g(z,t) | t \in B\}. \tag{4.4}$$

$z \in Z^P$ obviously is equivalent to $v(O(z)) \leq 0$. Moreover, for $\bar{z} \in Z^P$, the points $\bar{t}^l \in \bar{E}$ (cf. (3.1)), $l \in L$, are optimal solutions of $(O(z))$. Let us assume that the following holds.

Assumption 4.7. $\bar{E} = \{\bar{t}^1, \dots, \bar{t}^r\}$ is a finite set. There exist a neighborhood $U_{\bar{z}}$ of $\bar{z}$, neighborhoods $U_{\bar{t}^l}$ of $\bar{t}^l$, and continuous mappings

$$t^l : U_{\bar{z}} \to U_{\bar{t}^l} \cap B \tag{4.5}$$

such that

(i) $t^l(\bar{z}) = \bar{t}^l, \quad l = 1, \dots, r;$

(ii) for every $z \in U_{\bar{z}}$ and $l = 1, \dots, r$, $t^l(z)$ is the only local solution of $(O(z))$ in $U_{\bar{t}^l} \cap B$.

The following lemma is immediate.

LEMMA 4.8. *Let $\bar{z} \in Z^P$ be given and let Assumption* 4.7 *hold. Then there exists a neighborhood $\bar{U}$ of $\bar{z}$ such that for all $z \in \bar{U}$ we have $z \in Z^P$ if and only if*

$$G^l(z) := g(z, t^l(z)) \leq 0, \qquad l = 1, \dots, r. \tag{4.6}$$

We define the so-called locally reduced problem as follows.

Problem 4.9.

$$(P_{\text{red}}(\bar{z})) \qquad \max\{F(z) | G^l(z) \leq 0, l = 1, \dots, r; z \in \bar{U}\}.$$

THEOREM 4.10. *Let $\bar{U}$ be a neighborhood of $\bar{z} \in Z^P$ as in Lemma* 4.8. *Then we have*

A point $\tilde{z} \in \bar{U}$ is locally optimal for (P) if and only if it is for $(P_{\text{red}}(\bar{z}))$.

It is very easy to prove that for linear problems (P) the functions $G^l(z), l = 1, \dots, r$, are convex in $\bar{U}$ [1.8]. Thus, given Assumption 4.7, we can give—at least locally—a description of our feasible set in terms of a finite number of convex constraints which, however, are defined only implicitly via solutions of a parametric optimization problem. This again illustrates the difficulty of formulating a given linear semi-infinite problem as an ordinary convex optimization problem. Nevertheless, Theorem 4.10 is a vehicle to transmit theory and methods of finite programming to semi-infinite problems. For this it is helpful to study some "regular" cases which give more information about the functions $t^l(z)$, $G^l(z)$.

Assumption 4.11. Assume that the compact set B is a subset of $\mathbb{R}^m$ given by

$$B = \{t | h^j(t) \leq 0, j \in M\} \tag{4.7}$$

with $|M| < \infty$ and $h^j \in C^2(\mathbb{R}^m)$, $j \in M$. Moreover, in every point of B the linear independence constraint qualification (LICQ) holds: for every $t \in B$, the vectors

$$h_t^j(t), \qquad j \in M_t := \{i | h^i(t) = 0\}$$

are linearly independent.

We note for the sequel that Assumption 4.11 could be relaxed. However, in applications B can often be defined according to Assumption 4.11. Therefore, we choose this stronger description of B for simplicity. We immediately have the following result.

LEMMA 4.12. *Given $\bar{z} \in \mathbb{R}^n$. Assume that Assumption* 4.11 *holds. For every $\bar{t}^l \in \bar{E}$ of* (3.1) *of* §3 *define*

$$M^l = \{j | h^j(\bar{t}^l) = 0\} \tag{4.8}$$

and the Lagrangian of $(O(\bar{z}))$ (*Problem* 4.6) *with respect to $\bar{t}^l$* :

$$\mathcal{L}^l(t, \alpha^l) := g(\bar{z}, t) - \sum_{j \in M^l} \alpha_j^l h^j(t). \tag{4.9}$$

Then there exist unique multipliers $\bar{\alpha}_j^l \geq 0$ such that

$$\mathcal{L}_t^l(\bar{t}^l, \bar{\alpha}^l) = 0. \tag{4.10}$$

Assumption 4.13. Let $g \in C^2(\mathbb{R}^n \times \mathbb{R}^m)$. Assumption 4.11 holds, and for every $\bar{t}^l \in \bar{E}$ the following well-known strong second-order sufficient condition (SSOSC) for $\bar{t}^l$ to be a strict local maximum of $(O(\bar{z}))$ holds.

$\mathcal{L}_{tt}^l(\bar{t}, \bar{\alpha}^l)$ is negative definite on the tangent space

$$T^l = \{\eta | \bar{\alpha}_j^l \eta^T h_t^j(\bar{t}^l) = 0, j \in M^l\}. \tag{4.11}$$

Assumption 4.14. In addition to Assumption 4.13 for all $\bar{t}^l \in \bar{E}$, strict complementary slackness (SCS) is given, i.e., $\bar{\alpha}_j^l > 0$ for all $j \in M^l$.

We note that in [5.6] the setting under Assumption 4.14 is denoted as the "standard case" in semi-infinite programming since in some sense this assumption is generic.

Then we obtain the following theorem (cf. [5.6]).

THEOREM 4.15. (a) *Assume that Assumption* 4.14 *holds. Then Assumption* 4.7 *holds with continuously differentiable functions $t^l : U_{\bar{z}} \to U_{\bar{t}^l} \cap B$. Moreover, there exist continuously differentiable $\alpha^l : U_{\bar{z}} \to \mathbb{R}$ such that Assumption* 4.14 *holds for all triples $z, t^l(z), \alpha^l(z), z \in U_{\bar{z}}$. The derivatives $t_z^l = t_z^l(z)$, $\alpha_z^l = \alpha_z^l(z)$ are uniquely determined by*

$$\begin{pmatrix} \mathcal{L}_{tt}^l(t^l, \alpha^l) & H_t^l(t^l) \\ (H_t^l(t^l))^T & 0 \end{pmatrix} \begin{pmatrix} t_z^l \\ \alpha_z^l \end{pmatrix} = - \begin{pmatrix} g_{tz}^l(z, t^l) \\ 0 \end{pmatrix} \tag{4.12}$$

with

$$H_t^l(t^l) = (h_t^j(t^l))_{j \in M^l} \in \mathbb{R}^{m \times |M^l|}. \tag{4.13}$$

(b) *The constraint functions $G^l(z) = g(z, t^l(z))$ of the reduced problem are twice continuously differentiable in $U_{\bar{z}}$ and are given by*

$$G_z^l(z) = g_z(z, t^l(z)), \tag{4.14}$$

$$G_{zz}^l(z) = g_{zz}^l(z, t^l(z)) - (t_z^l(z))^T \mathcal{L}_{tt}^l(t^l(z), \alpha^l(z)) t_z^l(z). \tag{4.15}$$

The proof of (a) uses the Implicit Function Theorem and obvious continuity arguments. Equations (4.14) and (4.15) are derived from (4.12).

Remark 4.16. From a result in [4.4] (see also [5.9]) it follows that Assumption 4.13 is sufficient to have t^l, α^l Lipschitz continuous in $U_{\bar{z}}$. For all $\xi \in \mathbb{R}^n$, t^l has a directional derivative $Dt^l(\bar{z}; \xi)$ in $z = \bar{z}$ which is the unique solution η of the following quadratic programming problem.

Problem 4.17.

$$\max_{\eta \in T^l} \{\xi^T g_{zt}(\bar{z}, \bar{t}^l)\eta + \tfrac{1}{2}\eta^T \mathcal{L}^l_{tt}\eta\}$$

with T^l given by (4.11).

For G^l this implies continuous differentiability and existence of second-order directional derivatives.

Remark 4.18. Further relaxations of the assumptions are possible to define reduced problems. If $\bar{t}^l$ is assumed to be strongly stable (cf. [8.21]), it still follows that t^l exists in a neighborhood as a continuous function which implies that G^l is continuously differentiable (cf. [8.21], [4.1], [4.6]). In [7.52] a reduced problem is constructed with G^l only Lipschitz continuous on the basis of $t^l(z)$, which may even be discontinuous.

5. Second-order optimality conditions. The easiest way to obtain second-order optimality conditions for our semi-infinite problem (P) is by assuming Assumption 4.14 to hold and to apply known optimality conditions of finite optimization to the (locally equivalent) reduced problem (SIP_{red}) with twice continuously differentiable constraint functions $G^l(z)$ (Theorem 4.15).

Based on conditions for finite problems given in [5.5], in this way the following has been proved in [5.6].

THEOREM 5.1. *Suppose that at* $\bar{z} \in Z^P$ *Assumption* 4.14 *holds. Then we have the following.*

(a) (*necessary condition*). *If* $\bar{z}$ *solves* (P), *then for every* ξ *in* $\mathcal{K}$,

$$\mathcal{K} = \{\xi \mid \xi^T F_z \geq 0, \xi^T g^l_z \leq 0, l \in L\}, \tag{5.1}$$

there exists $\mu_0(\xi) \geq 0$, $\mu(\xi) \in \mathbb{R}^{(B)}_+$, $\text{supp}(\mu(\xi)) \in \bar{E}((\mu_0(\xi), \mu(\xi)) \neq (0, 0))$ *such that, with*

$$L(z, \mu_0, \mu, t) := \mu_0 F(z) - \sum_{l \in L} \mu_l g(z, t^l), \tag{5.2}$$

we have

$$L_z(\bar{z}, \mu_0(\xi), \mu(\xi), \bar{t}) = 0 \tag{5.3}$$

and

$$\begin{aligned} q(\xi) := {} & \xi^T L_{zz}(\bar{z}, \mu_0(\xi), \mu(\xi), \bar{t})\xi \\ & + \sum_{l \in L} \mu_l(\xi)(\bar{t}^l_z \xi)^T \mathcal{L}^l_{tt}(\bar{t}^l, \alpha(\bar{z}))(\bar{t}^l_z \xi) \leq 0. \end{aligned} \tag{5.4}$$

(b) (*sufficient condition*). *If for every* $\xi \in \mathcal{K}$ *there exist* $\mu_0(\xi)$, $\mu(\xi)$ *as in* (a) *with* (5.3) *and, for* $\xi \neq 0$,

$$q(\xi) < 0, \tag{5.5}$$

then $\bar{z}$ is a strict local solution of (P) (*i.e., there exists a neighborhood $\bar{U}$ of $\bar{z}$ such that for $z \in \bar{U} \cap Z^P$ we have $F(z) < F(\bar{z})$*).

(c) (*strong sufficient condition*). *Part* (b) *is especially satisfied if*

—*the g_z^l, $l \in L$ are linearly independent*;

—*there exist* (*unique*) $\mu_l \geq 0$ *such that the KKT condition* (5.3) *holds with* $\mu_0(\xi) = 1, \mu(\xi) = \mu$;

—(5.5) *holds* (*with* $\mu_0(\xi) = 1, \mu(\xi) = \mu$) *for all* $\xi \in \mathcal{T}$,

$$\mathcal{T} = \{\xi \mid \mu_l (g_z^l)^T \xi = 0, l \in L\}.$$

Remark 5.2. The two terms of $q(\xi)$ deserve some interpretation. The first one, $\xi^T L_{zz}\xi$, is the second-order term we would obtain for the discretized problem $(P(\bar{E}))$ (cf. (4.2)). The second term,

$$s(\xi) := \sum_{l \in L} \mu_l(\xi)(\bar{t}_z^l \xi)^T \mathcal{L}_{tt}^l (\bar{t}_z^l \xi), \tag{5.6}$$

reflects the "semi-infinite" structure inasmuch as it is generated by the shift of active constraints $t^l(z)$ as a function of z. We, therefore, denote $s(\xi)$ as a "shift term." Such a term first appeared in [5.15] (according to Kawasaki [5.10] it is the term which produces an "envelope-like" effect). The following example illustrates the importance of this term in optimality conditions.

Example 5.3. Consider the problem of maximizing

$$F(z) = (z_2 + 5)^2 + z_1^2 \quad \text{subject to}$$

$$g(z, t) = 2tz_1 + z_2 - t^2 \leq 0, \qquad t \in [-1, 1].$$

The feasible region Z^P is shown in Fig. 5.1, and is described by $g(z, t) \leq 0$, as the intersection of its supporting halfspaces.

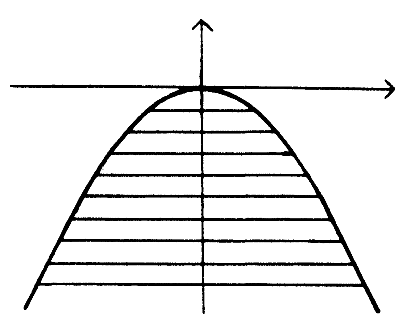

FIG. 5.1. *Feasible set of Example* 5.3.

Obviously $\bar{z} = 0$ is a strict local solution with

$$\bar{E} = E(\bar{z}) = \{\bar{t}^1 = 0\} \subset [-1, 1].$$

We obtain $\mathcal{K} = \{\binom{\xi_1}{0} \mid \xi_1 \in \mathbb{R}\}$, and for every ξ we have (5.3) with $\mu_0(\xi) = 1, \mu_1(\xi) = 10$ (unique up to a common positive factor).

This gives, for $\xi \in \mathcal{K}$,

$$\xi^T L_{zz}\xi = 2\xi_1^2, \quad s(\xi) = -20\xi_1^2, \quad \text{i.e., } q(\xi) = -18\xi_1^2,$$

proving, by Theorem 5.1, that $\bar{z} = 0$ is indeed a strict local solution.

Note, however, that $\xi^T L_{zz}\xi = 2\xi_1^2$ is positive definite on $\mathcal{K}$. This means that if only this term would be used in second-order conditions, then

—a sufficient condition could never hold;

—a necessary condition could be only the trivial one ($\xi^T L_{zz}\xi \leq 0$ for $\xi = 0$).

Now, observe that in $s(\xi)$ in (5.6) only the directional derivative

$$Dt^l(\bar{z}; \xi)(= \bar{t}^l_z \xi)$$

is employed, the existence of which is already implied by Assumption 4.13. This suggests that, to derive second-order conditions, it could be sufficient to replace Assumption 4.14 by Assumption 4.13, and $\bar{t}^l_z\xi$ by $Dt^l(\bar{z}, \xi)$. The difficulty is, that for the constraint functions $G^l(z) = g(z, t^l(z))$ of (SIP$_{\text{red}}$), we no longer have continuous second-order derivatives, but only existence of second-order directional derivatives [5.9]. In [5.9] it indeed could be shown that this is sufficient for a second-order sufficient condition based on such a second-order term. In a forthcoming paper [5.7] it will be shown that Theorem 5.1 actually remains true under Assumption 4.13 when $\bar{t}^l_z\xi$ is replaced by $Dt^l(\bar{z}, \xi)$.

For sufficient conditions a further relaxation is possible, which replaces $Dt^l(\bar{z}, \xi)$ by a solution of Problem 4.17 (if one exists).

For illustration we prove the following theorem.

THEOREM 5.4. *Suppose that Assumption* 4.11 *holds. Then the following is sufficient for* $\bar{z} \in Z^P$ *to be a strict local maximum of* (P).

For every $\xi \in \mathcal{K}$ *there exist* $\mu_0(\xi) \geq 0$, $\mu(\xi) \in \mathbb{R}^{(B)}_+$ (*not both zero*) *such that* (5.3) *holds and* (*if* $\xi \neq 0$)

$$\tilde{q}(\xi) = \xi^T L_{zz}(\bar{z}, \mu_0(\xi), \mu(\xi), \bar{t})\xi + \sum_{l \in L} \mu_l(\xi)\eta_l^T(\xi)\mathcal{L}^l_{tt}(\bar{t}^l, \alpha(\bar{z}))\eta_l(\xi) < 0,$$

where, depending on the solvability of $(QP_l(\xi))$ (*Problem* 4.17), $\eta_l(\xi)$ *is zero or a solution of* $(QP_l(\xi))$.

The proof uses the following two lemmas.

LEMMA 5.5. *Suppose that* $\eta_l(\xi)$ *solves* $(QP_l(\xi))$ *and Assumption* 4.11 *holds. Then the optimal value of* $(QP_l(\xi))$ *is given by*

$$\xi^T g^l_{zt}\eta_l(\xi) + \tfrac{1}{2}\eta_l^T(\xi)\mathcal{L}^l_{tt}\eta_l(\xi) = -\tfrac{1}{2}\eta_l^T(\xi)\mathcal{L}^l_{tt}\eta_l(\xi). \tag{5.7}$$

The proof is easily seen from the KKT conditions (which hold in $\eta_l(\xi)$ due to Assumption 4.11 and the definition of T^l).

LEMMA 5.6. *Suppose Assumption* 4.11 *holds. Then for each* $\eta \in T^l$ (*cf.* (4.11)) *there exists a twice continuously differentiable function* $\psi^l : [0, \tau_0] \to B(\tau_0 > 0)$ *such that* $\psi^l(0) = \bar{t}^l$, $\psi^l_\tau(0) = \eta$, *and* $h^j(\psi^l(\tau)) \equiv 0$, $j \in M^l$.

The proof is an easy consequence of the Implicit Function Theorem.

Proof of Theorem 5.4. Suppose that $\bar{z}$ is not a strict local maximum. Then, a sequence $z^k \in Z^P$, $z^k \to \bar{z}$ exists with $F(z^k) \geq F(\bar{z})$. Let

$$z^k = \bar{z} + \tau_k\xi^k, \qquad \|\xi^k\| = 1, \quad \tau_k > 0,$$

thus $\tau_k \searrow 0$. Without restriction we can assume that $\xi^k \to \bar{\xi}$, $\|\bar{\xi}\| = 1$. It is easy to show that $\bar{\xi} \in \mathcal{K}$. Let $\bar{\mu}_0 = \mu_0(\bar{\xi})$, $\bar{\mu} = \mu(\bar{\xi})$, and $\bar{\eta}_l = \eta_l(\bar{\xi})$ be as in the theorem. Note that $\bar{\eta}_l \in T^l$.

Let $\bar{\psi}^l(\tau)$ be as in Lemma 5.6, corresponding to $\bar{\eta}_l$.

