

A modified standard embedding with jumps in nonlinear optimization*

Jürgen Guddat¹⁾
Francisco Guerra Vázquez²⁾
Dieter Nowack¹⁾
Jan-J. Rückmann²⁾

Abstract. The paper deals with a combination of pathfollowing methods (embedding approach) and feasible descent direction methods (so-called jumps) for solving a nonlinear optimization problem with equality and inequality constraints. Since the method that we propose here uses jumps from one connected component to another one, more than one connected component of the solution set of the corresponding one-parametric problem can be followed numerically. It is assumed that the problem under consideration belongs to a generic subset which was introduced by Jongen, Jonker and Twilt. There already exist methods of this type for which each starting point of a jump has to be an endpoint of a branch of local minimizers. In this paper the authors propose a new method by allowing a larger set of starting points for the jumps which can be constructed at bifurcation and turning points of the solution set. The topological properties of those cases where the method is not successful are analyzed and the role of constraint qualifications in this context is discussed. Furthermore, this new method is applied to a so-called modified standard embedding which is a particular construction without equality constraints. Finally, an algorithmic version of this new method as well as computational results are presented.

Key words: Parametric programming, pathfollowing methods with jumps, genericity, Jongen-Jonker-Twilt regularity, modified standard embedding

AMS subject classification: 90 C 31, 90 C 26, 90 C 30, 65 K 05, 49 M 37

^{*)} This work was partially supported by Deutsche Forschungsgemeinschaft (DFG) under grant Gu 304/14-1 and the Sistema Nacional de Investigadores (SNI, México)

¹⁾ Humboldt-Universität zu Berlin, Institut für Mathematik, 10099 Berlin, Germany.

Emails: guddat@mathematik.hu-berlin.de, nowack@mathematik.hu-berlin.de

²⁾ Universidad de las Américas, Escuela de Ciencias, Sta. Catarina Mártir, Cholula, Puebla 72820, México. *Emails:* fguerra@mail.udlap.mx, rueckman@mail.udlap.mx

1 Introduction

Let \mathbb{R}^n be the n -dimensional space with the Euclidean norm $\|\cdot\|$ and $C^k(\mathbb{R}^n, \mathbb{R})$, $k \geq 1$ the space of k -times continuously differentiable functions. In this paper we consider the nonlinear optimization problem

$$(P) \quad \min \{f(x) \mid x \in M\} \quad (1.1)$$

where the nonempty feasible set is defined by finitely many equality and inequality constraints as

$$M = \{x \in \mathbb{R}^n \mid h_i(x) = 0, \ i \in I, \ g_j(x) \leq 0, \ j \in J\}$$

with $I = \{1, \dots, m\}$, $m < n$, $J = \{1, \dots, s\}$, and $f, h_i, g_j \in C^3(\mathbb{R}^n, \mathbb{R})$, $i \in I$, $j \in J$. Furthermore, we introduce the one-parametric nonlinear optimization problem

$$P(t) \quad \min \{f(x, t) \mid x \in M(t)\} \quad (1.2)$$

where $t \in \mathbb{R}$ is a real parameter,

$$M(t) = \{x \in \mathbb{R}^n \mid h_i(x, t) = 0, \ i \in I, \ g_j(x, t) \leq 0, \ j \in J\}$$

and $f, h_i, g_j \in C^3(\mathbb{R}^n \times \mathbb{R}, \mathbb{R})$, $i \in I$, $j \in J$. For sake of simplicity we assume that all functions in (1.1) and (1.2) are three times continuously differentiable although some of the results given here also hold for a lower degree of differentiability.

The *embedding approach* is a well-known method for the calculation of a solution point (local minimizer, stationary point, generalized critical point, etc.) of (P) ; its basic idea is to construct an auxiliary problem $P(t)$ which satisfies at least the following three conditions:

(A1) A solution point x^0 of $P(0)$ is known.

(A2) The set of solution points of $P(t)$ is nonempty for all $t \in [0, 1]$.

(A3) $P(1)$ is similar (in a certain sense) or equivalent with (P) .

Then, by using a so-called pathfollowing (or homotopy or continuation) method a solution point x^* of the original problem (P) can be obtained by following numerically a solution path connecting $(x^0, 0)$ and $(x^*, 1)$, i.e. one has to find a discretization

$$0 = t_0 < \dots < t_i < \dots < t_N = 1$$

of the interval $[0, 1]$ and corresponding solution points $x(t_i)$ of $P(t_i)$, $i = 0, \dots, N$ (cf. e.g. [1, 2, 3, 4, 5, 7, 8, 10, 12, 16, 18, 19, 21, 23]).

Example 1.1 As an example we present the so-called *standard embedding* which is defined by the one-parametric problem

$$P_{x^0}(t) \quad \min \{tf(x) + (1-t)\|x - x^0\|^2 \mid x \in M_{x^0}(t)\}$$

with the starting point $x^0 \in \mathbb{R}^n$ and the feasible set

$$M_{x^0}(t) = \left\{ x \in \mathbb{R}^n \left| \begin{array}{l} h_i(x) + (t-1)h_i(x^0) = 0, \quad i \in I \\ g_j(x) + (t-1)g_j(x^0) \leq 0, \quad j \in J \end{array} \right. \right\}.$$

Obviously, (A1) and (A3) are satisfied but (A2) cannot be guaranteed in general (cf. Example 5.1). In particular, the feasible set could be empty for some parameter values $t \in (0, 1)$.

However, in general, the existence of a solution curve to be followed is a very strong condition. In [16, 18], topological conditions are discussed which ensure an appropriate structure of the solution set of $P(t)$ (union of one-dimensional manifolds) for the use of pathfollowing methods. In particular, Jongen, Jonker and Twilt defined in [16] a particular open and dense subset $\mathcal{F} \subset C^3(\mathbb{R}^n \times \mathbb{R}, \mathbb{R})^{1+m+s}$ and described the corresponding topological structure of the solution set Σ_{gc} of $P(t)$ by defining five different types (the set of nondegenerate points and the singularities which may appear) such that Σ_{gc} can be divided in five disjoint subsets.

Assuming that the function vector $(f, h_i, g_j, i \in I, j \in J)$ in (1.2) belongs to this generic class \mathcal{F} , Guddat, Guerra Vázquez and Jongen presented in [12] some solution methods for $P(t)$ which combine pathfollowing methods with so-called “jumps”, where a “jump” refers to an appropriate feasible descent direction method (or, more general, NLP-solver) with the objective to calculate a solution point which belongs to another connected component of Σ_{gc} (“jump from one connected component of Σ_{gc} to another one”). Then, more than one connected component of Σ_{gc} can be followed numerically and, hence, there are more chances to attain the parameter value $t = 1$ (note that standard pathfollowing methods are restricted to one connected component only). In [12], the jumps are defined for points (starting points for the feasible descent direction method) which are end points of branches of local minimizers. By applying theoretical results from [14], in this paper we generalize the mentioned approach from [12] and introduce the method JUMP II* that combines pathfollowing methods with jumps which are defined for a larger class of starting points. By means of this new method (perhaps) more connected components of Σ_{gc} can be detected and numerically described. Besides the goal of attaining the parameter value $t = 1$ (in the context of the embedding approach) another motivation for combining pathfollowing methods with jumps is the solution of optimization problems that depend naturally on a parameter; here, one is often interested in following more than one (or as many as possible) connected components of the solution set.