Taylor expansion and utilization of (5.3) and (4.10) gives

$$0 \le \bar{\mu}_0(F(z^k) - F(\bar{z})) - \sum_{l\in L} \bar{\mu}_l \left[g(z^k, \bar{\psi}^l(\tau_k)) - \sum_{j\in M^l} \bar{\alpha}^l_j h^j(\bar{\psi}^l(\tau_k)) \right]$$

$$= \frac{1}{2}\tau_k^2 \left\{ (\xi^k)^T \left[\bar{\mu}_0 F_{zz}(z_\Theta) - \sum_{l\in L} \bar{\mu}_l g_{zz}(z_\Theta, \bar{\psi}^l_\Theta) \right] \xi^k \right.$$

$$-2\sum_{l\in L} \bar{\mu}_l (\xi^k)^T g_{zt}(z_\Theta, \bar{\psi}^l_\Theta)\bar{\eta}_l$$

$$\left. - \sum_{l\in L} \bar{\mu}_l \bar{\eta}_l^T \left[g_{tt}(z_\Theta, \bar{\psi}^l_\Theta) - \sum_{j\in M^l} \bar{\alpha}^l_j h^j_{tt}(\bar{\psi}^l_\Theta) \right] \bar{\eta}_l \right\}$$

$$- \sum_{l\in L} \bar{\mu}_l \bar{\psi}^l_{\tau\tau}(\Theta\tau_k) \left[g_t(z_\Theta, \bar{\psi}^l_\Theta) - \sum_{j} \bar{\alpha}^l_j h^j_t(\bar{\psi}^l_\Theta) \right] \psi^l_{\tau\tau}(\Theta\tau_k)$$

with a (fixed) $\Theta \in (0, 1)$, $z_\Theta = \bar{z} + \Theta\tau_k\xi^k$, $\bar{\psi}^l_\Theta = \bar{\psi}^l_\Theta(\Theta\tau_k)$.

For $k \to \infty$ this gives, due to Lemma 5.5,

$$0 \le \bar{\xi}^T L_{zz}\bar{\xi} - 2\sum_{l\in L} \bar{\mu}_l \left[\bar{\xi}^T g^l_{zt}\bar{\eta}_l + \frac{1}{2}\bar{\eta}_l^T \mathcal{L}^l_{tt}\bar{\eta}_l \right]$$

$$= \bar{\xi}^T L_{zz}\bar{\xi} + \sum_{l\in L} \bar{\mu}_l \bar{\eta}_l^T \mathcal{L}^l_{tt}\bar{\eta}_l = \tilde{q}(\xi),$$

a contradiction.

6. Duality theory in linear semi-infinite programming. We turn next to a review of duality theory for linear (and hence convex) SIP. In this section we study approaches to guaranteeing the duality equality between primal and dual program values (perfect duality), and then present a theory responding to the occurrence of duality gaps in these program values. For clarity we present the material in four parts (§§6.1–6.4).

6.1. The basic dual problem (D) in finite sequence space. Our choice for primal problem (P) is the following, where we retain "P" for primal as in §2. Let $B \subset \mathbb{R}^m$ be an arbitrary set, $a : B \to \mathbb{R}^n$, $b : B \to \mathbb{R}$ given functions and $c \in \mathbb{R}^n$ a given vector.

$$\text{(P)} \qquad \begin{aligned} &\text{Find } v(P) = \sup\,\{c^T z | z \in Z^P\}, \quad \text{where} \\ &Z^P = \{z \in \mathbb{R}^n | a^T(t)z \le b(t) \text{ for all } t \in B\}. \end{aligned}$$

We emphasize that in general we neither require B to be a compact set nor the functions a, b to be continuous.

In case of a finite set $B = \{t_1, \dots, t_r\}$ the usual LP-dual to P would be to determine the infimum of $\sum_{i=1}^r b(t_i)\mu(t_i)$ over the set of nonnegative multipliers $\mu(t_i)$ with $c = \sum_{i=1}^r a(t_i)\mu(t_i)$. Analogously, adopting the notion of $\mathbb{R}_+^{(B)}$ (cf. (3.8)) we define the dual in the space of generalized finite sequences (GFS) by

(D)
$$\text{Find } v(D) = \inf\left\{\sum_{t\in B} b(t)\mu(t) | \mu \in Z^D\right\}, \quad \text{where}$$
$$Z^D = \left\{\mu \in \mathbb{R}_+^{(B)} \middle| \sum_{t\in B} a(t)\mu(t) = c\right\}.$$

A trivial calculation shows that for any $\mu \in Z^D$, $z \in Z^P$,

$$c^T z \leq \sum_{t\in B} b(t)\mu(t), \tag{6.1}$$

implying $v(P) \leq v(D)$ and $v(P) \leq \sum_{t\in B} b(t)\mu(t)$ for every $\mu \in Z^D$. Therefore, an interpretation of D is: minimize the upper bound $\sum_{t\in B} b(t)\mu(t)$ for $v(P)$ given by $\mu \in Z^D$. The central question in duality theory is whether this upper bound can be made sharp, i.e., if $v(P) = v(D)$.

The following cones will play an important role:

$$\begin{aligned} M_n &= \operatorname{co}(\{a(t) | t \in B\}) \\ &= \left\{w = \sum_{t\in B} a(t)\mu(t) | \mu \in \mathbb{R}_+^{(B)}\right\} \subset \mathbb{R}^n \end{aligned} \tag{6.2}$$

$$M_{n+1} = \operatorname{co}\left(\left\{\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} | t \in B\right\}\right) \subset \mathbb{R}^{n+1}. \tag{6.3}$$

The next lemma is obvious from the definitions.

LEMMA 6.1.

(i) $Z^D \neq \phi$ *if and only if* $c \in M_n$;

(ii) $v(D) = v(D_G)$, *where* (D_G) *is the "geometrical dual"*

$$(D_G): \qquad \text{Find } v(D_G) = \inf\left\{d | \begin{pmatrix} c \\ d \end{pmatrix} \in M_{n+1}\right\}.$$

These simple observations nevertheless exhibit the close connection of SIP to moment problems and provide the possibility of treating an important class of applications with semi-infinite programming methods.

For an illustration we consider the case that the following holds.

Assumption 6.2. Set B is compact, and a, b are continuous on B. Assuming 6.2, a theorem of Rogosinsky [6.50] allows the following representations of the cones M_n, M_{n+1}.

LEMMA 6.3. *Suppose that Assumption* 6.2 *holds. Then, with* $M^+(B)$ *again the nonnegative regular Borel measures on* B, *we have*

$$M_n = \left\{w = \int_B a(t) d\mu(t) | \mu \in M^+(B)\right\},$$

$$M_{n+1} = \left\{w = \int_B \begin{pmatrix} a(t) \\ b(t) \end{pmatrix} d\mu(t) | \mu \in M^+(B)\right\}.$$

With regard to Lemmas 6.1 and 6.3 it is natural to formulate a dual in measure space:

$$(D_M): \qquad \text{Find } v(D_M) = \inf \left\{ \int_B b(t) d\mu(t) | \mu \in Z^{D_M} \right\}$$

$$z^{D_M} = \left\{ \mu \in M^+(B) | \int_B a(t) d\mu(t) = c \right\}.$$

Lemmas 6.1 and 6.3 immediately give the next lemma.

LEMMA 6.4. *Suppose that Assumption* 6.2 *holds. Then*

(i) $v(D_G) = v(D_M) = v(D)$;

(ii) *every solution* μ^* *of* (D) *solves* (D_M) *and, if* (D_M) *is attained, there is a* $\mu^* \in \mathbb{R}_+^{(B)} \subset M^+(B)$ *which solves* (D) *and* (D_M).

Due to Lemma 6.4, in most applications it is sufficient to consider the dual (D) in finite sequence space.

6.2. Duality: the superconsistent case for primal or dual program. The following sufficient conditions for $v(P) = v(D)$ have been well known for a long time (cf. [6.24] for a discussion).

THEOREM 6.5.

(i) *Suppose that* $v(D)$ *is finite and* $c \in \operatorname{int}(M_n)$. *Then* $v(D) = v(P)$ *and* (P) *is attained.*

(ii) *Suppose that* $v(P)$ *is finite and* M_{n+1} *is closed. Then* $v(D) = v(P)$ *and* (D) *is attained.*

Note that Theorem 6.5(ii) immediately gives $v(\tilde{D}_G) = v(P)(v(P)$ assumed to be finite) for the alternative geometrical dual

$$(\tilde{D}_G): \qquad \text{Find } v(\tilde{D}_G) = \inf \left\{ d | \begin{bmatrix} c \\ d \end{bmatrix} \in cl(M_{n+1}) \right\}$$

(cf. [1.9], [6.18], [6.42], [6.47], [6.52] for this approach).

The condition $c \in \operatorname{int}(M_n)$ strengthens the requirement that (D) is consistent (i.e., $c \in M_n$).

DEFINITION 6.6. (D) is called *superconsistent*, if $c \in \operatorname{int}(M_n)$.

Superconsistency implies a sort of stability: If Assumption 6.2 holds we can disturb the "data" a, b, c a small amount, without losing superconsistency or strong duality, $v(P) = v(D)$. A similar strengthening of primal consistency enforces closure of M_{n+1}, thus again guaranteeing strong duality by Theorem 6.5(ii).

DEFINITION 6.7. (P) is called *superconsistent* if Assumption 6.2 holds and if there is a $z^* \in Z^P$ such that $a^T(t)z^* < b(t)$.

Again superconsistency is stable in the sense that it is not lost by small perturbations of the "data" a, b, c. We have the following result [1.5], [6.10], [6.12], [6.20].

LEMMA 6.8. *Suppose that Assumption* 6.2 *holds. If* (P) *is superconsistent, then* M_{n+1} *is closed.*

This immediately gives the following nice symmetric result for our dual pair (P, D).

THEOREM 6.9. *Suppose that Assumption* 6.2 *holds. If either of the problems* (P) *or* (D) *is finite and superconsistent, then the other has attainment with* $v(P) = v(D)$.

Our next Theorem 6.11 below also shows that the implications of P or D superconsistency on the respective dual problem are indeed completely symmetric.

DEFINITION 6.10. P-level sets and D-level sets (for level κ) are respectively, defined by

$$L_{\geq}(Z^P, c, \kappa) = \{z \in Z^P | c^T z \geq \kappa\},$$

$$L_{\leq}(Z^D, b, \kappa) = \left\{\mu \in Z^D | \sum b(t)\mu(t) \leq \kappa\right\}.$$

For a given subset Z of a normed space X the recession cone $O^+(Z)$ is defined in the usual way by

$$O^+(Z) = \{d \in X | z + \lambda d \in Z \text{ for all } z \in Z, \lambda \geq 0\}.$$

THEOREM 6.11. *Let Assumption* 6.2 *prevail.*

(a) *If* (P) *is consistent, the following are equivalent:*

(a_1) *all* P*-level sets are bounded;*

(a_2) *then there exists* $\varepsilon > 0$ *such that for all* c', $\|c - c'\| < \varepsilon$, *one has* $\sup\{c'^T z | z \in Z^P\} < \infty$;

(a_3) $O^+(L_{\geq}(Z^P, c, \kappa)) = \{0\}$ *for all levels* κ;

(a_4) (D) *is superconsistent.*

(b) *If* (D) *is consistent, the following are equivalent:*

(b_1) *all* D*-level sets are bounded* (*i.e., for any* κ *there exist* ρ_κ *such that* $\sup_t |\mu(t)| \leq \rho_\kappa < \infty$ *for all* $\mu \in L_{\leq}(Z^D, c, \kappa)$);

(b_2) *there exists* $\varepsilon > 0$ *such that for all* b', $\|b - b'\|_\infty < \varepsilon$, *one has* $\inf\{\sum b'(t)\mu(t) | \mu \in Z^D\} > -\infty$;

(b_3) $O^+(L_{\leq}(Z^D, b, \kappa)) = \{0\}$ *for all levels* κ;

(b_4) (P) *is superconsistent.*

We note that (a_1) implies compactness of P-level sets. In [8.27] the equivalent of Theorem 6.11 is shown for the pair (P, D_M) with (b_1) sharpened to weak* compactness of the D_M-level sets. We give here an independent, elementary proof of part (b), while that of (a) is along the same lines (but easier). We will make use of the following lemma.

LEMMA 6.12. *Let Assumption* 6.2 *prevail. Suppose* $c \in cl(M_n)$ *and assume that a bounded sequence* $\{\mu^{(k)}\}_k$ *in* $\mathbb{R}_+^{(B)}$ *exists with* $c = \lim_k \sum a(t)\mu^{(k)}(t)$. *Then there exists* $\bar{\mu} \in \mathbb{R}_+^{(B)}$ *such that*

$$c = \sum a(t)\bar{\mu}(t) \quad \textit{and} \quad \sum b(t)\bar{\mu}(t) \leq \varliminf_k \sum b(t)\mu^{(k)}(t).$$

Proof. Caratheodory's lemma or the equivalent LP-based reduction coupled with a "purification" construction, say [6.17], [6.44], [6.45] guarantees for each k the existence of reals $\mu_{ki} \geq 0$, $1 \leq i \leq n+1$ and $\{t_{ki}\}_{i=1}^{n+1} \subset \operatorname{supp} \mu^{(k)}$ for which

$$\sum_{i=1}^{n+1} a(t_{ki})\mu_{ki} = \sum a(t)\mu^{(k)}(t) \quad \text{and} \quad \sum_{i=1}^{n+1} b(t_{ki})\mu_{ki} \leq \sum b(t)\mu^{(k)}(t).$$

Now using (6.1), the compactness of B, (6.2), the continuity of all functions, (6.3), and the uniform boundedness of μ_{ki} $(i = 1, \ldots, n+1;\ k = 1, 2, \ldots)$ we can apply the usual process of extracting successive subsequences of $\{t_{ki}, \mu_{ki}\}_k$ (first for $i = 1$, then $i = 2$ up to $i = n+1$) to obtain the desired result. (This process has been referred to as the Helly Selection Principle; see [2.42] and [2.38, Chap. 4, Thm. 4].)

Proof of Theorem 6.11(b). The equivalence of (b_1) and (b_3) is a standard result on convex sets (cf. [1.10, Thm. 8.4], for instance).

Next we prove the equivalence of (b_3) and (b_2). Assume that (b_3) is true. If on the other hand (b_2) were false, then for each $k = 1, 2, \ldots$ there exists $b^{(k)} \in C(B)$ such that

$$b(t) - \frac{1}{k} \leq b^{(k)}(t) \leq b(t) + \frac{1}{k} \quad \text{for } t \in B \tag{6.4a}$$

and

$$\inf\left\{\sum b^{(k)}(t)\mu(t) | \mu \in Z^D\right\} = -\infty. \tag{6.4b}$$

Hence, for each k, there exists $\mu^{(k)} \in Z^D$, $\mu^{(k)} \neq 0$, such that

$$\sum b(t)\mu^{(k)}(t) \leq -k + \frac{1}{k}\|\mu^{(k)}\|_1. \tag{6.5}$$

Therefore, $\{\|\mu^{(k)}\|_1\}_k$ must be unbounded, for otherwise the right side of (6.5) decreases arbitrarily with k, while the left side is bounded because $b \in C(B)$. Applying Lemma 6.12 (for $c = 0 \in cl(M_n)$) to the normalized sequence $\tilde{\mu}^{(k)} = \mu^{(k)}/\|\mu^{(k)}\|_1$ establishes the existence of $\bar{\mu}$ such that

$$\|\bar{\mu}\|_1 = 1, \quad \sum a(t)\bar{\mu}(t) = 0, \quad \text{and} \quad \sum b(t)\bar{\mu}(t) \leq 0, \tag{6.6}$$

i.e., $0 \neq \bar{\mu} \in O^+(L_{\leq}(Z^D, b, \kappa))$ for every $\kappa > v(D)$. This contradicts (b_3). Hence, (b_3) implies (b_2).

On the other hand, assume there exists a nonzero $\bar{\mu} \in O^+(L_{\leq}(Z^D, b, \kappa))$ for some κ. Since (D) is consistent, let $\hat{\mu} \in Z^D$ and form the sequence $\mu^{(k)} = \hat{\mu} + k\bar{\mu} \in Z^D$. Let $\delta > 0$ be otherwise arbitrary. Then, due to $\sum b(t)\bar{\mu}(t) \leq 0$,

$$\sum(b(t) - \delta)\mu^{(k)} \leq \sum b(t)\hat{\mu}(t) - \delta\|\hat{\mu}\|_1 - k\delta\|\bar{\mu}\|_1. \tag{6.7}$$

For $k \to \infty$ the right side of (6.7) $\to -\infty$. Hence, (b_2) cannot hold. Hence, (b_2) is equivalent to (b_3).

Finally we establish that (b_3) is equivalent to (b_4).

Assume (b_3) holds. This means that the following property holds:

$$\mu \in \mathbb{R}_+^{(B)}, \ \|\mu\|_1 = 1, \ \text{and} \ \sum a(t)\mu(t) = 0 \ \text{implies} \ \sum b(t)\mu(t) > 0. \tag{6.8}$$

This motivates the following pair of dual linear SIP programs:

$$\begin{aligned} (\tilde{P}): \quad & v(\tilde{P}) = \sup\{z_{n+1}\} \quad \text{subject to} \\ & a(t)^T z + z_{n+1} \leq b(t) \quad \text{for } t \in B. \end{aligned} \tag{6.9a}$$

$$\begin{aligned} (\tilde{D}): \quad & v(\tilde{D}) = \inf\left\{\sum b(t)\mu(t)\right\} \quad \text{subject to} \\ & \sum \begin{pmatrix} a(t) \\ 1 \end{pmatrix}\mu(t) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad \mu \in \mathbb{R}_+^{(B)}. \end{aligned} \tag{6.9b}$$

Obviously $(\tilde{P})$ is superconsistent and Assumption 6.2 holds. If $v(\tilde{P})$ were ≤ 0, Theorem 6.9

gives $v(\tilde{D}) \leq 0$ and attainment for $(\tilde{D})$. Hence there exists $\tilde{\mu} \in \mathbb{R}_+^{(B)}$ such that $\|\tilde{\mu}\|_1 = 1$, $\sum a(t)\tilde{\mu}(t) = 0$, and $\sum b(t)\tilde{\mu}(t) \leq 0$, a contradiction to (6.8). Thus, $v(\tilde{P}) > 0$, i.e., (b_4) holds. Hence, (b_3) implies (b_4).

The reverse is straightforward because, if z^0 were a Slater point and $0 \neq \bar{\mu} \in O^+(L_{\leq}(Z^D, b, \kappa))$, then we obtain the contradiction

$$0 \geq \sum (b(t) - a(t)^T z^0)\bar{\mu}(t) > 0.$$

This completes the proof.

Example 6.13. For the index set B, we select the ellipse in $\mathbb{R}^2$ given by

$$B = \{(t|2(t_1 - 0.5)^2 + t_2^2 = 0.5\}.$$

With

$$F(z) = z \quad \text{and} \quad Z^P = \{z | z(t_2 - t_1) \leq 1 - t_1, t \in B\} \subset \mathbb{R}$$

we have the problem

$$v(P) = \sup\{z | z \in Z^P\}.$$

For the dual we obtain

$$v(D) = \inf \left\{ \sum_{t \in B} (1 - t_1)\mu(t) | \mu \in Z^D \right\}$$

$$Z^D = \left\{ \mu \in \mathbb{R}_+^{(B)} | \sum_{t \in B} (t_2 - t_1)\mu(t) = 1 \right\}.$$

It is easy to prove that $Z^P = [0, 2]$, giving $\bar{z} = 2$ as the primal solution and $v(P) = 2$. $\bar{E}$ consists only of the point $\bar{t} = \frac{1}{3}\binom{1}{2}$.

For $z \in (0, 2)$ we have

$$z(t_2 - t_1) < 1 - t_1, \qquad t \in B$$

proving primal superconsistency and, by Theorem 6.9 in (D) and $v(P) = v(D) = 2$. One can of course directly construct an optimal $\bar{\mu}$ from elementary complementary slackness using $\bar{t}$, namely,

$$\bar{\mu}(t) = \begin{cases} 3, & \text{if } t = \bar{t}, \\ 0, & \text{otherwise.} \end{cases}$$

Alternatively, we can apply Theorem 6.11 upon observing that $O^+(L_{\geq}(Z^P, c, 0)) = \{0\}$, establishing superconsistency of (D). In addition, the role of program $(\tilde{D})$ (cf. (9b)) can be illustrated as

$$v(\tilde{D}) = \inf \left\{ \sum (1 - t_1)\mu(t) | \mu \in \mathbb{R}_+^{(B)}, \sum \mu(t) = 1 \text{ and } \sum (t_2 - t_1)\mu(t) = 0 \right\}.$$

With a little inspection one can determine a single "mass point" optimal solution $\tilde{\mu}$ according to

$$\tilde{\mu}(t) = \begin{cases} 1 & \text{if } t = \tilde{t} = \frac{1}{3}\binom{2}{2}, \\ 0 & \text{elsewhere,} \end{cases}$$

yielding $v(\tilde{D}) = \frac{1}{3} > 0$. Modifying the right side $1 - t_1$ of (P) to $k - t_1$, $k > \frac{2}{3}$, maintains $v(\tilde{D}) > 0$ and perfect duality. Thus, we are in a stable situation.

6.3. A closer look at perfect duality. Generally speaking, (P) and (D) are said to be in perfect duality if and only if (i) $v(P) = v(D)$ (including infinite values) or (ii) $v(P) = -\infty$ and $v(D) = +\infty$. In finite linear programming $v(P)$ and $v(D)$ are both attained when either is finite. This is not necessarily so in semi-infinite programming. However, by Theorem 6.5(ii) the closure of M_{n+1} is sufficient for perfect duality with attainment in (D).

Considering, instead of M_{n+1}, the cone

$$MH_{n+1} = \text{co}\,(M_{n+1} \cup \{e_{n+1}\})$$

with $e_{n+1} = (0, \ldots, 0, 1)^T \in \mathbb{R}^{n+1}$, a much finer analysis is possible. Observe that MH_{n+1} is closed if M_{n+1} is, but that the converse is false. For example, let $M_2 = \text{co}\{(-t^2, t)^T | t \in [0, 1]\}$. The boundary point $\binom{0}{1}$ is not in M_2, so M_2 is not closed. But with $\binom{0}{1}$ in the augmented cone MH_2, we see that MH_2 is closed.

THEOREM 6.14 [6.22]. *If (P) is consistent, the following are equivalent.*

(i) *For any $c \in \mathbb{R}^n$, (P) and (D) are in perfect duality with $v(D)$ attainment when finite.*

(ii) *MH_{n+1} is closed.*

(iii) *The coefficient set*

$$\left\{ \begin{pmatrix} a(t) \\ b(t) \end{pmatrix} \middle| \, t \in B \right\}$$

has a positive equivalent set (in the sense of [6.22]).

For many years canonical closure has meant that the coefficient set in (iii) is compact and that there exists an interior point z^0, i.e., $a(t)^T z^0 < b(t)$ for every $t \in B$. It has been known that canonical closure is sufficient for MH_{n+1} to be closed [6.10], [6.12], [6.13], [6.20]. Let us illustrate how condition (i) of Theorem 6.14 can be violated when MH_{n+1} is not closed.

Example 6.15. Let $B = [0, 1]$ and $Z^P = \{z | (-t)z \leq t^2, t \in B\} = \{z | z \geq 0\}$. For $c \in \mathbb{R}, c > 0$ we have $v(P) = v(D) = +\infty$, while for $c \leq 0$, $v(P) = v(D) = 0$. Hence, there is perfect duality, but because $v(D)$ is not attained for $c < 0$, we see that condition (i) is not satisfied.

Another sufficient condition for perfect duality, but which this time guarantees primal $v(P)$ attainment when finite, is for the convex cone

$$O^+(L_{\geq}(Z^P, c, 0)) = \{z | a(t)^T z \leq 0 \text{ for } t \in B, c^T z \geq 0\} \tag{6.10}$$

to be a line, including the degenerate line stated in (a_3) of Theorem 6.11 (cf. [6.23]). Adjoining a finite number of affine constraints to (P) poses no difficulties in stating sufficient conditions for primal or dual attainment. But (6.10), being only a line, does not give enough information about the range of variation in the objective function coefficients under which a finite supremum value is attained.

In [6.46] there appeared a dual pair of linear programs over closed convex cones in $\mathbb{R}^6$ for which $v(P) = v(D)$, both unattained.

6.4. A treatment of nonperfect duality in linear SIP. In finite linear programming Charnes introduced non-Archimedean field extensions into mathematical programming. He obtained a "complete regularization" of the primal/dual pair so that both problems are solvable with the same objective function value [6.9], i.e., the perfect duality case with attainment. The construction involves artificial variables and additional constraints. By examining the primal and dual solutions of the completely regularized problems one can determine which of the four mutually exclusive and collectively exhaustive (MECE) duality states prevail for the original finite linear programming pair (P, D):

(consistent, consistent) with common value,

(consistent and unbounded, inconsistent),

(inconsistent, consistent and unbounded), and

(inconsistent, inconsistent).

In linear semi-infinite programming the duality state characteristics must be refined and expanded. As illustrated in [2.29], there are now 11 MECE duality states which can occur, out of a conceivable 121. It is still an open question whether there exist complete regularizations of linear semi-infinite programs which can characterize the permissible duality states.

The classification theory has been extended to convex programs over linear topological spaces having the Hahn–Banach extension property; see [6.34], [6.39], [6.40].

The necessity of more duality state possibilities stems from "duality gap" phenomena, which arise when the values of primal and dual programs differ. We illustrate the occurrence of a duality gap from a favorite example of W. W. E. Wetterling, involving a fourth degree nonconvex polynomial having infimum zero but not attained.

Example 6.16. Define for a real parameter y the following function:

$$f(t_1, t_2, y) = (t_1t_2 - 1)^2 + (1 - y)t_2^2 \quad \text{for } t \in \mathbb{R}^2. \tag{6.11}$$

Then $\inf\{f(t_1, t_2, y) | t \in \mathbb{R}^2\} = 0$ whenever $y \leq 1$ with attainment only if $y = 1$.

The two variable polynomial (6.11) suggests a saddle value problem, where the inner minimization is the global minimization of $f(t_1, t_2, y)$:

$$\text{Find } \sup\{G(y) | y \in \mathbb{R}\}; \quad \text{where}$$

$$G(y) = \inf\{f(t_1, t_2, y) | t \in \mathbb{R}^2\}$$

defines the closed proper concave function on $\mathbb{R}$ given by

$$G(y) = \{0 \text{ if } y \leq 1 \text{ or } -\infty \text{ otherwise}\}.$$

The linear semi-infinite program equivalent is

$$v(P) = \sup z_1, \quad z_1, z_2, \in \mathbb{R} \quad \text{subject to}$$

$$z_1 + z_2t_2^2 \leq (t_1t_2 - 1)^2 + t_2^2 \quad \text{for } t \in \mathbb{R}^2.$$

In either form, note that (P) involves an intrinsically two-dimensional inner minimization when $z_2 < 1$. The immediate implication of this is that $v(P) = 0$. The generalized finite sequence dual of (P) is

$$v(D) = \inf \sum \{(t_1t_2 - 1)^2 + t_2^2\}\mu(t_1, t_2), \text{ from among } \mu \in \mathbb{R}_+^{(\mathbb{R}^2)} \text{ subject to}$$

$$\sum t_2^2\mu(t_1, t_2) = 0 \quad \text{and} \quad \sum \mu(t_1, t_2) = 1.$$

Observe that the dual program restricts the minimization of $f(t_1, t_2, 0)$ to the one-dimensional subspace $t_2 = 0$, with the consequent increase in extremal function value zero to 1, namely, $v(D) = 1$. This example really illustrates that a duality gap can arise when the duality itself restricts a higher-dimensional problem to a lower-dimensional one. This same phenomenon arises during the construction of saddle value problems of some ordinary convex programs: for example, replacing $f(t_1, t_2, y)$ of (6.11) with

$$\exp(-t_2) + y\{(t_1^2 + t_2^2)^{1/2} - t_1\}, \qquad t \in \mathbb{R}^2, y \geq 0 \quad \text{or}$$

$$\exp\left(-\sqrt{t_1 t_2}\right) + y t_2 \quad \text{for } t \geq 0, y \geq 0;$$

see [6.30].

One way of removing duality gaps is to extend the range of the variables in the linear inequality program. Instead of requiring (z_1, z_2) to be real numbers, permit (z_1, z_2) to range over an ordered extension of the reals such as polynomials in an indeterminate. It will not be necessary to consider division in the extension. To be precise the following "polynomial ring" is introduced.

DEFINITION 6.17. Let $\mathbb{R}[\Theta]$ denote the polynomial ring consisting of finite degree, real coefficient polynomials in an indeterminate Θ. A non-Archimedean ordering is derived by requiring $r < \Theta$ for all r in $\mathbb{R}$. A polynomial $p(\Theta) = \sum_{i=0}^{l} r_i \Theta^i$ is *positive* (*negative*) if the coefficient of the highest nonvanishing power of Θ is *positive* (*negative*). The polynomial $\sum_{i=1}^{l} r_i \Theta^i$ is termed the *infinite part* of $p(\Theta)$; see [6.33], [6.41], [6.42].

Let us illustrate the use of $\mathbb{R}[\Theta]$ to remove the duality gap in the Wetterling example. This amounts to restating (P) in terms of the extended variables:

$$v(P_\Theta) = \sup \tau; \quad \tau \in \mathbb{R}, \quad z_1(\Theta) \quad \text{and} \quad z_2(\Theta) \text{ in } \mathbb{R}[\Theta] \quad \text{subject to}$$

$$\tau - z_1(\Theta) \leq 0$$

$$z_1(\Theta) + z_2(\Theta) t_2^2 \leq (t_1 t_2 - 1)^2 + t_2^2 \qquad \text{for all } t \in \mathbb{R}^2.$$

An inspection of (P_Θ) will show that $v(P_\Theta) = 1$, attained, for example, with $\tau_* = 1$, $z_1^*(\Theta) = 1$, and $z_2^*(\Theta) = -\Theta$.

It is interesting that the Θ-polynomial solutions as illustrated above provide algebraic means of stating limiting feasible solutions to the original linear inequality system in real variables. This important connection will be made after first introducing the idea of asymptotic consistency for both primal and dual programs.

DEFINITION 6.18 (asymptotic consistency). Program (P) is *asymptotically consistent* (AC) if and only if every finite subsystem of $a(t)^T z \leq b(t)$, for all $t \in B$, is consistent. A sequence of generalized finite sequences $\{\mu^{(n)}\}_n$ is termed *asymptotically feasible* for Program (D) if and only if each $\mu^{(n)} \geq 0$ and

$$\lim_n \sum a(t) \mu^{(n)}(t) = c.$$

(D) itself is *asymptotically consistent* if it has asymptotically feasible solutions.

The connection between (P) asymptotic consistency and Θ-polynomial solutions appeared in the following theorem.

THEOREM 6.19 (see [6.33]). *Assume that* (P) *is* AC. *Then* (P) *has a feasible point* $z(\Theta) \in (\mathbb{R}(\Theta))^n$, *where each coordinate* $z_i(\Theta)$, $i = 1, \ldots, n$ *is a polynomial in* Θ *of degree at most* n.

The infinite system, $z \in \mathbb{R}^n$, $z_1 \geq k$, $-k z_1 + z_2 \geq 0, \ldots, -k z_{n-1} + z_n \geq 0$ for $k = 1, 2, \ldots$, shows that polynomials of full degree n may be necessary.

A main application of Theorem 6.19 is the construction of an asymptotic extension (P_Θ) of Program (P) as illustrated already in the above example:

$$v(P_\Theta) = \sup \tau; \qquad \tau \in \mathbb{R},\, z(\Theta) \in (\mathbb{R}[\Theta])^n \quad \text{subject to}$$

$$\tau - c^T z(\Theta) \le 0 \quad \text{and} \quad a(t)^T z(\Theta) \le b(t) \quad \text{for } t \in B.$$

Note that τ is required to be real, so that the supremum is over the reals only. Feasible solutions are pairs $(\tau, z(\Theta))$ lying in $\mathbb{R} \times (\mathbb{R}[\Theta])^n$. The above theorem has also been extended by Blair [6.2], and in a convex analysis setting by Borwein [6.5].

COROLLARY 6.20 (see [6.41], [6.42]). *Programs (P_Θ) and (D) are in perfect duality.*

In Theorem 6.11 we saw that with the condition $O^+(L_\ge(Z^P, c, \kappa)) = \{0\}$ good stability results followed. If this condition is substantially strengthened to $O^+(Z^P) = \{0\}$, i.e., $z = 0$ whenever $a(t)^T z \le 0$ for all $t \in B$, then it can be shown that (P_Θ) is consistent if and only if (P) is, and in this case $v(P_\Theta) = v(P)$. Assuming that (P) is consistent, is, of course, independent of the vector c. One can, therefore, use the perfect duality between (P_Θ) and (D) to prove that for every $c \in \mathbb{R}^n$ it is the case that $v(P) = v(D)$ and that $v(P)$ is attained.

There is a fundamental result about any linear inequality system which is inconsistent but asymptotically consistent.

THEOREM 6.21 (see [6.2]). *Assume $z \in \mathbb{R}^n$, $a(t)^T z \le b(t)$, for $t \in B$ of (P), is inconsistent but* AC. *Then there exists $c \in \mathbb{R}^n$ such that given any positive integer n, there is a finite subset B_n of B such that $a(t)^T z \le b(t)$ for $t \in B_n$ implies $c^T z \le -n$.*

Vectors c of Theorem 6.21 have been termed *descent vectors* (for the minimization associated with the dual program (D)). In general, the set of descent vectors is a convex cone not containing zero, which is contained in M_n, and which contains the relative interior of M_n, $ri(M_n)$, possibly strictly [6.3], [6.42].

To illustrate a set of descent vectors, consider the following linear inequality representation of the hypograph of the concave function $\log(z_1)$, $z_1 > 0$:

$$-z_1 \frac{1}{t} + z_2 \le \log(t) - 1 \quad \text{for all } t > 0.$$

Adjoin to the above system the single inequality $z_1 \le 0$. The total system is now inconsistent, buy any finite subsystem is consistent. In this case M_n for $n = 2$, is

$$\mathrm{co}\left(\left\{\begin{pmatrix} -\frac{1}{t} \\ 1 \end{pmatrix} \mid t > 0\right\} \cup \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right),$$

i.e., $M_2 = \{(x_1, x_2) | x_2 \ge 0\} \backslash \{(x_1, 0) | x_1 < 0\}$. We verify that for any $c_1 \in \mathbb{R}$ and $c_2 > 0$, (c_1, c_2) is a descent vector.

Given any integer n, there clearly exists $\varepsilon > 0$ such that (i) $c_2(\log(\varepsilon) - 1) < -n$ and (ii) $c_1 + \frac{1}{\varepsilon} c_2 > 0$. Consider the two inequality subsystems for the Blair condition (where $t = \varepsilon$):

$$-\frac{z_1}{\varepsilon} + z_2 \le \log(\varepsilon) - 1,$$

$$z_1 \le 0.$$

Multiplying the first inequality by c_2, the second by $c_1 + c_2/\varepsilon$, and adding gives

$$c_1 z_1 + c_2 z_2 \le c_2(\log(\varepsilon) - 1) \le -n.$$

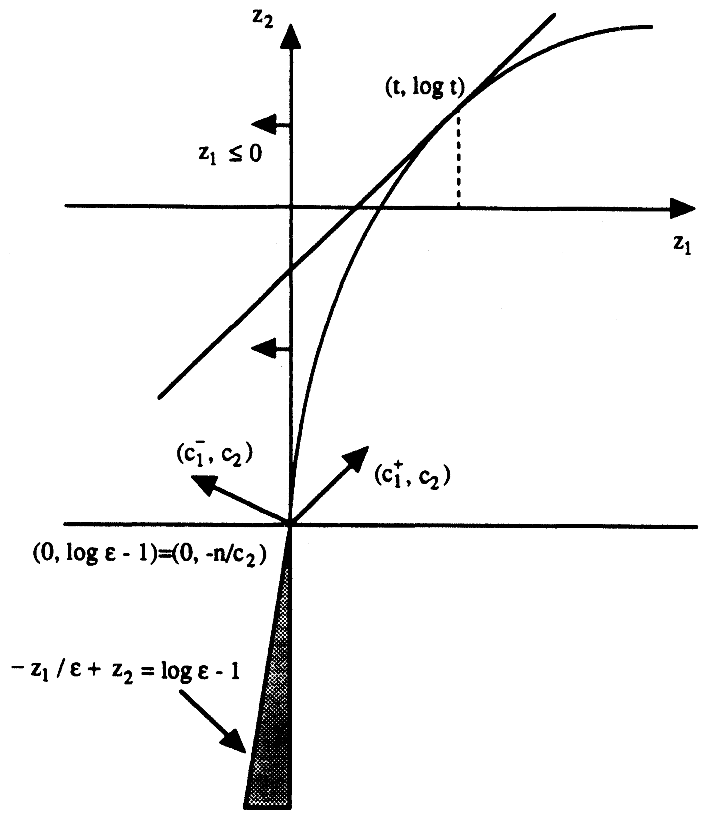

FIG. 6.1. *Illustration of the set of descent vectors for a logarithmic function. Example: Given any positive integer n, set $\varepsilon = \exp(1 - n/c_2)$. Shaded polyhedral cone lies in the two half spaces; $c_1^{\pm} z_1 + c_2 z \leq -n, c_1^+ > 0, c_1^- < 0$ arbitrary.*

This shows that the set of descent vectors $M_{DV} = \{(x_1, x_2) | x_2 > 0\}$, and that for this example, subsystems of two special linear inequalities work for the Blair condition. We illustrate geometrically in Fig. 6.1.