The goal of this paper is twofold:

Firstly, we discuss a particular one-parametric problem $P^m(t)$ which is called *modified standard embedding*; its construction yields that the feasible set is nonempty for all $t \in [0, 1]$ (which is a necessary condition for (A2) and that, in general, is not satisfied for the standard embedding given in Example 1.1). Furthermore, an important feature for the use of feasible descent direction methods is that $P^m(t)$ only has inequality constraints although the original problem (P) may also have equality constraints.

Secondly, we introduce the method JUMP II* which combines pathfollowing methods with feasible descent direction methods (“jumps”) where these jumps are defined for a larger class of starting points (not only for those which are end points of a branch of local minimizers). The goal of JUMP II* is to follow as many connected components of Σ_{gc} (restricted to a given parameter interval) as possible. We analyze the topological situations for which we cannot attain the parameter value $t = 1$ and discuss the role of constraint qualifications in this context.

The paper is organized as follows. Section 2 contains basic results and we recall the generic class \mathcal{F} which was introduced by Jongen, Jonker and Twilt (cf. [16]). In Section 3 we define the modified standard embedding and discuss its properties. Section 4 deals with the method JUMP II* which combines pathfollowing methods with jumps which are defined for solution points of Types 1-4. Each possible situation for a jump is characterized and an algorithmic version of JUMP II* is presented. Finally, Section 5 contains computational results which illustrate the application of JUMP II* to the modified standard embedding. These numerical tests have been realized by means of the program package PAFO [9].

2 Notations and theoretical background

In this section we present some notations, basic results and we recall the open and dense class \mathcal{F} of functions which was introduced by Jongen, Jonker and Twilt [16].

Critical point sets

Throughout the paper the problems (P) and $P(t)$ are defined as in (1.1) and (1.2), respectively. For $\bar{x} \in M(\bar{t})$ we denote the index set of active inequality constraints by

$$J_0(\bar{x}, \bar{t}) = \{j \in J \mid g_j(\bar{x}, \bar{t}) = 0\}.$$

A point $(\bar{x}, \bar{t}) \in \mathbb{R}^n \times \mathbb{R}$ is called a *generalized critical point* (*gc point*) of $P(\bar{t})$ (cf. [12, 16]) if $\bar{x} \in M(\bar{t})$ and the gradients $D_x f(\bar{x}, \bar{t})$, $D_x h_i(\bar{x}, \bar{t})$, $i \in I$, $D_x g_j(\bar{x}, \bar{t})$, $j \in J_0(\bar{x}, \bar{t})$ are linearly dependent. Obviously, if \bar{x} is a stationary point (for a definition, see [12, p.1]) or a local minimizer of $P(\bar{t})$, then (\bar{x}, \bar{t}) is also a gc point of $P(\bar{t})$. We introduce the following sets for the problems (P) and $P(t)$:

$$\begin{aligned} \Psi_{gc}(P(t)) &= \{(x, t) \in \mathbb{R}^n \times \mathbb{R} \mid (x, t) \text{ is a gc point of } P(t)\}, \\ \Psi_{stat}(P(t)) &= \{x \in \mathbb{R}^n \mid x \text{ is a stationary point of } P(t)\}, \\ \Psi_{loc}(P(t)) &= \{x \in \mathbb{R}^n \mid x \text{ is a local minimizer of } P(t)\}, \\ \Psi_{glob}(P(t)) &= \{x \in \mathbb{R}^n \mid x \text{ is a global minimizer of } P(t)\}, \\ \Psi_{stat}(P) &= \{x \in \mathbb{R}^n \mid x \text{ is a stationary point of } (P)\}, \\ \Psi_{loc}(P) &= \{x \in \mathbb{R}^n \mid x \text{ is a local minimizer of } (P)\}. \end{aligned}$$

Furthermore, define the unfolded sets

$$\begin{aligned} \Sigma_{gc} &= \{(x, t) \in \mathbb{R}^n \times \mathbb{R} \mid (x, t) \in \Psi_{gc}(P(t))\}, \\ \Sigma_{stat} &= \{(x, t) \in \mathbb{R}^n \times \mathbb{R} \mid x \in \Psi_{stat}(P(t))\}, \\ \Sigma_{loc} &= \{(x, t) \in \mathbb{R}^n \times \mathbb{R} \mid x \in \Psi_{loc}(P(t))\}. \end{aligned}$$

Constraint qualifications

We will use the subsequent well-known constraint qualifications LICQ and MFCQ.

The *Linear Independence constraint qualification* (briefly, *LICQ*) is said to hold at $\bar{x} \in M(\bar{t})$ if the gradients $D_x h_i(\bar{x}, \bar{t})$, $i \in I$, $D_x g_j(\bar{x}, \bar{t})$, $j \in J_0(\bar{x}, \bar{t})$ are linearly independent.

The *Mangasarian-Fromovitz constraint qualification* (briefly, *MFCQ*) holds at $\bar{x} \in M(\bar{t})$ if

(MF1) the gradients $D_x h_i(\bar{x}, \bar{t})$, $i \in I$ are linearly independent,

(MF2) there exists a vector $\xi \in \mathbb{R}^n$ such that

$$\begin{aligned} D_x h_i(\bar{x}, \bar{t})\xi &= 0, & i \in I, \\ D_x g_j(\bar{x}, \bar{t})\xi &< 0, & j \in J_0(\bar{x}, \bar{t}). \end{aligned}$$

Obviously (cf. e.g. [20]), if $\bar{x} \in \Psi_{loc}(P(\bar{t}))$ and MFCQ or LICQ holds at $\bar{x} \in M(\bar{t})$, then $\bar{x} \in \Psi_{stat}(P(\bar{t}))$.

The generic class \mathcal{F} of Jongen, Jonker and Twilt

In the sequel we consider the function vector $(f, H, G) \in C^3(\mathbb{R}^n \times \mathbb{R})^{1+m+s}$ which characterizes the one-parametric problem $P(t)$, where $H(x, t) = (h_1(x, t), \dots, h_m(x, t))$ and $G(x, t) = (g_1(x, t), \dots, g_s(x, t))$.

It is well-known that the use of pathfollowing methods for the calculation of gc points (or stationary points or local minimizers) of $P(t)$ requires a particular structure of the corresponding set Σ_{gc} (or Σ_{stat} or Σ_{loc}). In [16], Jongen, Jonker and Twilt introduced an open and dense subset \mathcal{F} of $C^3(\mathbb{R}^n \times \mathbb{R}, \mathbb{R})^{1+m+s}$ which can be described by the topological structure of the corresponding set Σ_{gc} for any $(f, H, G) \in \mathcal{F}$. In particular, $(f, H, G) \in \mathcal{F}$ implies that Σ_{gc} consists locally of a finite union of one-dimensional manifolds and, therefore, it has an appropriate structure for the use of pathfollowing methods. Below we cite our very short characterization from [12] of the class \mathcal{F} ; the complete description can be found in [16] (see [10] as well).