In this case the set of descent vectors happens to coincide with the relative interior of M_2.

Remark 6.22. There is an equivalent, duality theoretic way of characterizing descent vectors, related to the dual pair (P, D). A descent vector c places the pair in duality state 7, which is illustrated by [2.29]. See also [6.39], namely, (P) is improperly asymptotically consistent while (D) is consistent and unbounded.

Earlier we considered the perturbed program $(\tilde{D}_G)$ of (D), namely,

$$v(\tilde{D}_G) = \inf w; \quad w \in \mathbb{R} \quad \text{subject to}$$

$$\begin{pmatrix} c \\ w \end{pmatrix} \in cl(M_{n+1}).$$

However, rather than considering arbitrary paths in M_{n+1} converging to $\binom{c}{w}$, one need only follow along a ray determined by descent vectors for the partially homogenized system given as follows:

$$a(t)^T z + z_{n+1} b(t) \le 0 \quad \text{for all } t \in B,$$

$$z_{n+1} \le 0,$$

$$c^T z + z_{n+1} w \le 1,$$

where $w \in \mathbb{R}$ is a fixed parameter.

The specific, descent ray refinement of Program $\tilde{D}_G$ is as follows:

$$v(\tilde{D}) = \inf w \quad \text{subject to } w \in \mathbb{R} \quad \text{and}$$

$$\begin{pmatrix} v \\ v_{n+1} \end{pmatrix} \in \mathbb{R}^n \times \mathbb{R}, \quad \text{which satisfy:}$$

given any $\varepsilon > 0$ there exists $\mu \in \mathbb{R}_+^{(B)}$ such that

$$\sum a(t)\mu(t) = c + \varepsilon v \quad \text{and} \quad \sum b(t)\mu(t) \le w + \varepsilon v_{n+1}.$$

Feasible points are now the triple

$$\begin{pmatrix} w \\ v \\ v_{n+1} \end{pmatrix}$$

lying in $\mathbb{R} \times \mathbb{R}^n \times \mathbb{R}$. Corresponding to our review of descent vectors above, the subvector

$$V = \begin{pmatrix} v \\ v_{n+1} \end{pmatrix}$$

of a feasible point is a descent vector. As expected, Programs (P) and (D) are in perfect duality.

Descent vectors for $(\tilde{D})$ can be constructed as follows [6.41].

For a fixed w coming from a $\tilde{D}$-feasible point, in particular, the optimum value $w = v(\bar{D})$, consider the collection U in $\mathbb{R}^{n+1}$:

$$U = \left\{ \begin{pmatrix} a(t) \\ b(t) \end{pmatrix} \middle| \, t \in B \right\} \cup \left\{ \begin{pmatrix} -c \\ -w \end{pmatrix} \right\} \cup \left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}.$$

Let $\{Q_i\}_{i=1}^l$ be a subset of U forming a linear basis for the linear space (over the reals) spanned by U. Let $\delta_i > 0$, $i = 1, \ldots, l$, otherwise arbitrary, and set

$$\begin{pmatrix} v \\ v_{n+1} \end{pmatrix} = \sum_{i=1}^{l} \delta_i Q_i.$$

Then

$$\begin{pmatrix} w \\ v \\ v_{n+1} \end{pmatrix}$$

is a feasible point for Program $(\tilde{D})$.

7. Numerical methods. We now turn to the question of how to compute numerically a solution of the linear or nonlinear semi-infinite problem:

$$\text{(P)} \qquad \text{Maximize } F(z) \text{ subject to } z \in Z^P, \text{ with} \quad Z^P = \{z|g(z,t) \leq 0, t \in B\} \subset \mathbb{R}^n.$$

As we have seen in the foregoing sections the parametric problem

$$(O(z)) \qquad v(z) := v(O(z)) = \max\{g(z,t), t \in B\}$$

plays a crucial role in semi-infinite programming. Obviously $v(z)$ is continuous but not differentiable in general. To check the feasibility of a given $z \in \mathbb{R}^n$ requires checking if $v(z) \leq 0$. Moreover, in §4, we showed that (P) can be reduced to a problem $(P_{\text{red}}(\bar{z}))$ with only a finite number of constraints if some regularity conditions are given which permit the description of local solutions of $(O(z))$ as nice functions of z.

Writing (P) in the obviously equivalent form

$$\text{(P)} \qquad \text{Maximize } \{F(z)|v(z) \leq 0\}$$

shows that we could treat semi-infinite optimization in the framework of nondifferentiable optimization. This point of view has been taken by Polak in his review paper [7.43]. Descent methods, as the only generally applicable class of methods for nondifferentiable optimization, are discussed extensively in this paper. It should be noted that these methods are also applicable in cases where g is nondifferentiable with respect to t. On the other hand, descent methods in general have rates of convergence of at most one and consequently require a large number of function evaluations. This could be very costly because evaluating $v(z)$ requires finding a global solution of $(O(z))$. Moreover, to compute sufficiently good directions of descent one has to again take into account infinitely many constraints. In the implementable version of algorithms in [7.43], B is replaced by finite grids of successively finer meshsizes to guarantee convergence to a solution of the semi-infinite problem.

In summary, these methods are robust but rather expensive with respect to computing time. The interested reader is referred to [7.43] and the references given there for this approach. For descent methods in a more general setting see [7.34], [7.56]. We also have included a number of interesting papers for these methods in the reference list, which could not be reviewed here.

In this section, our main emphasis will be on superlinear convergent methods, which, however, require additional smoothness of g with respect to the variable t. All the algorithms to be considered now replace (P) by (a sequence of) finite programming problems, i.e., problems with only a finite number of constraints. These are solved by applying appropriate linear or nonlinear programming algorithms, for which we refer to an extensive literature. The general scheme of an algorithm for (P) is visualized in Fig. 7.1.

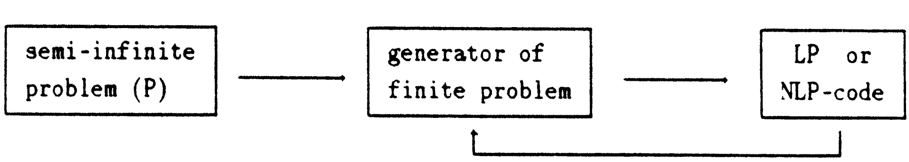

FIG. 7.1. *General scheme of* SIP *algorithms.*

According to the way finite problems are generated, we can roughly distinguish three types of methods: (A) exchange methods, (B) discretization methods, and (C) methods based on local reduction.

Throughout this section, if $\bar{B} \subset B$ is a given subset of B, by $(P(\bar{B}))$ we denote the problem

$$(7.1) \qquad \begin{aligned} &(P(\bar{B})) \quad \max\{F(z) | z \in Z^P(\bar{B})\}, \quad \text{with} \\ &Z^P(\bar{B}) = \{z | g(z,t) \leq 0, t \in \bar{B}, \|z\|_\infty \leq \Gamma\}, \end{aligned}$$

where we assume that $\Gamma < \infty$ exists such that

$$(7.2) \qquad Z^P \subset \{z | \|z\|_\infty \leq \Gamma\}.$$

This implies that Z^P is a compact set contained in a ball around zero with radius Γ. The additional (linear) constraints $\|z\|_\infty \leq \Gamma$ ensure that $(P(\bar{B}))$ has a solution for every $\bar{B} \subset B$.

7.1. Exchange methods. We use the term "exchange method" to denote all methods whose general step is determined algorithmically as follows.

ALGORITHM 7.1 (conceptional exchange method).

Step (i). Given $B_{i-1} \subset B, |B_{i-1}| < \infty$. Determine a solution z^i of $(P(B_{i-1}))$ and (approximate) solutions $t^1, \dots, t^{r_i}$ of the subproblem

$$(O(z^i)) \qquad \max\{g(z^i, t) | t \in B\}.$$

Stop, if $g(z^i, t^l) \leq 0, l = 1, \dots, r_i$. Otherwise choose B_i such that

$$(7.3) \qquad B_i \subset B_{i-1} \cup \{t^1, \dots, t^{r_i}\}.$$

Remark. The notation "exchange algorithm" refers to the fact that in every step a number of new constraints (corresponding to $t^1, \dots t^{r_i}$) are added and some of the old constraints, corresponding to $t \in B_{i-1}$, may conceivably be deleted, i.e., an exchange of constraints takes place.

For the following simple rule

$$(7.4) \qquad B_i = B_{i-1} \cup \{t^1, \dots t^{r_i}\}$$

we do not really have an exchange because no constraints are deleted. Denoting again by $v(z)$ the value function of $(O(z))$, one has the following result.

THEOREM 7.2. *Assume that there is a* $\Gamma > 0$ *with* (7.2) *and in every step* $v(z^i) = g(z^i, t^j)$ *holds for at least one* $j \in \{1, \dots, r_i\}$. *Then either Algorithm* 7.1 *with exchange rule* (7.4) *stops after a finite number of steps with a solution to* (P) *or the sequence* $\{z^\nu\}$ *has at least one point of accumulation and each of these solves* (P).

Proof. As $\{z^\nu\}$ is bounded, a point $\bar{z}$ of accumulation always exists. Without restriction assume $z^\nu \to \bar{z}$. Due to $B_\nu \subset B$ we have $v(P) \leq F(\bar{z})$. Therefore, it suffices to show that $\bar{z} \in Z^P$. Assume to the contrary, that $v(\bar{z}) > 0(v(z)$ the optimal value function of $(O(z)))$. Now, if $v(z^\nu) = g(z^\nu, \tilde{t}^\nu)$, we have

$$\begin{aligned} v(\bar{z}) &= v(z^\nu) + [v(\bar{z}) - v(z^\nu)] \\ &= g(z^\nu, \tilde{t}^\nu) + [v(\bar{z}) - v(z^\nu)] \\ &\leq [g(z^\nu, \tilde{t}^\nu) - g(\bar{z}, \tilde{t}^\nu)] + [v(\bar{z}) - v(z^\nu)], \end{aligned}$$

where we have used $\bar{z} \in \cap_{\nu=1}^\infty Z^P(B_\nu)$ because of $B_\nu \subset B_{\nu+1}$ for all ν.

Due to the continuity of v and g, the right-hand side becomes arbitrarily small for $\nu \to \infty$, which is a contradiction to $v(\bar{z}) > 0$.

For linear problems this has been proved in [1.8] (for $r_i = 1$ see also [7.3], [7.20], and for a slightly more general version see [7.32]).

Using (7.4) leads to a considerable increase in the number of constraints to be included in $(P(B_i))$. For algorithms for linear problems where exchanges occur, see [7.11], [7.30], [7.40], [7.53]. For the special case of linear Chebyshev approximation, Remes algorithms, and generalizations see [1.8], [7.37], [7.59]. Apart from the Haar case, either no proof of convergence has been given or the algorithm has to be extended by rather involved, additional strategies.

It is obvious, for the case $r_i = 1$, that the algorithm applied to linear (P) can be interpreted as being the cutting plane algorithm [7.35] for solving (P), considered as a convex problem (cf. [7.25]). Therefore, with the exception of the strongly unique case, (see [7.29]) good local convergence may not be expected (cf. [7.6]).

The determination of a global solution of $(O(z^i))$ is very costly for cases where $B \subset \mathbb{R}^m$, $m \geq 2$. This motivates strategies with $r > 1$, i.e., inclusion of all local solutions found. Another strategy (restricted to linear or convex problems) is to work with a $t = t^1$ such that $g(z^i, t^1) > 0$ only, i.e., an arbitrary violated constraint, and to try instead to cut off a piece of the feasible region containing no solution. In [7.16] such a strategy—originally proposed for finite convex problems [7.6]—is developed and has been applied to geometric programming with remarkable success [7.17]. The general convex case is done in [7.36], with numerical experiments performed on complex approximation problems; see also [2.18], [2.55], [2.77], [2.78].

7.2. Discretization methods. For a given vector $h \in \mathbb{R}^m$ of stepsizes $h_j > 0$, $j = 1, \ldots, m$, and a fixed $t^0 \in \mathbb{R}^m$, we define the grid

$$G_h = \{t \mid (t - t^0)_j = \alpha_j h_j, \alpha_j \in \mathbb{Z}, j = 1, \ldots, m\} \tag{7.5}$$

and

$$B_h = B \cap G_h. \tag{7.6}$$

In [4.5], [7.26], [7.27] for the linear and quadratic case algorithms are given which approximate a solution of (P) by solving a sequence of problems $(P(B_{h^i}))$. This concept may be applied to general nonlinear problems.

ALGORITHM 7.3 (conceptual discretization method).

Step (i). One is given h^i, a *selection* $\tilde{B}_{h^i} \subset B_{h^i}$ and a *solution* $\tilde{z}^i$ of problem $(P(\tilde{B}_{h^i}))$.

(a) Set $h^{i+1} = (1/n_i)h^i (n_i \in \mathbb{N}, n_i \geq 2)$.

(b) Select $\tilde{B}_{h^{i+1}} \subset B_{h^{i+1}}$ (on the basis of $\tilde{z}^i$, $\tilde{B}_{h^i}$, and eventually previous *trials* $\tilde{B}_{h^{i+1}}$ and $\tilde{z}^{i+1}$).

(c) Compute a *solution* $\tilde{z}^{i+1}$ of $(P(\tilde{B}_{h^{i+1}}))$. If $\tilde{z}^{i+1}$ is feasible for $(P(B_{h^{i+1}}))$ (within a given accuracy) continue with (d), otherwise repeat (b).

(d) If $i > i_0$ (prechosen number of refinement steps) stop. Otherwise Step $(i + 1)$.

Due to Z^P, $Z^P(B_{h^i}) \subset \{z \mid \|z\|_\infty \leq \Gamma\}$ (see (7.1), (7.2)) it follows from Theorem 4.4 that for linear problems every point of accumulation of solutions z^i of $(P(B_{h^i}))$ solves (P). If the accuracy criterion is sharpened to ensure that $\tilde{z}^i$ is feasible for $(P(B_{h^i}))$, then the same holds for $\{\tilde{z}^i\}$.

An essential point with respect to efficiency is to use as much information as possible from previous grids when solving $(P(\tilde{B}_{h^i}))$, since $B_{h^i} \subset B_{h^{i-1}}$, z^{i-1} is generally a good starting point in solving $(P(\tilde{B}_{h^i}))$. In the linear case it is easy to ensure that $\tilde{z}^{i-1}$ is a feasible vertex for the new problem; cf. [7.26].

Information from previous grids also should be used in Step (b) in deleting constraints from $(P(B_{h^i}))$. The most obvious way of selecting $\tilde{B}_{h^i} \subset B_{h^i}$ is a choice

$$\tilde{B}_{h^i} \supset B^{\gamma}_{h^i} := \{t \mid t \in B_{h^i}, g(\tilde{z}^{i-1}, t) \geq -\gamma\} \tag{7.7}$$

with $\gamma > 0$ being some chosen threshold. In a certain sense the determination of $B^{\gamma}_{h^i}$ replaces the global search for maxima t^j of $(O(z^i))$ in exchange methods. The hope is that by taking whole clusters of points in G_{h^i} this costly step has to be performed only a few times.

The choice of γ in (7.7) is crucial: A γ which is too large leads to many constraints in $(P(\tilde{B}_{h^i}))$. However, choosing γ too small, may have the same effect in subsequent problems, because parts of B may be overlooked. In [7.26], [7.27] $\gamma = 0$ is taken and (to circumvent the latter phenomenon) additional "critical" points are included in $\tilde{B}_{h^i}$. It is demonstrated with examples in [7.26] that problems (with $m = 1, 2$) can be solved on very fine grids rather efficiently by these methods, requiring only the solution of a small number of finite problems with rather few constraints.

Interesting in the context of discretization algorithms is the paper of Polak and He [7.46]. There the question of how to refine the grid is answered (i.e., how to choose numbers n_i in substep (a)) such that the rate of convergence of certain ascent methods [7.45] is preserved, when for each i such an ascent step is performed on $((P(B_{h^i})))$. A combination of this approach with the above grid selection strategies could be interesting for future investigations. A basis for such approaches is provided in [7.44].

7.3. Methods based on local reduction. Conceptually, these methods can be described as follows.

ALGORITHM 7.4 (conceptual reduction method).

Step (i). One is given z^i (not necessarily feasible).

(a) Determine all *local maxima* $t^1, \ldots, t^{r_i}$ of $(O(z^i))$.

(b) Apply k_i steps of a finite programming algorithm to the *reduced problem* (cf. Problem 4.9).

$$(P_{\text{red}}(z^i)) \qquad \max \quad \{F(z) \mid G^l(z) \leq 0, l = 1, \ldots, r_i\} \quad \text{with } G^l(z) = g(z, t^l(z)).$$

Let $z^{i,j}$, $j = 1, \ldots, k_i$ be the iterates.

(c) Set $z^{i+1} = z^{i,k_i}$ and continue with Step $(i + 1)$.

Substep (a) is very costly as it requires a global search for maxima of $g(z^i, t)$ on $B \subset \mathbb{R}^m$. The overall strategy must be to avoid execution of this step as much as possible.

Substep (a) tacitly assumes that there are only finitely many maxima of $(O(z^i))$. If this is not the case, a basic assumption for reduction fails to hold and another method (for instance, a discretization method) should be used. Note that, according to Remark 4.18, this case may be considered as degenerate.

Observe that in substep (b), the parametric problem $(O(z))$ has also to be considered in evaluating the constraints G^l. This, however, requires only local searches, which normally can be performed efficiently (using Newton's method, for example). We address the accuracy required for $t^l(z)$ in Theorem 7.10 below.

In substep (b) NLP methods with superlinear rate of convergence are advisable to keep the number of substeps (a) low. Up to now almost exclusively Sequential Quadratic Programming (SQP) methods have been used in the above context: Wilson's method [7.28], [7.31], [7.33], SQP with augmented Lagrangians and quasi-Newton update of its Hessian [7.41], "globalized"

SQP (see below) with exact Hessian of the Lagrangian [7.5], [7.61] or with quasi-Newton update of this Hessian [7.14]. Some details of the latter algorithm will be given below. For other variants of SQP, see [7.50], [7.52], [7.54], [7.58].

In the first instance SQP methods are locally convergent. In the earlier papers [7.7], [7.19], [7.20], [7.41] but also in [7.40] hybrid techniques are used, combining robust globally convergent ascent methods with the SQP approach. This means the question of criteria for switching from one method to another must be addressed. Easier and more convenient seem the globalization techniques for SQP methods introduced by Han [7.22], which are based on a stepsize rule to increase a certain merit function. These merit functions are defined by means of exact penalty functions. In [7.5], [7.14], [7.61] the merit function

$$\Theta_p(z) := F(z) - p\sum_{l=0}^{r_i}[G^l(z)]_+, \tag{7.8}$$
$$[G^l(z)]_+ = \max\{0, G^l(z)\}$$

is used, which is known to be an exact penalty function for p sufficiently large. The penalty term in (7.8) is termed as an L_1-penalty. Note that (7.8) again is based on the possibility of local reduction. A proof of global convergence in [7.14], for $k_i \equiv 1$ (cf. substep (b) in Algorithm 7.4) assumes that along the piecewise linear path connecting the iterates $z^{i,j}$, all local maxima t^l of $(O(z))$ are at least strongly stable (cf. [4.6], [8.21]). This weakens considerably the assumptions made in [7.5]. For $k_i > 1$ this same result can be obtained by a modification of Algorithm 7.4 in that a backtracking technique is added requiring little additional effort (cf. [2.30]).

Alternative penalty functions for semi-infinite problems are given in [7.4], [7.44], [7.52], [7.54]. In [7.54] the L_∞-penalty term

$$-\,p\max_l\{[G^l(z)]_+\} \tag{7.9}$$

is used to define a variant of SQP (a sort of trust region method) by including the penalty term (7.9) in the objective of the quadratic program defining the step. In [7.52] the approach is similar, but the reduced problem is defined more generally, admitting discontinuous $t^l(z)$ with resulting $G^l(z)$ only Lipschitz continuous. This approach could be interesting in dealing with exceptional points. The penalty functions in [7.4], [7.42] are not based on reduction and, therefore, require integration over the regions of B where $g(z,t) > 0$. Up to now no algorithms seem to be implemented using these penalty functions.

Historically, it is interesting that the second algorithm of Remes for Chebyshev approximation may be interpreted as a special case of the above class, applied and adapted to problems with strongly unique solutions (cf. [7.29], [7.64]). An old version of the reduction algorithm for linear semi-infinite problems is given in [2.24], [7.7], [7.20], which can be interpreted as Newton's method applied to the Karush–Kuhn–Tucker system of (P). In [7.23] this has been extended to nonlinear problems, and to a special class in [7.21].

To be explicit we will specify substep (b) in Algorithm 7.4 by use of a special SQP version ("SQP augmented Lagrangian BFGS method"), and discuss the local convergence properties of the resulting, still conceptual Algorithm 7.5.

ALGORITHM 7.5 (conceptual reduction method with SQP solver).

Step (i). Given z^i and a positive definite matrix B_i.

(a) Determine all *local maxima* $t^1, \dots, t^{r_i}$ of $(O(z^i))$.

(b) Set $z^{i,0} = z^i$, $B_{i,0} = B_i$. For $j = 1, \dots, k_i$ do $(b_1) - (b_3)$.

(b_1) Compute a *solution* s^j and an *optimal multiplier vector* $\lambda^{i,j} \geq 0$ of the quadratic programming problem

$$\text{maximize } F_z^T(z^{i,j-1})s + \tfrac{1}{2}s^T B_{i,j-1}s$$

$$\text{subject to } G^l(z^{i,j-1}) + G_z^l(z^{i,j-1})^T s \le 0, l = 1, \dots, r_i$$

$$\text{with } G^l(z) := g(z, t^l(z)).$$

(b_2) Compute a *step length* α_j (see below).

(b_3) Let $z^{i,j} = z^{i,j-1} + \alpha_j s^j$, $B_{i,j} = \text{BFGS}(\alpha_j s^j, y^j, B_{i,j-1})$, where

$$y^j = L_z^c(z^{i,j}, \lambda^{i,j}) - L_z^c(z^{i,j-1}, \lambda^{i,j})$$

with the *augmented Lagrangian* (of $(P_{\text{red}}(z^i))$)

$$L^c(z, \lambda) = F(z) - \sum_{l=1}^{r_i} \lambda_l G^l(z) - \frac{c}{2}\sum_{l=1}^{r_i}(G^l(z))^2$$

(c a prechosen parameter; see Assumption 7.8).

(c) Set $z^{i+1} = z^{i,k_i}$, $B_{i+1} = B_{i,k_i}$ and continue with step $(i + 1)$.

BFGS(s, y, B) denotes the common Broyden–Fletcher–Goldfarb–Shanno update formula which computes an updated matrix B_+ according to

$$B_+ = B + \frac{yy^T}{y^T s} - \frac{(Bs)(Bs)^T}{s^T Bs}.$$

The step length α_j is computed according to the above remark following Algorithm 7.4. The strategy should be such that $\alpha_j = 1$ for $j > j_0$, j_0 some index, otherwise the good local convergence (see below) could be destroyed (cf. [7.2]).

To investigate local convergence of Algorithm 7.5, we take $\alpha_j \equiv 1$ and make Assumptions 7.6–7.8 on the solution $\bar{z}$ of the semi-infinite problem (P).

Assumption 7.6. F, g, and h^j are twice continuously differentiable with Lipschitz continuous second-order derivatives. Let $\bar{t}^l$, $l \in L$, denote all local solutions of

$$(O(\bar{z})) \qquad \text{maximize } g(\bar{z}, t), \qquad t \in B = \{t | h^j(t) \le 0, j \in M\}.$$

Then for all $l \in L$, it is assumed that $\bar{t}^l$ satisfies Assumption 4.14.

Note that by Theorem 4.15, Assumption 7.6 ensures the existence of continuously differentiable $t^l(z)$ in a neighborhood of $\bar{z}$.

Assumption 7.7. In addition to Assumption 7.6 the strong sufficient second-order optimality condition of Theorem 5.1 (c) holds.

Assumption 7.8. c has been chosen such that the Hessian $L_{zz}^c(\bar{z}, \bar{\lambda})$ of the augmented Lagrangian is positive definite on $\mathbb{R}^n$.

Then we have the following theorem on the local convergence of Algorithm 7.5.

THEOREM 7.9. *Assume that in the solution $\bar{z}$ of the semi-infinite programming problem (P) Assumptions* 7.7 *and* 7.8 *hold. Then there exist δ_z, δ_B such that for z^0, B_0 with*

$$\|\bar{z} - z^0\| < \delta_z, \qquad \|L_{zz}^c(\bar{z}, \bar{\lambda}) - B_0\| < \delta_B$$

the sequence

$$\begin{pmatrix} z^{i,j} \\ \lambda^{i,j} \end{pmatrix}$$

converges q-superlinearly to

$$\begin{pmatrix} \bar{z} \\ \bar{\lambda} \end{pmatrix}$$

and $z^{i,j}$ *converges q-superlinearily to* $\bar{z}$.

Sketch of proof. From Assumption 7.7 one concludes that in a neighborhood $U_{\bar{z}}$ of $\bar{z}$, (P) becomes equivalent not only to $(P_{\text{red}}(\bar{z}))$ (Theorems 4.10 and 4.15), but also to the equality constrained problem:

$$(P_{\text{red}}(\bar{z})) \qquad \text{Maximize } F(z) \text{ subject to } G^l(z) \leq 0, \qquad l \in \bar{L} = \{l \mid g(z, \bar{t}^l) = 0\}.$$

The G^l are not only twice continuously differentiable (Theorem 4.15), but also have Lipschitz continuous second-order derivatives due to Assumption 7.6. Moreover, starting with z^0, B_0 sufficiently close to $\bar{z}$, $L^c_{zz}(\bar{z}, \bar{\lambda})$, respectively, Algorithm 7.5 generates the same sequence as the SQP generalized Lagrangian BFGS method applied to $(P_{\text{red}}(\bar{z}))$.

Assumptions 7.7 and 7.8 immediately imply that the assumptions of [7.9, Cor. 5.5] hold, and this corollary gives the required q-superlinear convergence results.

We emphasize that the assumptions of Theorem 7.9 are not very restrictive. From [4.6] it follows that Assumption 7.7 may be considered to hold generically, whereas Assumption 7.8 can be enforced by choosing c sufficiently large.

Of course, the SQP version used in Algorithm 7.5 just provides one example. Instead, any other SQP method from nonlinear programming could be used, and the corresponding results on convergence be translated similarly to Theorem 7.9.

Up to now we have assumed that the $t^l(z)$—and this means the functions $G^l(z) = g(z, t^l(z))$—are evaluated exactly. Note, that an error in $t^l(z)$ implies an inexact update of the matrices $B_{i,j}$ in the SQP method. This raises the question about what errors in the approximations $\tilde{t}^l(z)$ of $t^l(z)$ might be tolerable without destroying q-superlinear convergence of Algorithm 7.5. The following theorem has been proved in [7.14] (see also [7.15]).

THEOREM 7.10. *Consider Algorithm* 7.5 *with approximations* $\tilde{t}^l(z)$ *to* $t^l(z)$ *and* $c = 0$ *(i.e.,* L^c *becomes the Lagrangian itself). Assume that Assumption* 7.7 *holds and Assumption* 7.8 *with* $c = 0$.

Assume further that there exist a neighborhood $U_{\bar{z}}$ *of* $\bar{z}$ *and a* $K > 0$ *such that for all* $z \in U_{\bar{z}}$ *and next iterate* $z_+ \in U_{\bar{z}}$ *we have*

$$\|t^l(z_+) - \tilde{t}^l(z_+)\| + \|t^l(z) - \tilde{t}^l(z)\| \leq K\sigma(z, z_+)\|z - z_+\|, \tag{7.10}$$
$$(\sigma(z, z_+) = \max\{\|z_+ - \bar{z}\|, \|z - \bar{z}\|\}).$$

Then the q-superlinear convergence results of Theorem 7.9 *are preserved.*

Remark. Inequality (7.10) states that asymptotically $\tilde{t}^l$ approximates t^l faster than z^l approximates $\bar{z}$. For instance, $||t^l(z) - \tilde{t}^l(z)|| = O(||z - \bar{z}||^2)$ would ensure (7.10) in an appropriate neighborhood, but not, for instance, $||t^l(z) - \tilde{t}^l(z)|| = O(||z - \bar{z}||)$. As a consequence, $t^l(z)$ must be determined very accurately. This is in accordance with results of Polak et al. (cf., for instance, [7.48]), where it is shown that superlinear convergence of an SQP algorithm in combination with successively finer discretization can be maintained only by reducing the meshsize (and that means the accuracy of $t^l(z)$) in a similar ratio as indicated by inequality (7.10) in Theorem 7.10.

Example 7.11. Let us close this section by an example which gives an indication why Algorithm 7.5 is superior to the approach of discretizing B and applying SQP to the discretized problem.

TABLE 7.8
k: number of spline knots ($z \in \mathbb{R}^n$, $n = k+4$); i: number of SQP iterations; ac: average number of constraints considered in the reduced problems; mc: maximal number of constraints considered in the reduced problems.

k	i	ac	mc
7	17	87	196
15	11	57	92
31	9	75	120
63	8	135	245

In Example 2.1, a model for robot trajectory planning was presented, which leads to a problem of minimizing a linear function $c^T z$ subject to

$$g_l^{i,j}(z, t) \leq 0, \quad t \in [0, 1], \quad i = 1, \dots, R, \quad j = 1, 2, 3, \quad l = 1, 2.$$

For a specific robot with $R = 3$ free coordinates computations have been reported in [2.32].

A typical result is given in Table 7.8 (computed using the Harwell subroutine VF02AD based on [7.51] as NLP-solver).

Comparing this formulation with a mere discretization having only three additional discretization points between the knots, the latter, for $k = 63$, would require solving a problem with 4608 constraints (two thirds of them nonlinear) by SQP instead of a maximal number of 245 per SQP step in the above approach.

Remark 7.12. It is difficult to render advice concerning the question of which method should be selected for which type of problem. The reason is that on the one hand, various types of methods have been applied mostly to special problem classes while on the other hand, the total reported numerical experience on really hard problems is still very restricted. So we prefer to set forth a number of rather general remarks, extracted from our own restricted experience, on the three types of methods considered in this paper, which could be helpful.

Exchange (including cutting plane) methods almost exclusively have been applied to linear or convex SIP. In these cases, the methods usually yield low accuracy approximations rather efficiently, which is enhanced if multiple exchange is used. Due to the slow convergence it is not advisable to require high accuracy.

Discretization methods have a drawback. When the index set B is higher-dimensional, i.e., $B \subset \mathbb{R}^m$, $m \geq 2$, the number of constraints which have to be considered in the subproblems tends to become very large. In these cases the methods become inefficient.

Both exchange and discretization methods suffer from the drawback that in almost every step all local maxima of $g(z^i, \cdot)$ over B have to be computed. This situation is better for reduction methods, where global search usually can be restricted to every third or fourth step. Reduction methods also have advantages. The convergence (to high accuracy) is fast and the number of constraints in the subproblems (also for $m \geq 2$) usually can be kept low. However, if Assumption 7.7 does not hold, the method may fail.

In developing a hybrid approach one could anticipate using a reduction method as the basic method which is complemented by exchange steps as a starting phase, and with discretization steps to be employed if difficulties in the reduction steps are encountered.

8. Parametric problems. Results from finite parametric programming have been used in §4 to reduce SIP locally to a finite optimization problem ($P_{\text{red}}(\bar{z})$). This approach has proved to be essential in previous sections on various theoretical and numerical aspects. In addition, in Example 2.1 (see also [8.17]), we have seen that there are applications leading to SIP which themselves depend on parameters. It would go beyond the scope of this review to treat the huge field of parametric optimization to its full extent. We restrict ourselves to those aspects which are especially important for SIP.

In §4, to derive the reduced problem $(P_{\mathrm{red}}(\bar{z}))$ in a given point $\bar{z}$, we considered Problem 4.6:

$$(8.1) \qquad (O(z)) \qquad \max\{g(z,t)|h^j(t) \le 0,\ j \in J\},$$

a finite parametric problem with g, h^j twice continuously differentiable. With respect to our goal, we are exclusively interested in those results in the theory of parametric programming that lead to functions $t^l(z)$, which give locally unique solutions depending on z. The existence of such $t^l(z)$ usually is justified by applying the Implicit Function Theorem to the Karush–Kuhn–Tucker systems; see [8.9]:

$$(8.2) \qquad \begin{aligned} \mathcal{L}_t^l(z,t,\alpha^l) = g_t(z,t) - \sum_{j\in M^l} \alpha_j^l h^j(t) = 0, \\ h^j(t) = 0, \quad j \in M^l, \end{aligned}$$

where M_l is the set of indices identifying active constraints for the solution $\bar{t}^l$ of $(O(\bar{z}))$ and $\mathcal{L}^l$ the Lagrangian (see (4.8), (4.9)). Therefore, some stability concept for the KKT point $(\bar{t}^l, \bar{\alpha}^l)$ is required, ensuring the existence of a KKT manifold $(t^l(z), \alpha^l(z))$ (cf. [8.19]). A sufficient condition, for instance, is that $(\bar{t}^l, \bar{\alpha}^l)$ is nondegenerate (implied by Assumption 4.14), or strongly regular in the sense of Robinson [8.24] (implied by Assumption 4.13), or strongly stable in the sense of Kojima [8.21]. In these cases the functions $(t^l(z), \alpha^l(z))$ exist locally and are continuously differentiable, Lipschitz continuous, and continuous, respectively.

The above observation remains true for problems of the type

$$(8.3) \qquad (\tilde{O}(z)) \qquad \max\{g(z,t)|h^j(z,t) \le 0,\ j \in J\},$$

where the feasible set also depends on z; therefore, reduction can be performed analogously in generalized SIP (cf. [2.21], [5.14]), where B may depend on z.

Numerical methods for one-parameter problems usually track KKT curves by applying path following methods to system (8.2) (see, e.g., [8.11], [8.14]). A good introduction into the above may be found in [8.13], where an extensive list of references is also given.

Let us turn now to a SIP problem depending on a real parameter vector $\lambda \in \mathbb{R}^p$,

$$(\mathrm{P}(\lambda)) \qquad \begin{aligned} v(\lambda) = \max\{F(\lambda,z)|z \in Z(\lambda)\} \quad \text{with} \\ Z(\lambda) = \{z|g(\lambda,z,t) \le 0, t \in B\}. \end{aligned}$$

In Example 2.2 we have presented an application to membrane eigenvalue problems, where the eigenvalues are approximated by local minima of the value function $v(\lambda)$ of a linear problem $(P_l(\lambda))$, $\lambda \in \mathbb{R}$,

$$(\mathrm{P}_l(\lambda)) \qquad \begin{aligned} v(\lambda) = \max\{c^T z|z \in Z_l(\lambda)\} \quad \text{with} \\ Z_l(\lambda) = \{z|a^T(\lambda,t)z \le b(\lambda,t), t \in B\}. \end{aligned}$$

General rational Chebyshev approximation [2.69], [2.37] provides another application where a zero of $v(\lambda)$ is required.

An application amounting to a nonlinear $(P(\lambda))$ may be found in [2.34], where the eigenvalues of elliptic membranes are considered as functions of the eccentricity of the ellipse.

To determine local minima or zeros of $v(\lambda)$, $\lambda \in \mathbb{R}$, it is valuable to be able to compute (one-sided) derivatives of $v(\lambda)$ if these exist. In [8.27], for $(P_l(\lambda))$ it has been proved that these one-sided derivatives exist at $\lambda = \bar{\lambda}$ if $(P_l(\bar{\lambda}))$ and its dual are both superconsistent (cf. Definitions 6.6 and 6.7). Moreover, an alternative representation is given there, which in applications often is more useful than the usual inf-sup representations from convex programming (cf. [8.18]). In [2.37] this has been used to prove superlinear convergence of an algorithm for general constrained rational approximation under very weak assumptions.

In [8.25], [8.26] path following methods for nonlinear problems $(P(\lambda))$, $\lambda \in \mathbb{R}$, are investigated. To this end, Rupp generalized the results in [8.19] on Kuhn–Tucker curves for finite parametric problems. He showed [8.26] that under generic assumptions the set of nondegenerate Kuhn–Tucker points of $(P_{\text{red}}(\lambda, z))$, traced in the space augmented with variable $\lambda(\lambda \in \mathbb{R})$, consists of disjoint, one-dimensional differentiable manifolds, some of them connected in isolated degenerate Kuhn–Tucker points. These isolated, exceptional points can be characterized in a way so that their numerical identification is possible during computation. This forms the theoretical basis for a path following algorithm (cf. also [8.25]) which proved stable and efficient in a number of difficult test problems (see [8.26]).

For computational purposes in one way or another the existence of sufficiently smooth Kuhn–Tucker curves must be assumed. It follows from the results of Rupp that generically this may be assumed.