If $(f, H, G) \in \mathcal{F}$, then Σ_{gc} can be divided into five types of gc points:

Type 1: A point $\bar{z} = (\bar{x}, \bar{t}) \in \Sigma_{gc}$ is of Type 1 (non-degenerate gc point) if the following conditions are satisfied:

There exist $\bar{\lambda}_i, \bar{\mu}_j \in \mathbb{R}$, $i \in I$, $j \in J_0(\bar{z})$ with

$$\left(D_x f + \sum_{i \in I} \bar{\lambda}_i D_x h_i + \sum_{j \in J_0(\bar{z})} \bar{\mu}_j D_x g_j \right) \Big|_{z=\bar{z}} = 0, \quad (2.1)$$

$$\text{LICQ is satisfied at } \bar{x} \in M(\bar{t}), \quad (2.2a)$$

(therefore $\bar{\lambda}_i, \bar{\mu}_j$, $i \in I$, $j \in J_0(\bar{z})$ are uniquely defined)

$$\bar{\mu}_j \neq 0, \quad j \in J_0(\bar{z}), \quad (2.2b)$$

$$D_x^2 L(\bar{z}) \Big|_{T(\bar{z})} \text{ is nonsingular,} \quad (2.2c)$$

where D_x^2L is the Hessian of the Lagrangian

$$L(z) = f(z) + \sum_{i \in I} \bar{\lambda}_i h_i(z) + \sum_{j \in J_0(\bar{z})} \bar{\mu}_j g_j(z),$$

and the uniquely determined numbers $\bar{\lambda}_i, \bar{\mu}_j$ are taken from (2.1). Furthermore,

$$T(z) = \{\xi \in \mathbb{R}^n \mid D_x h_i(z)\xi = 0, i \in I, D_x g_j(z)\xi = 0, j \in J_0(z)\}$$

is the tangent space at z . $D_x^2L(z)|_{T(z)}$ represents $V^T D_x^2L V$, where V is a matrix whose columns form a basis of $T(z)$. The nondegeneracy of a point of Type 1 implies that the subset of Σ_{gc} which consists of all points of Type 1 forms a one-dimensional manifold. The points of the Types 2-5 represent four basic degeneracies (see [16] for more details):

Type 2 – violation of (2.2b)

Type 3 – violation of (2.2c)

Type 4 – violation of (2.2a) and $|I| + |J_0(\bar{z})| < n + 1$

Type 5 – violation of (2.2a) and $|I| + |J_0(\bar{z})| = n + 1$.

Figure 2.1 illustrates for each of the five types the local structure of Σ_{gc} in a neighbourhood of a stationary point \bar{z} : the full curve stands for the set Σ_{stat} and the dotted curve represents the set of gc points which are not stationary points.

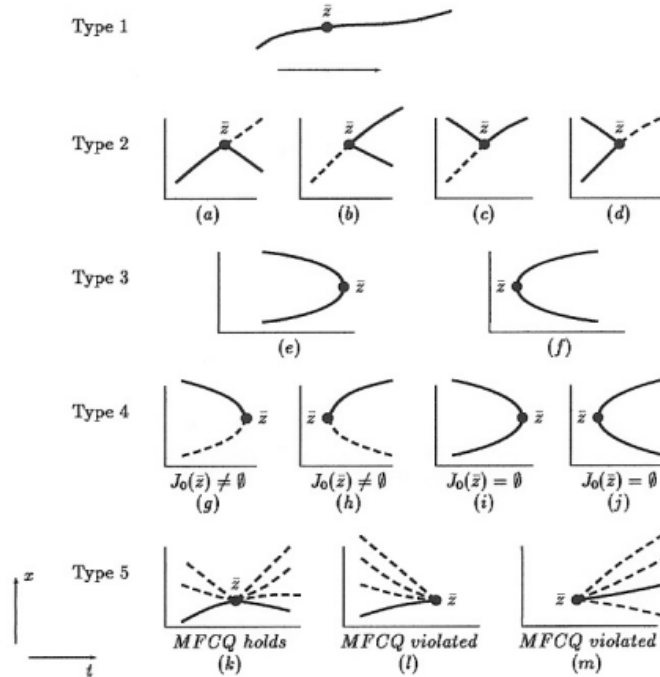


Figure 2.1: The structure of Σ_{gc} in a neighbourhood of $\bar{z} \in \Sigma_{stat}$.

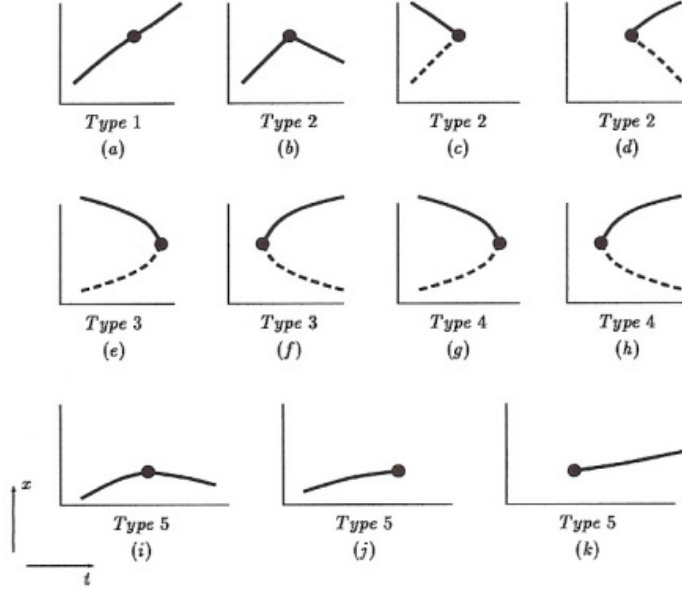


Figure 2.2: The structure of Σ_{stat} in a neighbourhood of a local minimizer.

Figure 2.2 illustrates for each of the five types the local structure of Σ_{stat} in a neighbourhood of a local minimizer: the full curve stands for the set Σ_{loc} and the dotted curve in (c), (d), (e), (f) represents stationary points which are not local minimizers. The dotted curve in (g), (h) represents stationary points (which are not local minimizers) when $J_0(\bar{x}) = \emptyset$.

The points of the Types 2-5 constitute a discrete subset of Σ_{gc} ; in particular, the whole set Σ_{gc} is the closure of the set of all points of Type 1. Therefore, $(f, H, G) \in \mathcal{F}$ implies an appropriate structure of Σ_{gc} for the application of path-following methods.