Beyond that, much more general investigations on the dependence of the feasible set or the solution set on more general parameters required can be found in the literature. Often the functions F and g themselves are considered as parameters. See [8.3] for an early extensive study. One question concerns weak assumptions that assure (upper or lower semi-) continuity of the related set valued mappings associating the feasible set or solution set to the parameter [8.1], [8.4], [8.6], [8.7], [8.8], [8.15], [8.16]. Other questions concern the existence of continuous selections from these mappings [8.10] and stability results for (locally unique) solutions [8.2], [8.5], [8.23] and feasible sets. With respect to the latter question, in [8.20] it has been shown that the feasible set transforms homeomorphically in F and g (in an appropriate topology) if and only if the Mangasarian–Fromovitz Constraint Qualification holds.

9. Conclusion. As either editor or coeditor of a Springer-Verlag Lecture Notes Series, [1.4], [1.7], each of us separately has written a preface about an attitude towards semi-infinite programming. In simplest terms, the common ground in those descriptions also pervades this review, namely, the need for a theory which addresses optimization problems having no alternate formulations other than ones with a finite number of variables appearing in infinitely many constraints. Nevertheless, the successful approaches that we have described in this paper, must, as one should expect, make fruitful connections with finite problems in order to be computationally realized. We have gone into considerable depth about how a nonlinear SIP can be connected to a finite nonlinear programming problem, which, in general, can only be done in a local sense. With regard to this question recent structural results on parametric optimization, obtained by Jongen and others, have led to a considerable progress during the last years. It was possible to specify the differentiability properties of these implicitly defined finite problems and to classify exceptional irregularities which (generically) have to be taken into account. On this basis it was possible to create effective and robust numerical approaches, which have received broad attention in our review.

In the linear case, rather early in the 60s when linear SIP was viewed as a next level of extension of ordinary finite linear programming, the relation to finite problems was studied intensively. Any "discretization" would yield a linear program, but questions focused on whether one discretization really could suffice, or alternately would one necessarily be forced to consider exchange methods or discretization methods. Addressing these questions quickly leads into the phenomena of duality gaps, and one is forced to investigate various regularity conditions for avoiding them. In this paper we have reviewed the particular infinite-finite connections endemic to semi-infinite programming, already in the linear case. Perhaps one of the prototype roles of the linear case per se is to demonstrate perfect duality in a more readily available context. In addition to doing this we have reviewed ways for achieving perfect duality for cases of nonperfect duality also. For future consideration it could be interesting in this context to derive numerical procedures for identifying which of the 11 permissible duality states applies to a given linear SIP, out of a possible 121 duality states (cf. §6.4 and the references given there).

Interesting questions for future numerical research could also be addressed by developing methods for problems where B has dimension higher than two or three and for problems where B again depends on z. Moreover, the development of hybrid algorithms still is preliminary, and the same holds for numerical approaches to parametric problems. Fields of applications in probability theory and in optimal control problems with nontrivial state and control constraints have not yet been sufficiently exploited.

Acknowledgment. We thank all of those who have contributed to this review by giving valuable suggestions on the text and the references. We are indebted especially to C. E. Blair, B. Jongen, E. Polak, R. Tichatschke, and an unknown referee.

REFERENCES

[1] *Introduction*

[1.1] E. J. ANDERSON AND P. NASH, *Linear Programming in Infinite Dimensional Spaces*, John Wiley & Sons, New York, 1987.

[1.2] E. J. ANDERSON AND A. B. PHILPOTT, EDS., *Infinite programming*, Lecture Notes in Econom. and Math. Systems 259, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1985.

[1.3] E. BLUM AND W. OETTLI, *Mathematische Optimierung*, Springer-Verlag, Berlin, Heidelberg, New York, 1975.

[1.4] A. V. FIACCO AND K. O. KORTANEK, EDS., *Semi-infinite programming and applications*, Lecture Notes in Econom. and Math. Systems 215, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1983.

[1.5] K. GLASHOFF AND S.-Å. GUSTAFSON, *Linear Optimization and Approximation*, Springer-Verlag, Berlin, Heidelberg, New York, 1983.

[1.6] S.-Å. GUSTAFSON AND K. O. KORTANEK, *Semi-infinite programming and applications*, in Mathematical Programming: the State of the Art, A. Bachem, M. Grötschel, and B. Korte, eds., Springer-Verlag, Berlin, New York, 1983, pp. 132–157.

[1.7] R. HETTICH, ED., *Semi-infinite programming*, Lecture Notes in Control and Inform. Sci. 15, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1979.

[1.8] R. HETTICH AND P. ZENCKE, *Numerische Methoden der Approximation und semi-infiniten Optimierung*, Teubner Studienbücher Mathematik, Stuttgart, 1982.

[1.9] W. KRABS, *Optimization and Approximation*, John Wiley & Sons, Chichester, New York, Brisbane, Toronto, 1979.

[1.10] R. T. ROCKAFELLAR *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.

[1.11] J. STOER AND C. H. WITZGALL, *Convexity and Optimization in Finite Dimension* I, Springer-Verlag, Berlin, Heidelberg, New York, 1970.

[2] *Applications*

[2.1] B. BROSOWSKI, *Zur Stabilität von Näherungsverfahren für die Potentialgleichung*, Z. Angew. Math. Mech., 62 (1982), pp. T322–323.

[2.2] E. CHRISTIANSEN, *Limit analysis in plasticity as a mathematical programming problem*, Calcolo, 17 (1980), pp. 41–65.

[2.3] ———, *Computation of limit loads*, Internat. J. Numer. Methods Engrg., 17 (1981), pp. 1547–1570.

[2.4] ———, *Limit analysis for plastic plates*, SIAM J. Math. Anal., 11 (1980), pp. 514–522.

[2.5] ———, *Examples of collapse solutions in limit analysis*, Utilitas Math., 22 (1982), pp. 77–102.

[2.6] E. CHRISTIANSEN AND K. O. KORTANEK, *Computing material collapse displacement fields on a CRAY X-MP/48 by the LP primal affine scaling algorithm*, Ann. Oper. Res., 22 (1990), pp. 355–376.

[2.7] E. Christiansen and K. O. Kortanek, *Computation of the collapse state in limit analysis using the LP primal affine scaling algorithm*, J. Comput. Appl. Math., 34 (1991), pp. 47–63.

[2.8] L. Collatz, *Functional Analysis and Numerical Mathematics*, Academic Press, New York, London, 1966.

[2.9] L. Collatz and W. Krabs, *Approximationstheorie*, B. G. Teubner, Stuttgart, 1973.

[2.10] D. Colton, *The inverse scattering problem for time-harmonic acoustic waves*, SIAM Rev., 26 (1984), pp. 323–350.

[2.11] D. Colton and R. Reemtsen, *The numerical solution of the inverse Stefan problem in two space variables*, SIAM J. Appl. Math, 44 (1984), pp. 996–1013.

[2.12] G. Dahlquist, S. Eisenstat, and G. H. Golub, *Bounds for the error of linear systems of equations using the theory of moments*, J. Math. Anal. Appl., 37 (1972), pp. 151–166.

[2.13] G. Dahlquist, G. H. Golub, and S. G. Nash, *Bounds for the error in linear systems*, in Semi-Infinite Programming, Lecture Notes in Control and Inform. Sci. 15, R. Hettich, ed., Springer-Verlag, Berlin, Heidelberg, New York, 1979, pp. 154–172.

[2.14] A. Ech-Cherif, J. G. Ecker, M. Kupferschmid, and S. P. Marin, *Robot trajectory planning by nonlinear programming*, General Motors Research Publication, GMR - 6460, December, 1988.

[2.15] J. P. Evans, *Duality in mathematical programming over Markov chains and countable action spaces in finite horizon problems*, Management Sci., 15 (1969), pp. 626–638.

[2.16] V. V. Fedorov, *Theory of Optimal Experiments*, translated and edited by W. J. Studden and E. M. Klimko, Academic Press, New York, 1972.

[2.17] K. Glashoff and S.-Å. Gustafson, *Numerical treatment of a parabolic boundary value control problem*, J. Optim. Theory Appl., 19 (1976), pp. 645–663.

[2.18] K. Glashoff and K. Roleff, *A new method for Chebyshev approximation of complex-valued functions*, Math. Comp., 36 (1981), pp. 233–239.

[2.19] W. L. Gorr, S.-Å. Gustafson, and K. O. Kortanek, *Optimal control strategies for air quality standards and regulatory policies*, Environment and Planning, 4 (1972), pp. 183–192.

[2.20] W. L. Gorr and K. O. Kortanek, *Numerical aspects of pollution abatement problems: optimal control strategies for air quality standards*, in Proc. Oper. Res., M. Henke, A. Jaeger, R. Wartman, H. J. Zimmermann, eds., Physica-Verlag, Würzburg, Vienna, 1971, pp. 34–58.

[2.21] T. J. Graettinger and B. H. Krough, *The acceleration radius: a global performance measure for robotic manipulators*, IEEE J. of Robotics and Automation, 4 (1988), pp. 60–69.

[2.22] P. R. Gribik, *Selected applications of semi-infinite programming*, in Constructive Approaches to Mathematical Models, C. V. Coffman and G. J. Fix, eds., Academic Press, New York, London, 1979, pp. 171–187.

[2.23] P. R. Gribik and K. O. Kortanek, *Equivalence theorems and cutting plane algorithms for a class of experimental design problems*, SIAM J. Appl. Math., 32 (1977), pp. 232–259.

[2.24] S.-Å. Gustafson, *On the computational solution of a class of generalized moment problems*, SIAM J. Numer. Anal., 7 (1970), pp. 343–357.

[2.25] ———, *On computational applications of the theory of the moment problem*, Rocky Mountain J. Math., 4 (1974), pp. 227–240.

[2.26] ———, *On the numerical treatment of a multi-dimensional parabolic boundary-value control problem*, in Optimization and Optimal Control, R. Bulirsch, W. Oettli and J. Stoer, eds., Lecture Notes in Math. 477, Springer-Verlag, Berlin, Heidelberg, New York, 1975, pp. 121–126.

[2.27] ———, *On semi-infinite programming in numerical analysis*, in Semi-Infinite Programming, Lecture Notes in Control and Inform. Sci. 15, R. Hettich, ed., Springer-Verlag, Berlin, Heidelberg, New York, 1979, pp. 137–153.

[2.28] S.-Å. Gustafson and K. O. Kortanek, *Determining sampling equipment locations by optimal experimental design with applications to environmental protection and acoustics*, in Proc. Comp. Sci. Statist. 7th Symp. on the Interface, W. F. Kennedy, ed., Iowa State University, Ames, IA, 1973, pp. 332–338.

[2.29] S.-Å. Gustafson, K. O. Kortanek, and W. Rom, *Non-Chebyshevian moment problems*, SIAM J. Numer. Anal., 7 (1970), pp. 335–342.

[2.30] E. Haaren-Retagne, *A semi-infinite programming algorithm for robot trajectory planning*, Ph.D. thesis, Universität Trier, Germany, in preparation.

[2.31] E. Haaren-Retagne, R. Hettich, and G. Still, *A numerical comparison of two approaches to compute membrane eigenvalues by defect minimization*, in Numerical Treatment of Eigenvalue Problems, J. Albrecht et al., eds., Vol. 5, Internat. Ser. Numer. Math. 96, Birkhäuser, Basel, Boston, Berlin, 1991, pp. 209–224.

[2.32] E. Haaren-Retagne, *Robot trajectory planning with semi-infinite programming*, IEEE Trans. Autom. Control, submitted.

[2.33] R. Hettich, *On the computation of membrane-eigenvalues by semi-infinite programming methods*, in Infinite Programming, E. J. Anderson and A. B. Philpott, eds., Lecture Notes in Econom. and Math. Systems 259, Springer-Verlag, Berlin, Heidelberg, New York, 1985, pp. 79–89.

[2.34] R. Hettich, E. Haaren, M. Ries, and G. Still, *Accurate numerical approximations of eigenfrequencies and eigenfunctions of elliptic membranes*, Z. Angew. Math. Mech., 67 (1987), pp. 589–597.

[2.35] R. Hettich and G. Still, *Computation of eigenvalues and eigenfunctions of symmetric operators by defect-minimization*, in Numerical Treatment of Eigenvalue Problems, J. Albrecht, L. Collatz, W. Velte, W. Wunderlich, eds., Vol. 4, Internat. Ser. Numer. Math. 83, Birkhäuser-Verlag, Basel, Boston, 1987, pp. 124–129.

[2.36] ———, *Semi-infinite programming models in robotics*, in Parametric Optimization and Related Topics II, J. Guddat et al., eds., Math. Res. 62, Akademie Verlag, Berlin, 1991, pp. 112–118.

[2.37] R. Hettich and P. Zencke, *An algorithm for general rational Chebyshev approximation*, SIAM J. Numer. Anal., 27 (1990), pp. 1024–1033.

[2.38] E. W. Hobson, *The Theory of Functions of a Real Variable and the Theory of Fourier Series*, Vol. 1, Dover Publications, New York, 1927 and 1957.

[2.39] P. Jochum, *The inverse Stefan problem as a problem of nonlinear approximation theory*, J. Approx. Theory, 30 (1980), pp. 81–98.

[2.40] F. Juhnke, *Inradius and thickness of convex bodies under optimization-theoretical viewpoints*, Beiträge Algebra Geom., 27 (1988), pp. 13–20.

[2.41] ———, *The circumsphere problem and linear semi-infinite optimization*, Beiträge Algebra Geom., 28 (1989), pp. 147–156.

[2.42] S. Karlin and W. J. Studden, *Tchebycheff Systems: with Applications in Analysis and Statistics*, Interscience Publishers, New York, London, Sydney, 1966.

[2.43] J. H. B. Kemperman, *On the role of duality in the theory of moments*, in Semi-Infinite Programming and Applications, A. V. Fiacco and K. O. Kortanek, eds., Lecture Notes In Econom. and Math. Systems 215, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1983, pp. 63–92.

[2.44] K. C. Kiwiel, *A linearization algorithm for optimizing control systems subject to singular value inequalities*, IEEE Trans. Autom. Control AC-31, 1986, pp. 595–602.

[2.45] P. Knabner, *Control of Stefan problems by means of linear-quadratic defect minimization*, Numer. Math., 46 (1985), pp. 429–442.

[2.46] K. O. Kortanek, *Using generalized finite sequences in mathematical physics*, Department of Mathematics, Carnegie Mellon University, Pittsburgh, PA, August 1984.

[2.47] ———, *Semi-infinite programming duality and finite elements in plane stress plasticity*, Utilitas Math., 28 (1985), pp. 219–242.

[2.48] ———, *Semi-infinite programming and continuum physics*, in Infinite Programming, E. J. Anderson and A. B. Philpott, eds., Lecture Notes Econom. and Math. Systems 259, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1985, pp. 65–78.

[2.49] ———, *Vector-Supercomputer experiments with the primal affine linear programming scaling algorithm*, SIAM J. Sci. Comput., 14 (1993), pp. 279–294.

[2.50] K. O. Kortanek and M. Yamasaki, *Semi-infinite transportation problems*, J. Math. Anal. Appl., 88 (1982), pp. 555–565.

[2.51] N. Levinson, *A class of continuous linear programming problems*, J. Math. Anal. Appl., 16 (1966), pp. 73–83.

[2.52] S. P. Marin, *Optimal parametrization of curves for robot trajectory design*, IEEE Trans. on Automatic Control, Vol. AC-33, 1988, pp. 209–214.

[2.53] H. Matthies, G. Strang, and E. Christiansen, *The saddle point of a differential program*, in Energy Methods in Finite Element Analysis, R. Glowinski, E. Rodin and O. C. Zienkiewicz, eds., John Wiley & Sons, New York, 1979, pp. 309–319.

[2.54] G. Opfer, *Computation of T-approximations by complex planar splines*, Z. Angew. Math. Mech., 62 (1982), pp. T342–344.

[2.55] ———, *Solving complex approximation problems by semi-infinite optimization techniques: a study on convergence*, Numer. Math., 39 (1982), pp. 411–420.

[2.56] M. L. Parks and A. L. Soyster, *Semi-infinite programming and fuzzy set programming*, in Semi-Infinite Programming and Applications, A. V. Fiacco and K. O. Kortanek, eds., Lecture Notes in Econom. and Math. Systems 215, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1983, pp. 219–235.

[2.57] A. Pazman, *Foundations of Optimum Experimental Design*, D. Reidel, Boston, MA, 1986.

[2.58] E. Polak, *Semi-infinite programming in engineering design*, in Semi-Infinite Programming and Applications, A. V. Fiacco and K. O. Kortanek, eds., Lecture Notes in Econom. and Math. Systems 215, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1983, pp. 236–248.

[2.59] E. Polak, S. E. Salcudean, and D. Q. Mayne, *Adaptive control of* ARMA *plants using worst-case design by semi-infinite optimization*, IEEE Trans. Autom. Control AC-32, 1987, pp. 388–396.

[2.60] E. Polak, D. Q. Mayne, and D. M. Stimler, *Control system design via semi-infinite optimization*, IEEE Proc., 1984, p. 1777.

[2.61] E. Polak, K. S. Pister, and D. Ray, *Optimal design of framed structures subject to earthquakes*, Engrg. Opt., 12 (1976), pp. 65–71.

[2.62] E. Polak and T. L. Wuu, *On the design of stabilizing compensators via semi-infinite optimization*, IEEE Trans. Autom. Control AC-34, 1989, pp. 196–200.

[2.63] F. Pukelsheim, *On linear regression designs which maximize information*, J. Statist. Plan. Inference, 4 (1980), pp. 339–364.

[2.64] R. Reemtsen, *Defect minimization in operator equations: theory and application*, Pitman Res. Notes in Math. Ser., 163, Longman, Harlow and New York, 1987.

[2.65] R. Reemtsen and A. Kirsch, *A method for the numerical solution of the one-dimensional inverse Stefan problem*, Numer. Math., 45 (1984), pp. 253–273.

[2.66] R. Reemtsen and C. Lozano, *An approximation technique for the numerical solution of a Stefan problem*, Numer. Math., 38 (1981), pp. 141–154.

[2.67] T. Robertson, F. T. Wright, and R. L. Dykstra, *Order Restricted Statistical Inference*, John Wiley and Sons, Chichester, New York, Brisbane, Toronto, Singapore, 1988.

[2.68] A. Sard, *Linear Approximation*, Math. Survey No. 9, American Mathematical Society, Providence, RI, 1963.

[2.69] R. Sommer, *Behandlung semi-infiniter Optimierungsaufgaben bei der approximativen Lösung elliptischer Randwertaufgaben*, Wiss. Schr. Tech. Hochsch. Karl-Marx-Stadt, 23 (1981), pp. 471–478.