Let Σ_{gc}^v be that subset of Σ_{gc} which consists of all points of Type v , $v = 1, \dots, 5$. Then, we define the class \mathcal{F} as

$$\mathcal{F} = \{(f, H, G) \in C^3(\mathbb{R}^n \times \mathbb{R}, \mathbb{R})^{1+m+s} \mid \Sigma_{gc} = \bigcup_{v=1}^5 \Sigma_{gc}^v\}.$$

For further analysis we need the following definition:

Definition 2.1 Let $K \subset \mathbb{R}$ be an interval. The one-parametric problem $P(t)$ is called *regular in the sense of Jongen, Jonker and Twilt* (briefly, *JJT-regular*) with respect to K if

$$(\mathbb{R}^n \times K) \cap \Sigma_{gc} = (\mathbb{R}^n \times K) \cap \bigcup_{v=1}^5 \Sigma_{gc}^v.$$

The following two theorems give insight about the assumption that $(f, H, G) \in \mathcal{F}$. The first one states a generic property that we already mentioned above.

Theorem 2.1 (Genericity theorem, cf. [16]). *The class \mathcal{F} is C_s^3 -open and C_s^3 -dense in $C^3(\mathbb{R}^n \times \mathbb{R}, \mathbb{R})^{1+m+s}$, where C_s^3 denotes the strong (or Whitney-) C_s^3 -topology (cf. [15, 17]).*

The next theorem provides a special perturbation of (f, H, G) with additional parameters that can be chosen arbitrarily small and such that the perturbed functions belong to the class \mathcal{F} . Let the space of symmetric $n \times n$ -matrices be identified by $\mathbb{R}^{n(n+1)/2}$.

Theorem 2.2 (Perturbation theorem, cf. [22]). *Let $(f, H, G) \in C^3(\mathbb{R}^n \times \mathbb{R}, \mathbb{R})^{1+m+s}$. Then, for almost all $(b, A, c, D, e, F) \in \mathbb{R}^n \times \mathbb{R}^{n(n+1)/2} \times \mathbb{R}^m \times \mathbb{R}^{mn} \times \mathbb{R}^s \times \mathbb{R}^{sn}$, we have*

$$(f(x, t) + b^T x + x^T A x, H(x, t) + c + D x, G(x, t) + e + F x) \in \mathcal{F}.$$

Here “almost all” means: each measurable subset of

$$\{(b, A, c, D, e, F) \mid (f(x, t) + b^T x + x^T A x, H(x, t) + c + D x, G(x, t) + e + F x) \notin \mathcal{F}\}$$

has the Lebesgue-measure zero.

Exemplarily, the following theorem presents some conditions which imply the existence of a curve in Σ_{stat} connecting a (known) stationary point x^0 of $P(0)$ with a (unknown) stationary point x^* of $P(1)$.

Theorem 2.3 (cf. [7]). *Assume that*

- (C1) $M(t)$ is non-empty and there exists a compact set containing $M(t)$ for all $t \in [0, 1]$.
- (C2) $P(t)$ is JJT-regular with respect to $[0, 1]$.
- (C3) There exists a $t_1 > 0$ and a continuous function $x : [0, t_1) \rightarrow \mathbb{R}^n$ such that $x(t)$ is the unique stationary point for $P(t)$ for $t \in [0, t_1)$.
- (C4) MFCQ holds at all $x \in M(t)$ for all $t \in [0, 1]$.

Then there exists a piecewise C^3 -path (briefly, PC^3 -path) in Σ_{stat} that connects $(x^0, 0)$ (with $x(0) = x^0$) with some point $(x^, 1)$.*

For a discussion of the assumptions of the latter theorem we refer to [7]. However, it can easily be seen from the topological properties of Σ_{gc} that the condition (C4) is generically not fulfilled, i.e. in general it cannot be satisfied by appropriate arbitrarily small perturbations of the function vector (f, H, G) . We will return to this point later in the discussion of the heuristic method JUMP II* in Section 4, Remark 4.5.

3 A modified standard embedding

In order to ensure the existence of a curve in Σ_{gc} connecting a gc point of $P(0)$ with a gc point of $P(1)$ one has to assume that

$$M(t) \neq \emptyset \text{ for all } t \in [0, 1]. \quad (3.1)$$

In general, this assumption (3.1) is not fulfilled for the standard embedding presented in Example 1.1 (see [6] for a corresponding example which will be considered again in Example 5.1). In this section we introduce the so-called modified standard embedding which satisfies (3.1). The *modified standard embedding* $P^m(t)$ for the problem (P) in (1.1) is defined as follows (with $q \in \mathbb{R}$, $q > 0$ and $x^0 \in \mathbb{R}^n$):

$$P^m(t) = \min \{f(x, t) \mid x \in M^m(t)\}, \quad t \in \mathbb{R}, \quad (3.2)$$

where

$$\begin{aligned} f(x, t) &= tf(x) + (1-t)\|x - x^0\|^2, \\ M^m(t) &= \{x \in \mathbb{R}^n \mid g_j(x, t) \leq 0, \quad j = 1, \dots, m+s+2\}, \\ g_i(x, t) &= th_i(x) + t - 1, \quad i \in I, \\ g_{m+j}(x, t) &= tg_j(x) + t - 1, \quad j \in J, \\ g_{m+s+1}(x, t) &= \|x\|^2 - q, \\ g_{m+s+2}(x, t) &= -t \sum_{i \in I} h_i(x) + t - 1. \end{aligned}$$

Remark 3.1 Several modifications of the standard embedding have been discussed in [6]. In (3.2), the additional ‘‘compactification constraint’’ g_{m+s+1} implies that $M(t)$ is compact for all $t \in [0, 1]$; its consequences for the similarity of $P^m(1)$ and (P) (cf. (A3)) are discussed in the subsequent Proposition 3.2. Furthermore, the fact that $P^m(t)$ does not have any equality constraint allows the application of corresponding descent methods for the realization of the so-called jumps (see Section 4). Note that the starting point x^0 can be chosen arbitrarily.

In the next two propositions we will summarize several properties of the modified standard embedding (cf. [6] for a more general discussion).

Proposition 3.1 *Let $q > 0$ in $P^m(t)$ be chosen such that $M \cap \{x \in \mathbb{R}^n \mid \|x\|^2 \leq q\} \neq \emptyset$ and $\|x^0\|^2 < q$. Then we have:*

- (i) $M^m(t)$ is nonempty and compact for all $t \in [0, 1]$. In particular, $M^m(1) = M \cap \{x \in \mathbb{R}^n \mid \|x\|^2 \leq q\}$.
- (ii) $M^m(t_2) \subset M^m(t_1)$ for all $t_1 < t_2$, $t_1, t_2 \in [0, 1]$.
- (iii) $\Psi_{glob}(P^m(0)) = \Psi_{stat}(P^m(0)) = \{x^0\}$.
- (iv) $\Psi_{glob}(P^m(t))$ is nonempty for all $t \in [0, 1]$.
- (v) If $I \neq \emptyset$, then the MFCQ does not hold at any point $x \in M^m(1)$.

(vi) If $I \neq \emptyset$, then $M^m(t) = \emptyset$ for all $t > 1$.