[2.70] A. L. Soyster, *A semi-infinite game*, Management Sci., 21 (1975), pp. 806–812.

[2.71] ———, *Inexact linear programming with generalized resource sets*, European J. Oper. Res., 3 (1979), pp. 316–321.

[2.72] G. Still, *Computable bounds for eigenvalues and eigenfunctions of elliptic differential operators*, Numer. Math., 54 (1988), pp. 201–223.

[2.73] ———, *Defektminimierungsmethoden zur Lösung elliptischer Rand—und Eigenwertaufgaben—Spezielle Klassen von Ansatzfunktionen und ihre Approximationseigenschaften*, Habilitationsschrift, Universität Trier, 1989.

[2.74] G. Strang, *A minimax problem in plasticity theory*, in Functional Analysis Methods in Numerical Analysis, M. Z. Nashed, ed., Lecture Notes in Math. 701, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1979, pp. 319–333.

[2.75] ———, *Maximal flow through a domain*, Math. Programming, 26 (1983), pp. 123–143.

[2.76] H. Strauss, *On semi-infinite optimization and Chebyshev approximation with constraints*, Numer. Funct. Anal. Optim., 8 (1985), pp. 153–171.

[2.77] R. L. Streit, *Solutions of systems of complex linear equations in the* l_∞ norm *with constraints on the unknowns*, SIAM J. Sci. Statist. Comput., 7 (1986), pp. 132–149.

[2.78] R. L. Streit and A. H. Nutall, *A note on the semi-infinite programming approach to complex approximation*, Math. Comp., 40 (1983), pp. 599–605.

[2.79] N. Sturm, *Semi-infinite Optimierung: Lösungsvorschlag mittels der Momentenmethode von Markov*, Schwarzenbek, 1978.

[2.80] R. Tichatschke, *Semi-infinite Optimierung: Eine Übersicht über Theorie, Lösungsverfahren und Anwendungen*, Mit. Math. Ges. DDR, 4 (1986), pp. 20–34.

[2.81] R. Tichatschke, R. Hettich and G. Still, *Connections between generalized, inexact, and semi-infinite linear programming*, Z. Oper. Res., 33 (1989), pp. 367–382.

[2.82] S. H. Tijs, *Semi-infinite linear programs and semi-infinite matrix games*, Nieuw Arch. Wisk., 27 (1979), pp. 197–214.

[2.83] A. L. Tits, *On the optimal design centering, tolerancing and tuning problem*, J. Optim. Theory Appl., 45 (1985), pp. 487–499.

[2.84] D. M. Titterington, *Geometric approaches to design of experiments*, Math. Oper. Statist. Ser. Statist., 11 (1980), pp. 151–163.

[2.85] B. Tornsey, *Moment inequalities via optimal design theory*, Linear Algebra Appl., 82 (1986), pp. 237–253.

[2.86] G. T. Venkov and N. T. Krushev, *Nondifferentiable design optimization problems with application to control engineering*, Technical University of Sofia, Bulgaria, preprint, 1990.

[2.87] W. Wetterling, *Quotienteneinschließung beim ersten Membraneigenwert, Numerische Behandlung von Differentialgleichungen*, Bd. 3, Tag. Clausthal 1980, Internat. Ser. Numer. Math., 56 (1981), pp. 193–198.

[2.88] M. YAMASAKI, *Semi-infinite programs and conditional Gauss variational problems*, Hiroshima Math. J., 1 (1971), pp. 177–226 (and corrections in 2 (1972), p. 547).

[2.89] D. ZWICK, *A special one-sided approximation problem*, J. Approx. Theory, 62 (1990), pp. 70–93.

[3] *First-order optimality conditions*

[3.1] E. W. CHENEY, *Introduction to Approximation Theory*, McGraw-Hill, New York, 1966.

[3.2] L. CROMME, *Eine Klasse von Verfahren zur Ermittlung bester nichtlinearer Tchebyscheff–Approximationen*, Numer. Math., 25 (1975), pp. 447–459.

[3.3] TH. FISCHER, *Strong unicity in normed linear spaces*, Numer. Funct. Anal. Optim., to appear.

[3.4] G. NUERNBERGER, *Global unicity in semi-infinite optimization*, Numer. Funct. Anal. Optim., 8 (1985), pp. 173–191.

[3.5] ———, *Global unicity in optimization and approximation*, Z. Angew. Math. Mech., 65 (1985), pp. T319–T321.

[4] *Discretization and (local) reduction to finite problems*

[4.1] J. GAUVIN AND J. W. TOLLE, *Differential stability in non-linear programming*, SIAM J. Control Optim., 15 (1977), pp. 294–311.

[4.2] R. D. GRIGORIEFF AND R. M. REEMTSEN, *Discrete approximations of minimization problems*, I. theory, II. applications, preprints 260 and 263, 1990, Technical University, Berlin, Germany.

[4.3] S.-Å. GUSTAFSON, *On numerical analysis in semi-infinite programming*, in Semi-Infinite Programming, R. Hettich, ed., Lecture Notes Control and Inform. Sci., 15, Springer-Verlag, Berlin, Heidelberg, New York, 1979, pp. 51–65.

[4.4] K. JITTORNTRUM, *Solution point differentiability without strict complementarity in nonlinear programming*, in Sensitivity, Stability, and Parametric Analysis, A. V. Fiacco, ed., Math. Progr. Study, 21 (1984), pp. 127–138.

[4.5] R. REEMTSEN, *Discretization methods for the solution of semi-infinite programming problems*, J. Optim. Theory Appl., 71 (1991), pp. 85–104.

[4.6] G. ZWIER, *Structural analysis in semi-infinite programming*, Ph.D. thesis, University of Twente, Enschede, The Netherlands, 1987.

[5] *Second-order optimality conditions*

[5.1] A. BEN-TAL, *Second order theory of extremum problems*, in Extremal Methods and Systems Analysis, A. V. Fiacco and K. O. Kortanek, eds., Lecture Notes in Econom. and Math. Systems 174, Springer-Verlag, Berlin, Heidelberg, New York, 1980, pp. 336–356.

[5.2] A. BEN-TAL, M. TEBOULLE, AND J. ZOWE, *Second order necessary optimality conditions for semi-infinite programming problems*, in Semi-Infinite Programming, R. Hettich, ed., Springer-Verlag, Berlin, Heidelberg, New York, 1979, pp. 17–30.

[5.3] R. HETTICH, *Extremalkriterien für Optimierungs- und Approximations-aufgaben*, Ph.D. thesis, TH Twente, The Netherlands, 1973.

[5.4] ———, *Kriterien zweiter Ordnung für lokal beste Approximation*, Numer. Math., 22 (1974), pp. 409–417.

[5.5] R. HETTICH AND H. TH. JONGEN, *On first and second order conditions for local optima for optimization problems in finite dimensions*, Methods Oper. Res., 23 (1977), pp. 82–97.

[5.6] ———, *Semi-infinite programming: conditions of optimality and applications*, in Optimization Techniques, J. Stoer, ed., Part 2, Lecture Notes in Control and Inform. Sci., 7, Springer-Verlag, Berlin, Heidelberg, New York, 1978, pp. 1–11.

[5.7] R. HETTICH AND G. STILL, *Second order conditions for semi-infinite optimization*, in preparation.

[5.8] A. D. JOFFE, *Second-order conditions in nonlinear nonsmooth problems of semi-infinite programming*, in Semi-Infinite Programming and Applications, A. V. Fiacco and K. O. Kortanek, eds., Lecture Notes in Econom. and Math. Systems 215, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1983, pp. 262–280.

[5.9] H. TH. JONGEN, W. WETTERLING AND G. ZWIER, *On sufficient conditions for local optimality in semi-infinite programming*, Math. Operationsforsch. Statist., Ser. Optim., 18 (1987), pp. 165–178.

[5.10] H. KAWASAKI, *An envelope-like effect of infinitely many inequality constraints on second-order necessary conditions for minimization problems*, Math. Programming, 41 (1988), pp. 73–96.

[5.11] ———, *Second order necessary optimality conditions for minimizing a sup-type function*, Math. Programming, 49 (1991), pp. 213–229.

[5.12] D. KLATTE, B. KUMMER, AND R. WALZEBOK, *Conditions for optimality and strong stability in nonlinear programs without assuming twice differentiability of data*, preprint.

[5.13] D. KLATTE, AND K. TAMMER, *On second-order sufficient optimality conditions for $C^{1,1}$-optimization problems*, Optimization, 19 (1988), pp. 169–179.

[5.14] A. SHAPIRO, *Second-order derivatives of extremal-value functions and optimality conditions for semi-infinite programs*, Math. Oper. Res., 10 (1985), pp. 207–219.

[5.15] W. WETTERLING, *Definitheitsbedingungen für relative Extrema bei Optimierungs- und Approximationsaufgaben*, Numer. Math., 15 (1970), pp. 122–136.

[6] *Duality theory in linear semi-infinite programming*

[6.1] A. BEN-TAL, A. BEN-ISRAEL, AND E. ROSINGER, *A Helly-type theorem and semi-infinite programming*, in Constructive Approaches to Mathematical Models, C. V. Coffman and G. J. Fix, eds., Academic Press, New York, 1979, pp. 127–135.

[6.2] C. E. BLAIR, *A note on infinite systems of linear inequalities in R^n*, J. Math. Anal. Appl., 48 (1974), pp. 150–154.

[6.3] ———, *Ascent ray theorems and some applications*, in Semi-Infinite Programming and Applications, A. V. Fiacco and K. O. Kortanek, eds., Lecture Notes in Econom. and Math. Systems 215, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1983, pp. 2–9.

[6.4] J. M. BORWEIN, *A note on perfect duality and limiting Lagrangians*, Math. Programming, 18 (1980), pp. 330–337.

[6.5] ———, *Lexicographic multipliers*, J. Math. Anal. Appl., 78 (1980), pp. 309–327.

[6.6] ———, *Direct theorems in semi-infinite convex programming*, Math. Programming, 21 (1981), pp. 301–318.

[6.7] ———, *Semi-infinite programming: how special is it?*, in Semi-Infinite Programming and Applications, A. V. Fiacco and K. O. Kortanek, eds., Lecture Notes in Econom. and Math. Systems 215, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1983, pp. 10–36.

[6.8] J. M. BORWEIN AND K. O. KORTANEK, *Fenchel duality and separably-infinite programs*, Optimization, Math. OP. & Stat., 14 (1983), pp. 1–12.

[6.9] A. CHARNES AND W. W. COOPER, *The strong Minkowski–Farkas–Weyl theorem for vector spaces over ordered fields*, Proc. Nat. Acad. Sci. USA, 44 (1958), pp. 914–916.

[6.10] A. CHARNES, W. W. COOPER, AND K. O. KORTANEK, *Duality, Haar programs and finite sequence spaces*, Proc. Nat. Acad. Sci. USA, 48 (1962), pp. 783–786.

[6.11] ———, *A duality theory for convex programs with convex constraints*, Bull. Amer. Math. Soc., 68 (1962), pp. 605–608.

[6.12] ———, *Duality in semi-infinite programs and some works of Haar and Caratheodory*, Management Sci., 9 (1963), pp. 209–228.

[6.13] ———, *On representations of semi-infinite programs which have no duality gaps*, Management Sci., 12 (1965), pp. 113–121.

[6.14] ———, *On the theory of semi-infinite programming and some generalizations of Kuhn–Tucker saddle point theorems for arbitrary convex functions*, Naval Res. Logist., 16 (1969), pp. 41–51.

[6.15] A. CHARNES, P. R. GRIBIK, AND K. O. KORTANEK, *Separably-infinite programming*, Z. Oper. Res., 24 (1980), pp. 33–45.

[6.16] ———, *Polyextremal principles and separably–infinite programs*, Z. Oper. Res., 25 (1980), pp. 211–234.

[6.17] A. CHARNES AND K. O. KORTANEK, *An opposite sign algorithm for purification to an extreme point solution*, ONR Memo. No. 84, Northwestern University, Evanston, IL 1963.

[6.18] U. DIETER, *Optimierungsaufgaben in topologischen Vektorräumen* I: *Dualitätstheorie*, Z. Wahrscheinlichkeitstheorie verw. Geb, 5 (1966), pp. 89–117.

[6.19] R. J. DUFFIN, *Infinite programs*, in Linear Inequalities and Related Systems, H. W. Kuhn and A. Tucker, eds., Princeton University Press, Princeton, NJ, 1956, pp. 157–170.

[6.20] R. J. DUFFIN AND L. A. KARLOVITZ, *An infinite linear program with a duality gap*, Management Sci., 12 (1965), pp. 122–134.

[6.21] R. J. DUFFIN, *Convex analysis treated by linear programming*, Math. Programming, 4 (1973), pp. 125–143.

[6.22] R. J. DUFFIN, R. G. JEROSLOW, AND L. A. KARLOVITZ, *Duality in semi-infinite linear programming*, in Semi-Infinite Programming and Applications, A. V. Fiacco and K. O. Kortanek, Eds., Lecture Notes in Econom. and Math. Systems 215, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1983, pp. 50–62.

[6.23] K. FAN, *Asymptotic cones and duality of linear relations*, J. Approx. Theory, 2 (1969), pp. 152–159.

[6.24] K. GLASHOFF, *Duality theory of semi-infinite programming*, in Semi-Infinite Programming, R. Hettich, ed., Lecture Notes in Control and Inform. Sci. 15, Springer-Verlag, Berlin, Heidelberg, New York, 1979, pp. 1–16.

[6.25] M. A. GOBERNA, M. A. LOPEZ, AND J. PASTOR, *Farkas–Minkowski systems in semi-infinite programming*, Appl. Math. Optim, 7 (1981), pp. 295–308.

[6.26] M. A. GOBERNA AND M. A. LOPEZ, *Conditions for the closedness of the characteristic cone associated to an infinite linear system*, in Infinite Programming, E. J. Anderson and A. Philpott, eds., Lecture Notes in Control and Inform. Sci., Springer-Verlag, Berlin, Heidelberg, New York, 1984, pp. 16–28.

[6.27] ———, *A theory of linear inequality systems*, Linear Algebra Appl., 106 (1988), pp. 77–115.

[6.28] W. GOCHET, Y. SMEERS, AND K. O. KORTANEK, *Using semi-infinite programming in geometric programming*, in Proc. 20th Inter. TIMS: On Management Sciences, Developing Countries, and National Priorities, E. Shlifer, ed., Jerusalem Acad. Press and North Holland, Amsterdam, 1973, pp. 430–438.
[6.29] ———, *On a classification scheme for geometric programming and complementarity theorems*, Appl. Anal., 6 (1976), pp. 47–59.
[6.30] E. G. GOL'STEIN, *Theory of Convex Programming*, Trans. Math. Mono. Amer. Math. Soc., Providence, RI, 1972.
[6.31] R. G. JEROSLOW, *A limiting Lagrangian for infinitely constrained convex optimization in R^n*, J. Optim. Theory Appl., 33 (1981), pp. 479–495.
[6.32] ———, *Uniform duality in semi-infinite convex optimization*, Math. Programming, 27 (1983), pp. 144–154.
[6.33] R. G. JEROSLOW AND K. O. KORTANEK, *On semi-infinite systems of linear inequalities*, Israel J. Math., 10 (1971), pp. 252–258.
[6.34] C. KALLINA AND A. C. WILLIAMS, *Linear programming in reflexive spaces*, SIAM Rev., (1971), pp. 350–376.
[6.35] D. F. KARNEY, *Duality gaps in semi-infinite linear programming, an approximation problem*, Math. Programming, 20 (1981), pp. 129–143.
[6.36] ———, *Clark's theorem for semi-infinite convex programs*, Adv. Appl. Math., 2 (1981), pp. 7–13.
[6.37] ———, *A duality theorem for semi-infinite convex programs and their finite subprograms*, Math. Programming, 27 (1983), pp. 75–82.
[6.38] ———, *In a semi-infinite program only a countable subset of the constraints is essential*, J. Approx. Theory, 44 (1985), pp. 69–72.
[6.39] K. O. KORTANEK, *Classifying convex extremum problems over linear topologies having separation properties*, J. Math. Anal. Appl., 46 (1974), pp. 725–755.
[6.40] ———, *Extended abstract of classifying convex extremum problems*, Zentralblatt f. Math., 283 (1975), pp. 491–496.
[6.41] ———, *Perfect duality in generalized linear programming*, in Proc. 9th Intern. Symp. Math. Prog., A. Prékopa, ed., Hungarian Academy of Sci., Budapest, Amsterdam, North-Holland, 1976, pp. 43–58.
[6.42] ———, *Constructing a perfect duality in infinite programming*, Appl. Math. Optim., 3 (1977), pp. 357–372.
[6.43] ———, *Perfect duality in semi-infinite and generalized convex programming*, in Methods of Operations Research 25, A. Angermann, R. Kaerkes, K.-P. Kistner, K. Neumann, B. Rauhut, and F. Steffens, Eds., Meisenheim am Glan: Verlag Anton Hain, 1977, pp. 79–87.
[6.44] K. O. KORTANEK AND H. M. STROJWAS, *On constraint sets of infinite linear programs over ordered fields*, Math. Programming, 33 (1985), pp. 146–161.
[6.45] K. O. KORTANEK AND JISHAN ZHU, *New purification algorithms for linear programming*, Nav. Res. Logist., 35 (1988), pp. 571–683.
[6.46] K. O. KORTANEK AND W. O. ROM, *Classification schemes for the strong duality of linear programming over cones*, Oper. Res., 19 (1971), pp. 1571–1585.
[6.47] R. LEHMANN AND W. OETTLI, *The theorem of the alternative, the key theorem, and the vector-maximum problem*, Math. Programming, 8 (1975), pp. 332–344.
[6.48] P. O. LINDBERG, *Fenchel duality from LP duality*, Optimization Math. OP & Stat., 11 (1980), pp. 171–180.
[6.49] M. A. LOPEZ AND E. VERCHER, *Optimality conditions for nondifferentiable convex semi-infinite programming*, Math. Programming, 27 (1983), pp. 307–319.
[6.50] W. W. ROGOSINSKY, *Moments of non-negative mass*, Proc. Roy. Soc. London Ser. A, 245 (1958), pp. 1–27.
[6.51] R. TRÖLTZSCH, *Duality theorems for a class of continuous linear programming problems in a space of Bochner-integrable abstract functions*, Math. Operationsforsch. Statist. Ser. Optim., 11 (1980), pp. 375–388.
[6.52] R. M. VAN SLYKE AND R. J. B. WETS, *A duality theory for abstract mathematical programs with applications to optimal control theory*, J. Math. Appl., 22 (1968), pp. 679–706.
[6.53] M. YAMASAKI, *Duality theorems in mathematical programming and their applications*, J. Sci. Hiroshima Univ., 30 (1968), pp. 331–356.
[6.54] Y. J. ZHU, *Generalization of some fundamental theorems on linear inequalities*, Acta Math. Sinica, 16 (1966), pp. 25–40.
[7] *Numerical methods*
[7.1] E. J. ANDERSON AND A. S. LEWIS, *An extension of the simplex algorithm for semi-infinite linear programming*, Math. Programming, 44 (1989), pp. 247–269.
[7.2] R. M. CHAMBERLAIN, M. J. D. POWELL, C. LEMARECHAL, AND H. C. PEDERSEN, *The watchdog technique for forcing convergence in algorithms for constrained optimization*, Math. Programming Study, 16 (1982), pp. 1–17.
[7.3] E. W. CHENEY AND A. A. GOLDSTEIN, *Newton's method for convex programming and Tchebycheff approximation*, Numer. Math., 1 (1959), pp. 253–268.