Proof. From the definition of $P^m(t)$ we obtain easily (i) – (iv). Now, let $I \neq \emptyset$. Then we have for $x \in M^m(1)$:

$$g_i(x, 1) = h_i(x) = 0, \quad i \in I \text{ and } g_{m+s+2}(x, 1) = -\sum_{i \in I} h_i(x) = 0.$$

Obviously, there does not exist a vector $\xi \in \mathbb{R}^n$ satisfying (MF2) and we obtain (v). For the proof of (vi) assume that there exist $\bar{t} > 1$ and $\bar{x} \in M^m(\bar{t})$. Then, we obtain the contradiction

$$h_i(\bar{x}) \leq \frac{1 - \bar{t}}{\bar{t}} < 0, \quad i \in I \text{ and } -\sum_{i \in I} h_i(\bar{x}) \leq \frac{1 - \bar{t}}{\bar{t}}. \quad \triangle$$

The next proposition discusses properties of $P^m(0)$ as well as the similarity of the problems $P^m(1)$ and (P) . We omitted the proof because it is obvious.

Proposition 3.2 The following statements are true:

(i) If $0 < \|x^0\|^2 < q$, then each gc point of $P^m(0)$ is non-degenerate (cf. the definition of a gc point of Type 1 in Section 2).

(ii) If $\|\bar{x}\|^2 < q$ and $\bar{x} \in \Psi_{stat}(P^m(1))$ (resp. $\bar{x} \in \Psi_{loc}(P^m(1))$), then $\bar{x} \in \Psi_{stat}(P)$ (resp. $\bar{x} \in \Psi_{stat}(P)$).

(iii) $\Psi_{gc}(P^m(1)) = M^m(1) \times \{1\}$ and $M^m(1) = M \cap \{x \in \mathbb{R}^n \mid \|x\|^2 \leq q\}$.

(iv) Let $\bar{x} \in \Psi_{loc}(P)$ and $\|\bar{x}\|^2 \leq q$. Then, $\bar{x} \in \Psi_{loc}(P^m(1))$.

(v) Let $\bar{x} \in \Psi_{stat}(P)$ and $\|\bar{x}\|^2 \leq q$. Then, $(\bar{x}, 1) \in \Psi_{gc}(P^m(1))$.

(vi) If $M \subseteq \{x \in \mathbb{R}^n \mid \|x\|^2 \leq q\}$, then $P^m(1)$ and (P) are equivalent, i.e. $f(x, 1) = f(x)$ and $M^m(1) = M$. △

Note that, by Proposition 3.1 (iii) and Proposition 3.2 (i), the starting point $(x^0, 0)$ is a non-degenerate gc point and x^0 a global minimizer of $P^m(0)$. Now, assume for a moment that each gc point of $M^m(1)$ is of one of the five types of Jongen, Jonker and Twilt (i.e. it belongs to $\bigcup_{v=1}^5 \Sigma_{gc}^v$, cf. Section 2). Then, by Proposition 3.2 (iii) and Proposition 3.1 (v), we have that $\Psi_{gc}(P^m(1)) = M^m(1) \times \{1\}$ and MFCQ does not hold at each $\bar{x} \in M^m(1)$ (i.e. $(\bar{x}, 1)$ is not a point of Type 1) and, therefore, $M^m(1)$ has to be a discrete subset (the points of the Types 2–5 constitute a discrete subset, cf. Section 2).

Since, in general, $M^m(1)$ is not a discrete subset it makes no sense to assume that the problem $P^m(t)$ is JJT-regular with respect to an interval K that contains the parameter value $t = 1$ (which would imply that each gc point of $P^m(1)$ belongs to $\bigcup_{v=1}^5 \Sigma_{gc}^v$). Having that in mind we modify the Theorem 2.3 as follows.

Corollary 3.1 *Assume that*

(C2*) $P^m(t)$ is JJT-regular with respect to $[0, 1)$.

(C4*) MFCQ holds at all $x \in M^m(t)$ for all $t \in [0, 1)$.

Then, there exists for each $\hat{t} \in (0, 1)$ a PC^3 -path in Σ_{stat} that connects $(x^0, 0)$ with some point (\hat{x}, \hat{t}) .

Proof. It follows directly from Theorem 2.3 and Proposition 3.1 (i), (iii). \triangle

Under the assumptions of Corollary 3.1 a sequence $(x^i, t^i) \in \Sigma_{stat}$ with $t^i \rightarrow 1$ can be created whose limit points are gc points of $P^m(1)$ (since the set Σ_{gc} is closed).

As already mentioned at the end of Section 2, the assumption (C4) of Theorem 2.3 - and, now, the assumption (C4*) of Corollary 3.1 as well - are very strong conditions in the sense that they cannot be satisfied by perturbing (arbitrarily small) the function vector (f, H, G) (cf. Remark 4.5). However, next we present a condition on the constraints $h_i(x)$, $g_j(x)$, $i \in I$, $j \in J$ of (P) (which do not depend on the parameter t) that ensures (C4*). This kind of condition was discussed in [7, 13]; in [13], in the framework of penalty, exact penalty and Lagrange multiplier methods, it was introduced as *Enlarged Mangasarian Fromovitz constraint qualification* (briefly, EnMFCQ). We define the following modification of EnMFCQ, where $B_q = \{x \in \mathbb{R}^n \mid \|x\|^2 \leq q\}$ for $q > 0$.

Definition 3.1 The *Modified EnMFCQ* is said to hold at B_q if for each $x \in B_q$ there exists a vector $\xi \in \mathbb{R}$ such that:

$$\begin{aligned} h_i(x) + Dh_i(x)\xi &\leq 0, \quad i \in \{i \in I \mid h_i(x) > 0\}, \\ g_j(x) + Dg_j(x)\xi &\leq 0, \quad j \in \{j \in J \mid g_j(x) > 0\}, \\ 2x^T\xi &< 0 \text{ if } \|x\|^2 = q, \\ -\sum_{i \in I} h_i(x) - \left(\sum_{i \in I} Dh_i(x) \right) \xi &\leq 0 \text{ if } \sum_{i \in I} h_i(x) < 0. \end{aligned}$$

Corollary 3.2 *Assume that the Modified EnMFCQ holds at B_q . Then, the condition (C4*) is fulfilled.*

Proof. First, let $t = 0$, $\bar{x} \in M^m(0)$ and $\|\bar{x}\|^2 = q$. Then, by the Modified EnMFCQ, we have $2\bar{x}^T\xi < 0$ and we are done. Now, let $\bar{t} \in (0, 1)$, $\bar{x} \in M^m(\bar{t})$ and consider any active constraint of (\bar{x}, \bar{t}) , e.g.

$$\bar{t}h_i(\bar{x}) + \bar{t} - 1 = 0 \text{ for some } i \in I.$$

Then, $\bar{t}h_i(\bar{x}) = 1 - \bar{t} > 0$ and, by the Modified EnMFCQ, we have $\bar{t}h_i(\bar{x}) + \bar{t}Dh_i(\bar{x})\xi \leq 0$ and, hence, $\bar{t}Dh_i(\bar{x})\xi < 0$. Since the active constraint was chosen arbitrarily, the proof is complete. \triangle

Next we present a “justification theorem” for the assumption (C2*) in Corollary 3.1

which is an application of Theorem 2.2. In particular, the perturbations used in the following Theorem 3.1 are related to the function vector $(f, h_i(x), g_j(x), i \in I, j \in J)$ which describes the original problem (P) (and which does not depend on t). Let the parameter vector $(b, A, c, D) \in \mathbb{R}^n \times \mathbb{R}^{n(n+1)/2} \times \mathbb{R}^{m+s+1} \times \mathbb{R}^{(m+s+1)n}$ with $D = (d^1, \dots, d^{m+s+1})$, $d^i \in \mathbb{R}^n$, $i = 1 \dots, m + s + 1$ be given and “almost all” be defined in an analogous way as in Theorem 2.2.

Theorem 3.1 For almost all (b, A, c, D) the modified standard embedding with the perturbed function vector

$$\begin{aligned} & tf(x) + (1-t)\|x - x^0\|^2 + (1-t)(b^\top x + x^\top Ax) \\ & th_i(x) + t - 1 + (1-t)(c_i + (d^i)^\top x), \quad i \in I \\ & tg_j(x) + t - 1 + (1-t)(c_{m+j} + (d^{m+j})^\top x), \quad j \in J \\ & \|x\|^2 - q \\ & -t \sum_{i \in I} h_i(x) + t - 1 + (1-t)(c_{m+s+1} + (d^{m+s+1})^\top x) \end{aligned}$$

is JJT-regular with respect to $(-\infty, 1)$.

The proof is left to the reader because it applies straightforwardly the technique used in the proof of Theorem 2.2 (cf. [22]).

In the subsequent section we will describe a pathfollowing method with jumps where the problem under consideration is assumed to be JJT-regular. However, the strong condition (C4*) need not be fulfilled and we will see that the failure of (C4*) may imply the method to stop at some parameter value $t^* < 1$, i.e. a gc point of the original problem (P) is not attained. In that sense the method to be presented is a heuristic one.

4 The heuristic method JUMP II*

Throughout this section we refer to the one-parametric problem $P(t)$ given in (1.2). We present the heuristic method JUMP II* in a general form for the class of problems $P(t)$. Then, in Section 5, we will apply numerically JUMP II* to problems of the form $P^m(t)$ (cf. (3.2)) in order to obtain a gc point (or a stationary point or a local minimizer) of a given problem (P) .

In [12], a pathfollowing method for the set Σ_{gc} (called PATH III, cf. [12, Section 4.5]) is described which can be combined with so-called *jumps* (the corresponding methods are called JUMP I cf. [12, Section 5.2] and JUMP II cf. [12, Section 5.3]). A *jump* refers to an appropriate feasible descent direction method (or, more general, NLP-solver) applied to a starting point $(x^1, t^1) \in \Sigma_{gc}$ with the objective to obtain a point $(x^2, t^1) \in \Sigma_{gc}$ that belongs to another connected component of Σ_{gc} than (x^1, t^1) (“jump from one connected component of Σ_{gc} to another one”). Next we explain very briefly the main ideas of PATH III and JUMP II; for more details we refer to [12].

PATH III

The method computes, for a given interval $[t_A, t_B] \subset \mathbb{R}$, $t_A < 0 < t_B$, a numerical description of a compact connected component of $\Sigma_{gc} \cap (\mathbb{R}^n \times [t_A, t_B])$ assuming that $P(t)$ is JJT-regular with respect to $[t_A, t_B]$. In particular, PATH III finds a discretization of $[t_A, t_B]$ and corresponding gc points starting at $(x^0, 0) \in \Sigma_{gc}$ (cf. (A1) in Section 1). The method is based on an active-index-set strategy and uses a so-called predictor-corrector scheme for those parts with constant active index sets. A Newton-like corrector is applied which implies a superlinear rate of convergence. We mention two important features of this approach: the computation of the new index sets for all possible continuations at points of Types 2 and 5 (cf. Figure 2.1) as well as the pathfollowing of the turning points of Types 3 and 4 (cf. Figure 2.1).

Remark 4.1 If $[0, 1] \subset [t_A, t_B]$ and if there exists a PC³-path in Σ_{stat} connecting $(x^0, 0)$ and $(x^*, 1)$ (cf. Theorem 2.3), then PATH III also refers to the standard procedure of the embedding approach: A finite discretization

$$0 = t_0 < t_1 < \dots < t_i < \dots < t_N = 1$$

of the interval $[0, 1]$ and corresponding approximations $\tilde{x}(t_i)$ of stationary points $x(t_i)$ of $P(t_i)$ are obtained.

JUMP II

The goal of JUMP II is to find a numerical description of finitely many connected components of $\Sigma_{gc} \cap (\mathbb{R}^n \times [t_A, t_B])$ by combining PATH III (for the numerical description of each connected component) with jumps in Σ_{gc} . The jumps (feasible descent direction method, NLP-solver) are defined for gc points of Type 1 and for gc points from $cl\Sigma_{loc}$ (where cl denotes the closure) which are of Type 2, 3 or 4. For more details on JUMP II we refer to the Examples 5.3.1, 5.3.2 and 5.3.3 as well as to the Figures 5.17, 5.18 and 5.19 in [12, Section 5.3].

Now we introduce the heuristic method JUMP II* as a generalization of JUMP II. We refer to the forthcoming Remark 4.4 after the presentation of JUMP II* where some differences between JUMP II and JUMP II* are discussed.

*JUMP II**

As for JUMP II, the goal of JUMP II* is to find a numerical description of finitely many connected components of $\Sigma_{gc} \cap (\mathbb{R}^n \times [t_A, t_B])$. However, in JUMP II* there exist more possibilities for “jumping from one connected component of $\Sigma_{gc} \cap (\mathbb{R}^n \times [t_A, t_B])$ to another one” than in JUMP II.

Let $[t_A, t_B] \subset \mathbb{R}$ with $t_A < 0 < t_B$ and assume the following conditions:

- (B1) $P(t)$ is JJT-regular with respect to $[t_A, t_B]$.

(B2) $(x^0, 0) \in \Sigma_{gc}^1$ is known.

(B3) $M(t)$ is nonempty and there exists a compact set containing $M(t)$ for all $t \in [t_A, t_B]$.

By using PATH III with the starting point $(x^0, 0) \in \Sigma_{gc}^1$ (cf. (B2)) assume that a connected component C of $\Sigma_{gc} \cap (\mathbb{R}^n \times [t_A, t_B])$ with $(x^0, 0) \in C$ is described numerically: we have a (sufficiently fine) discretization of the interval $[t_A, t_B]$:

$$t_A \leq t_{-q_1} < \dots < t_0 = 0 < \dots < t_i < t_{i+1} < \dots < t_{r_1} \leq t_B$$

as well as the corresponding sets $C(t_i) \subset \Sigma_{gc}$ where $C(t_i) = \{(x, t) \in C \mid t = t_i\}$, $i \in S = \{-q_1, \dots, 0, \dots, r_1\}$ with

$$C \cap \left(\bigcup_{v=2}^5 \Sigma_{gc}^v \right) \subset \bigcup_{i \in S} C(t_i). \quad (4.1)$$

The latter situation is depicted in [12, Figure 5.16]. By (B3) and since the gc points of Types 2-5 form a discrete set, the sets $C \cap \Sigma_{gc}^v$, $v = 2, \dots, 5$ are finite and, hence, the condition (4.1) can be fulfilled by a suitable choice of the step length in PATH III. We calculate for each t_i , $i \in S$ the points $(\hat{x}(t_i), t_i) \in C$ and $(\bar{x}(t_i), t_i) \in C$ satisfying

$$f(\hat{x}(t_i), t_i) \leq f(x, t_i) \text{ for all } (x, t_i) \in C(t_i) \quad (4.2)$$

and

$$f(\bar{x}(t_i), t_i) \geq f(x, t_i) \text{ for all } (x, t_i) \in C(t_i). \quad (4.3)$$

Note that $\hat{x}(t_i)$ and $\bar{x}(t_i)$ need not be unique.