[7.4] A. R. Conn and N. I. M. Gould, *An exact penalty function for semi-infinite programming*, Math. Programming, 37 (1987), pp. 19–40.

[7.5] I. D. Coope and G. A. Watson, *A projected Lagrangian algorithm for semi-infinite programming*, Math. Programming, 32 (1985), pp. 337–356.

[7.6] J. Elzinga and Th. G. Moore, *A central cutting plane algorithm for the convex programming problem*, Math. Programming, 8 (1975), pp. 134–145.

[7.7] K. Fahlander, *Computer programs for semi-infinite optimization*, TRITA-AN-7312, Dept. of Numerical Analysis, Royal Institute of Technology, S-10044 Stockholm 70, 1973.

[7.8] M. C. Ferris and A. B. Philpott, *An interior point algorithm for semi-infinite programming*, Math. Programming, 43 (1989), pp. 257–276.

[7.9] R. Fontecilla, T. Steihaug, and R. A. Tapia, *A convergence theory for a class of Quasi-Newton methods for constrained optimization*, SIAM J. Numer. Anal., 24 (1987), pp. 1133–1151.

[7.10] M. A. Goberna and M. A. Lopez, *Reduction and discrete approximation in linear semi-infinite programming*, Optimization, 18 (1987), pp. 643–658.

[7.11] M. A. Goberna and V. Jornet, *Geometric fundamentals of the simplex method in semi-infinite programming*, Dept. of Math. and Stat., Univ. Alicante, Spain, preprint.

[7.12] C. Gonzaga and E. Polak, *On constraint dropping schemes and optimality functions for a class of outer approximations algorithms*, SIAM J. Control Optim., 17 (1979), pp. 477–493.

[7.13] C. Gonzaga, E. Polak, and R. Trahan, *An improved algorithm for optimization problems with functional inequality constraints*, IEEE Trans. Automat. Control, AC-25 (1980), pp. 49–54.

[7.14] G. Gramlich, SQP-*Methoden für semi-infinite Optimierungsprobleme*, Ph.D. thesis, University of Trier, 1990.

[7.15] G. Gramlich, R. Hettich, and E. Sachs, SQP *methods in semi-infinite programming*, in preparation.

[7.16] P. R. Gribik, *A central cutting plane algorithm for semi-infinite programming problems*, in Semi-Infinite Programming, R. Hettich, ed., Lecture Notes in Control and Inform. Sci. 15, Springer-Verlag, Berlin, Heidelberg, New York, 1979, pp. 66–82.

[7.17] P. R. Gribik and D. N. Lee, *A comparison of two central cutting plane algorithms for prototype geometric programming problems*, in Meth. Oper. Res., W. Oettli and F. Steffens, eds., 31 (1978), pp. 275–287.

[7.18] S.-Å. Gustafson, *Nonlinear systems in semi-infinite programming*, in Numerical Solutions of Nonlinear Algebraic Systems, G. B. Byrnes and C. A. Hall, eds., Academic Press, New York, 1973, pp. 63–99.

[7.19] ———, *A three-phase algorithm for semi-infinite programs*, in Semi-Infinite Programming and Applications, A. V. Fiacco and K. O. Kortanek, eds., Lecture Notes in Econom. and Math. Systems 215, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1983, pp. 138–157.

[7.20] S.-Å. Gustafson and K. O. Kortanek, *Numerical solution of a class of semi-infinite programming problems*, Naval Res. Logist., 20 (1973), pp. 477–504.

[7.21] ———, *Numerical solution of a class of convex programs*, Meth. Oper. Res., 16 (1973), pp. 138–149.

[7.22] S. P. Han, *A globally convergent method for nonlinear programming*, J. Optim. Theory Appl., 22 (1977), pp. 297–309.

[7.23] R. Hettich, *A Newton method for nonlinear Chebyshev approximation*, in Approximation Theory, Lecture Notes in Math. 556, R. Schaback and K. Scherer, eds., Springer-Verlag, Berlin, Heidelberg, New York, 1976, pp. 222–236.

[7.24] ———, *A comparison of some numerical methods for semi-infinite programming*, in Semi-Infinite Programming, R. Hettich, ed., Lecture Notes in Control and Inform. Sci. 15, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1979, pp. 112–125.

[7.25] R. Hettich, *A review of numerical methods for semi-infinite optimization*, in Semi-Infinite Programming and Applications, A. V. Fiacco and K. O. Kortanek, eds., Lecture Notes in Econom. and Math. Systems 215, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1983, pp. 158–178.

[7.26] ———, *An implementation of a discretization method for semi-infinite programming*, Math. Programming, 34 (1986), pp. 354–361.

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

[7.28] R. Hettich and W. Van Honstede, *On quadratically convergent methods for semi-infinite programming*, in Semi-Infinite Programming, R. Hettich, ed., Lecture Notes in Control and Inform. Sci. 15, Springer-Verlag, Berlin, Heidelberg, New York, 1979, pp. 97–111.

[7.29] R. Hettich and P. Zencke, *Superlinear konvergente Verfahren für semi-infinite Optimierungsprobleme im stark eindeutigen Fall*, in Konstruktive Methoden der finiten nichtlinearen Optimierung, L. Collatz, G. Meinardus, and W. Wetterling, eds., Birkhäuser-Verlag, 55 (1980), pp. 103–120.

[7.30] K. H. Hoffmann and A. Klostermaier, *A semi-infinite linear programming procedure and applications to approximation problems in optimal control*, in Approximation Theory II, Lorentz et al., eds., Academic Press, New York, San Francisco, London, pp. 379–389.

[7.31] W. Van Honstede, *An approximation method for semi-infinite programs*, in Semi-Infinite Programming, R. Hettich, Ed., Lecture Notes in Control and Inform. Sci. 15, Springer-Verlag, Berlin, Heidelberg, New York, 1979, pp. 126–136.

[7.32] H. Hu, *A one-phase algorithm for semi-infinite linear programming*, Math Programming, 46 (1990), pp. 85–104.

[7.33] M. Huth, *Superlinear konvergente Verfahren zur Lösung semi-infiniter Optimierungsaufgaben*, Ph.D. dissertation, Pädagog. Hochsch. Halle, 1987.

[7.34] A. Kaplan and R. Tichatschke, *Iterative methods for solving incorrect convex variational problems*, Sov. Dokl., 315 (1990), pp. 275–278.

[7.35] J. E. Kelley, Jr., *The cutting plane method for solving convex programs*, SIAM J., 8 (1960), pp. 703–712.

[7.36] K. O. Kortanek and Hoon No, *A central cutting plane algorithm for convex semi-infinite programming problems*, SIAM J. Optimization, to appear.

[7.37] P. J. Laurent and C. Carasso, *An algorithm of successive minimization in convex programming*, RAIRO Numer. Anal., 12 (1978), pp. 377–400.

[7.38] D. Q. Mayne and E. Polak, *A quadratically convergent algorithm for solving infinite dimensional inequalities*, J. Appl. Math. Optim., 9 (1982), pp. 25–40.

[7.39] D. Q. Mayne, E. Polak, and R. Trahan, *An outer approximations algorithm for computer-aided design problems*, J. Appl. Optim., 28 (1979), pp. 331–352.

[7.40] V. Nebeling, *Eine Schnittmethode zur Lösung semi-infiniter quadratischer Optimierungsaufgaben*, Diss. A, Pädagog. Hochsch. Halle, 1985.

[7.41] K. Oettershagen, *Ein superlinear konvergenter Algorithmus zur Lösung semi-infiniter Optimierungsprobleme*, Ph.D. dissertation, Universität Bonn, 1982.

[7.42] T. Pietrzykowski, *The potential method for conditional maxima in the locally compact metric spaces*, Numer. Math., 14 (1970), pp. 325–329.

[7.43] E. Polak, *On the mathematical foundations of nondifferentiable optimization in engineering design*, SIAM Rev., 29 (1987), pp. 21–89.

[7.44] ———, *On the use of consistent approximations in the solution of semi-infinite optimization and optimal control problems*, Memo. No. UCB/ERL M92/36, Univ. of California, Berkeley, 1992.

[7.45] E. Polak and L. He, *A unified Phase* I – *Phase* II *method of feasible directions for semi-infinite optimization*, Memo. No. UCB/ERL M89/7, Univ. of California, Berkeley, 1989.

[7.46] ———, *Rate preserving discretization strategies for semi-infinite programming and optimal control*, Memo. No. UCB/ERL M89/112, Univ. of California, Berkeley, CA, 1989.

[7.47] ———, *Algorithm models for nondifferentiable optimization*, SIAM J. Control Optim., 23 (1985), pp. 477–491.

[7.48] E. Polak, L. He, and J. E. Higgins, *On the extension of Newton's method to semi-infinite minimax problems*, Math. Programming, to appear.

[7.49] E. Polak, L. He, and Y. Wardi, *On the extension of constrained optimization algorithms from differentiable to nondifferentiable problems*, SIAM J. Control Optim., 21 (1983), pp. 179–204.

[7.50] E. Polak and A. L. Tits, *A recursive quadratic programming algorithm for semi-infinite optimization problems*, J. Appl. Math. Optim., 8 (1982), pp. 325–349.

[7.51] M. J. D. Powell, *A fast algorithm for nonlinearly constrained optimization calculations*, in Numerical Analysis, G. A. Watson, ed., Lecture Notes in Math. 630, Springer-Verlag, Berlin, 1978, pp. 144–157.

[7.52] C. J. Price and I. D. Coope, *An exact penalty function algorithm for semi-infinite programs*, BIT, 30 (1990), pp. 723–734.

[7.53] K. Roleff, *A stable multiple exchange algorithm for linear* SIP, in Semi-Infinite Programming, R. Hettich, ed., Lecture Notes in Control and Inform. Sci. 15, Springer-Verlag, Berlin, Heidelberg, New York, 1979, pp. 83–96.

[7.54] Y. Tanaka, M. Fukushima, and T. Ibaraki, *A globally convergent* SQP *method for semi-infinite nonlinear optimization*, J. Comp. Appl. Math., 23 (1988), pp. 141–153.

[7.55] R. Tichatschke, *Untersuchungen zur numerischen Lösung semi-infiniter linearer Optimierungsaufgaben*, Diss. B., Techn. Hochsch. Karl, Marx, Stadt, 1978.

[7.56] R. Tichatschke and H. Kirsten, *About the efficiency of a method of feasible directions for solving variational inequalities*, in Recent Advances and Historical Development of Vector Optimization, J. Jahn and W. Knabs, eds., Lecture Notes in Econom. and Math. Systems 194, Springer-Verlag, Berlin, Heidelberg, New York, 1987, pp. 379–400.

[7.57] R. Tichatschke and B. Schwartz, *A method for solving convex semi-infinite programming problems*, Part I Optimisacija, 43 (1988), pp. 62–73 and II Optimisacija, 44 (1988), pp. 64–81.

[7.58] A. L. Tits, *Lagrangian based superlinearly convergent algorithms for ordinary and semi-infinite optimization problems*, Ph.D. dissertation, Dept. of Electrical Engineering and Computer Science, University of California, Berkeley, CA, 1980.

[7.59] H. J. Töpfer, *Tschebyscheff–Approximation bei nicht erfüllter Haarscher Bedingung*, Internat. Ser. Numer. Math., 7 (1967), pp. 71–89.

[7.60] R. Trahan and E. Polak, *A derivative free algorithm for a class of infinitely constrained problems*, IEEE Trans. Automat. Control, AC-25 (1980), pp. 54–61.

[7.61] G. A. Watson, *Globally convergent methods for semi-infinite programming*, BIT, 21 (1981), pp. 362–373.

[7.62] ———, *Numerical experiments with globally convergent methods for semi-infinite programming problems*, in Semi-Infinite Programming and Applications, A. V. Fiacco and K. O. Kortanek, eds., Lecture Notes in Econom. and Math. Systems 215, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1983, pp. 193–205.

[7.63] ———, *Lagrangian methods for semi-infinite programming problems*, in Infinite Programming, E. J. Anderson and A. B. Philpott, eds., Lecture Notes in Econom. and Math. Systems 259, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1985, pp. 90–107.

[7.64] W. Wetterling, *Anwendung des Newtonschen Iterationsverfahrens bei der Tschebyscheff–Approximation, insbesondere mit nichtlinear auftretenden Parametern*, MTW Teil I: pp. 61–63; Teil II: pp. 112–115, 1963.

[8] *Parametric problems*

[8.1] B. Brosowski, *Zur stetigen Parameterabhängigkeit der Menge der Minimalpunkte bei Minimaxaufgaben*, Meth. Oper. Res., 25 (1976), pp. 28–33.

[8.2] ———, *An optimality condition and its application to parametric semi-infinite optimization*, Oper. Res. in progress; Theory Decis. Libr. 32, 1982, pp. 3–16.

[8.3] ———, *Parametric semi-infinite programming*, Methoden und Verfahren der Mathematischen Physik, Bd. 22, Verlag Peter Lang, Frankfurt (Main)-Bern, 1982.

[8.4] ———, *Parametric semi-infinite linear programming* I: *continuity of the feasible set and the optimal value*, Math. Progr. Stud., 21 (1984), pp. 18–42.

[8.5] ———, *A refinement of an optimality criterion and its application to parametric programming*, J. Optim. Theory Appl., 42 (1984), pp. 367–382.

[8.6] R. Colgen, *Necessary conditions for upper semicontinuity in parammetric semi-infinite programming*, J. Optim. Theory Appl., 48 (1986), pp. 65–79.

[8.7] ———, *Necessary conditions for upper semicontinuity in parametric semi-infinite programming*, J. Optim. Theory Appl., to appear.

[8.8] R. Colgen and K. Schnatz, *Continuity properties in semi-infinite parametric linear optimization*, Numer. Funct. Anal. Optim., 3 (1981), pp. 451–460.

[8.9] A. V. Fiacco, *Introduction to Sensitivity and Stability Analysis in Nonlinear Programming*, Academic Press, New York, 1983.

[8.10] T. Fischer, *Stetige Selektionen bei semi-infiniter Optimierung*, R. G. Fischer-Verlag, Frankfurt/Main, 1986.

[8.11] H. Gfrerer, J. Guddat, H. J. Wacker, and W. Zulehner, *Path-following methods for Kuhn–Tucker curves by an active index set strategy*, in Systems and Optimization, A. Bagchi and H. Th. Jongen, eds., Lecture Notes in Control and Inform. Sci. 66, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1985, pp. 111– 131.

[8.12] J. Guddat, *Parametric optimization: pivoting and predictor-corrector continuation—a survey*, in Parametric Optimization and Related Topics, J. Guddat, H. Th. Jongen, B. Kummer, and F. Nozicka, eds., Math. Res. 35, Akademie-Verlag, Berlin, 1987, pp. 309–363.

[8.13] J. Guddat, F. Guerra Vasquez, and H. Th. Jongen, *Parametric Optimization: Singularities, Pathfollowing and Jumps*, Teubner, Stuttgart and John Wiley & Sons, Chichester, New York, Brisbane, Toronto, Singapore, 1990.

[8.14] J. Guddat, H. Th. Jongen, B. Kummer, and F. Nozicka, eds., *Parametric Optimization and Related Topics*, Math. Res. 35, Akademie-Verlag, Berlin, 1987.

[8.15] S. Helbig, *Parametrische semi-infinite Optimierung in total geordneten Gruppen*, R. G. Fischer-Verlag, Frankfurt (Main), 1987.

[8.16] ———, *Parametric semi-infinite optimization in certain lattices: Continuity of the feasible set*, Trends in Mathematical Optimization, 4th French–German Conf., Irsee/FRG 1986, Internat. Ser. Numer. Math., 84 (1988), pp. 103–115.

[8.17] R. Hettich and P. Zencke, *Two case-studies in parametric semi-infinite programming*, in Systems and Optimization, A. Bagchi, H. Th. Jongen, eds., Lecture Notes in Control and Inform. Sci. 66, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1985.

[8.18] W. W. Hogan, *Directional derivatives for extremal-value functions with applications to the completely convex case*, Oper. Res., 21 (1973), pp. 188–209.

[8.19] H. Th. Jongen, P. Jonker, and F. Twilt, *Critical sets in parametric optimization*, Math. Programming, 34 (1986), pp. 333–353.

[8.20] H. Th. Jongen, F. Twilt, and G. W. Weber, *Semi-infinite optimization: Structure and stability of the feasible set*, Memorandum No. 838, University of Twente, the Netherlands, 1989.

[8.21] M. Kojima, *Strongly stable stationary solutions in nonlinear programs*, in Analysis and Computation of Fixed Points, S. M. Robinson, ed., Academic Press, New York, 1980, pp. 93–138.

[8.22] R. Lehmann, *An algorithm for solving one-parametric optimization problems based on an active index set strategy*, in Parametric Optimization and Related Topics, J. Guddat, H. Th. Jongen, B. Kummer, and F. Nozicka, eds., Math. Res. 35, Akademie-Verlag, Berlin, 1987, pp. 268–301.

[8.23] G. Nuernberger, *Parametric optimization and related topics*, Akademie-Verlag Math. Res., 35 (1987), pp. 316–326.

[8.24] S. M. Robinson, *Strongly regular generalized equations*, Math. Oper. Res., 5 (1980), pp. 43–62.

[8.25] Th. Rupp, *Kontinuitätsmethoden zur Lösung einparametrischer semi-infiniter Optimierungsprobleme*, thesis, University of Trier, Germany, 1988.

[8.26] ———, *Kuhn–Tucker curves for one-parametric semi-infinite programming*, Optimization, 20 (1989), pp. 61–77.

[8.27] P. Zencke and R. Hettich, *Directional derivatives for the value function in semi-infinite programming*, Math. Programming, 38 (1987), pp. 323–340.