JUMP II* deals with the following 6 situations where for each of them a corresponding feasible descent direction exists: by applying a feasible descent direction method one obtains a gc point which does not belong to C (“jump from C to another connected component of $\Sigma_{gc} \cap (\mathbb{R}^n \times [t_A, t_B])$ ”) which will be described numerically by applying PATH III again). The Situations 1, 2 and 5 are cited from [12, Section 5.3] since they also appear in JUMP II. The proof and theoretical background for the descent directions given in Situations 3, 4 and 6 can be found in [14].

Situation 1: There exists an $i_0 \in S$ with $(\hat{x}(t_{i_0}), t_{i_0}) \in \Sigma_{gc}^1 \setminus \Sigma_{stat}$.

Then we obtain a feasible first order descent direction $\xi \in \mathbb{R}^n$ by solving the system

$$\begin{aligned} D_x f(\hat{x}(t_{i_0}), t_{i_0}) \xi &< 0, & D_x h_i(\hat{x}(t_{i_0}), t_{i_0}) \xi &= 0, & i \in I \\ D_x g_j(\hat{x}(t_{i_0}), t_{i_0}) \xi &\leq 0, & j \in J_0(\hat{x}(t_{i_0}), t_{i_0}). \end{aligned}$$

Situation 2: There exists an $i_0 \in S$ with $(\hat{x}(t_{i_0}), t_{i_0}) \in \Sigma_{stat}^1 \setminus \Sigma_{loc}$.

Then we obtain a feasible second order descent direction $\xi \in \mathbb{R}^n$ by solving the problem (where L is the Lagrangian, cf. (2.2c)):

$$\min \left\{ \xi^T D_x^2 L(\hat{x}(t_{i_0}), t_{i_0}) \xi \mid \begin{array}{l} \|\xi\| = 1, D_x h_i(\hat{x}(t_{i_0}), t_{i_0}) \xi = 0, \quad i \in I, \\ D_x g_j(\hat{x}(t_{i_0}), t_{i_0}) \xi \leq 0, \quad j \in J_0(\hat{x}(t_{i_0}), t_{i_0}) \end{array} \right\}$$

Situation 3: There exists an $i_0 \in S$ with $(\hat{x}(t_{i_0}), t_{i_0}) \in \Sigma_{gc}^2 \cap \Sigma_{stat}$ and $(\hat{x}(t_{i_0}), t_{i_0})$ is a turning point in Σ_{stat} (cf. Figure 2.1 (Type 2, (b) and (d)) and Figure 4.1).

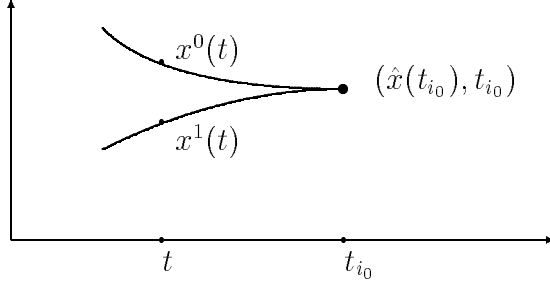


Figure 4.1: $x^0(t)$ and $x^1(t)$ describe locally two branches of Σ_{stat}

Then, $u := \lim_{t \rightarrow t_{i_0}} \frac{x^0(t) - x^1(t)}{\|x^0(t) - x^1(t)\|}$ exists and u or $-u$ is a feasible second order descent direction.

Situation 4: There exists an $i_0 \in S$ with $(\hat{x}(t_{i_0}), t_{i_0}) \in \Sigma_{gc}^3$. Then,

$$u := \lim_{t \rightarrow t_{i_0}} \frac{x^0(t) - x^1(t)}{\|x^0(t) - x^1(t)\|}$$

exists and u or $-u$ is a feasible third order descent direction, where $x^0(t)$ and $x^1(t)$ again describe the two different branches of Σ_{gc} locally around $(\hat{x}(t_{i_0}), t_{i_0})$.

Remark 4.2 Let $\bar{z} = (\hat{x}(t_{i_0}), t_{i_0})$, $\bar{z} \in \Sigma_{gc}^4 \cap cl\Sigma_{loc}$ and $J_0(\bar{z}) = \{1, \dots, p\}$. Then there exists a vector $\bar{q} \in \mathbb{R}^{m+p}$ which is uniquely determined - up to a common multiple - and satisfies (cf. [16])

$$\sum_{i \in I} \bar{q}_i D_x h_i(\bar{z}) + \sum_{j=1}^p \bar{q}_{m+j} D_x g_j(\bar{z}) = 0, \quad \bar{q} \neq 0 \text{ with } \bar{q}_{m+j} > 0, \quad j = 1, \dots, p.$$

Define

$$\delta = \text{sign} \left(\sum_{i \in I} \bar{q}_i D_t h_i(\bar{z}) + \sum_{j=1}^p \bar{q}_{m+j} D_t g_j(\bar{z}) \right).$$

If $\delta = -1$ ($\delta = 1$), then there exists (there does not exist) a jump to another connected component of $\Sigma_{gc} \cap (\mathbb{R}^n \times [t_A, t_b])$ (cf. [12] for more details and see Figure 4.2 as well as Remarks 4.4 and 4.5).

Situation 5: There exists an $i_0 \in S$ with $(\hat{x}(t_{i_0}), t_{i_0}) \in \Sigma_{gc}^4 \cap cl\Sigma_{loc}$ and $\delta = -1$ (cf. Remark 4.2). Then there exists a point of Σ_{gc} beyond the turning point $(\hat{x}(t_{i_0}), t_{i_0})$, say $(x_{max}(t), t)$ (cf. Figure 4.2) such that $x_{max}(t)$ is a local maximizer of $P(t)$. We can start with a feasible descent method at $(x_{max}(t), t)$ in order to jump to another connected component of $\Sigma_{gc} \cap (\mathbb{R}^n \times [t_A, t_B])$.

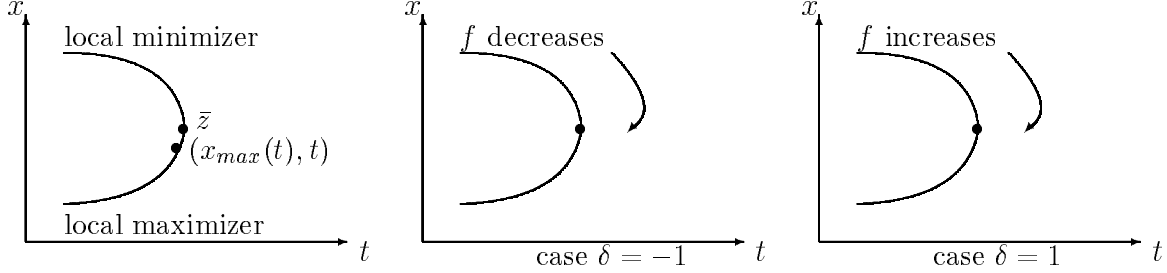


Figure 4.2: The behaviour of Σ_{gc} and f in a neighbourhood of $\bar{z} \in \Sigma_{gc}^4 \cap cl\Sigma_{loc}$

Situation 6: There exists an $i_0 \in S$ with $(\hat{x}(t_{i_0}), t_{i_0}) \in \Sigma_{gc}^4 \setminus cl\Sigma_{loc}$.

Then, analogously to the case $\delta = -1$ in Situation 5, there always exists a feasible descent direction in order to jump to another connected component of $\Sigma_{gc} \cap (\mathbb{R}^n \times [t_A, t_B])$ (the construction of this direction is very technical and is contained in [14]).

Remark 4.3 The Situations 1-6 can be modified as follows. One can substitute the point $(\hat{x}(t_{i_0}), t_{i_0})$ by the point $(\bar{x}(t_{i_0}), t_{i_0})$ (cf. (4.3)) and calculate a corresponding feasible descent direction with respect to the negative objective function $-f$.

Next we present an algorithmic version of JUMP II* which for sake of simplicity is restricted to the numerical description of at most seven connected components of $\Sigma_{gc} \cap (\mathbb{R}^n \times [t_A, t_B])$. Obviously, it can easily be generalized to the case of more than seven components.

*An algorithmic version of JUMP II**

1. Let $(x^0, t_0) \in \Sigma_{gc}^1$ (starting point). $k := 1$, $t_-^1 := t_A$, $t_+^1 := t_B$, $x_0^1 := x^0$, $t_0^1 := t_0$.
2. Starting point (x_0^k, t_0^k) . Apply algorithm PATH III and obtain a numerical description of the connected component C_k of $\Sigma_{gc} \cap (\mathbb{R}^n \times [t_-^k, t_+^k])$ which contains (x_0^k, t_0^k) .
3. If $k = 7$, then STOP.
4. Let $t_{-q_k}^k < \dots < t_0^k < \dots < t_i^k < t_{i+1}^k < \dots < t_{r_k}^k$ be the discretization of the interval $[t_-^k, t_+^k]$ obtained such that $C_k \cap \left(\bigcup_{v=2}^5 \Sigma_{gc}^v\right) \subset \bigcup_{i \in S^k} C_k(t_i^k)$, where

$$S^k = \{-q_k, \dots, 0, \dots, r_k\} \text{ and } C_k(t_i^k) = \{(x, t) \in C_k \mid t = t_i^k\}.$$

5. If $k \leq 3$, then determine $\hat{x}^k(t_i^k)$ and $\bar{x}^k(t_i^k)$ analogously to $\hat{x}(t_i)$ and $\bar{x}(t_i)$ in (4.2) and (4.3), respectively, $i \in S^k$.
6. $k := k + 1$.
7. If $k = 2$, then $t_-^2 := t_A$, $t_+^2 := t_B$, $\hat{x}(t_i) := \hat{x}^1(t_i^1)$, $i \in S^1$, go to 16.
8. If $k = 3$, then $t_-^3 := t_A$, $t_+^3 := t_B$, $\hat{x}(t_i) := \bar{x}^1(t_i^1)$, $i \in S^1$, go to 16.
9. If $k > 4$, then go to 13.
10. If $t_{r_2}^2 \leq \max(t_{r_1}^1, t_{r_3}^3)$ or $t_{-q_3}^3 \geq \min(t_{-q_1}^1, t_{-q_2}^2)$, then STOP.
11. $t_+^* := \max(t_{r_1}^1, t_{r_3}^3)$, $t_-^* := \min(t_{-q_1}^1, t_{-q_2}^2)$.
12. $t_-^4 := t_A$, $t_+^4 := t_B$, $\hat{x}(t_i) := \hat{x}^2(t_i^2)$, $i \in S^2$, go to 16.
13. If $k = 5$, then $t_-^5 := t_+^*$, $t_+^5 := t_B$, $\hat{x}(t_i) = \bar{x}^2(t_i^2)$, $i \in \{i \in S^2 \mid t_i^2 \geq t_+^*\}$, go to 16.
14. If $k = 6$, then $t_-^6 := t_A$, $t_+^6 := t_-^*$, $\hat{x}(t_i) := \hat{x}^3(t_i^3)$, $i \in \{i \in S^3 \mid t_i^3 \leq t_-^*\}$, go to 16.
15. If $k = 7$, then $t_-^7 := t_A$, $t_+^7 := t_B$, $\hat{x}(t_i) := \bar{x}^3(t_i^3)$, $i \in S^3$.
16. Determine the first index i such that one of the following conditions holds (choose the first condition that is satisfied):
 - (E1) $(\hat{x}(t_i), t_i) \in \Sigma_{gc}^1 \setminus \Sigma_{stat}$
 - (E2) $(\hat{x}(t_i), t_i) \in \Sigma_{stat}^1 \setminus \Sigma_{loc}$
 - (E3) $(\hat{x}(t_i), t_i) \in \Sigma_{gc}^2 \cap \Sigma_{stat}$ and $(\hat{x}(t_i), t_i)$ is a turning point in Σ_{stat}
 - (E4) $(\hat{x}(t_i), t_i) \in \Sigma_{gc}^3$
 - (E5) $(\hat{x}(t_i), t_i) \in \Sigma_{gc}^4 \cap cl\Sigma_{loc}$ and $(\hat{x}(t_i), t_i)$ satisfies $\delta = -1$ (cf. Remark 4.2)
 - (E6) $(\hat{x}(t_i), t_i) \in \Sigma_{gc}^4 \setminus cl\Sigma_{loc}$
17. If (E1) ((E2), (E3), (E4), (E5), (E6), respectively) holds, then calculate a feasible descent direction of f if $k \in \{2, 4, 6\}$ and of $-f$ if $k \in \{3, 5, 7\}$ (cf. Remark 4.3) as described in Situation 1 (2, 3, 4, 5, 6, respectively). Apply an appropriate NLP-solver (“jump”) and calculate a new point (\tilde{x}, t_i) which belongs to another connected component. $x_0^k := \tilde{x}$, $t_0^k := t_i$, go to 2.

Remark 4.4 (Difference between JUMP II and JUMP II*) As already mentioned above, the goal of JUMP II and JUMP II* is to find (as many as possible) connected components of $\Sigma_{gc} \cap (\mathbb{R}^n \times [t_A, t_B])$ and to obtain a numerical description of them. The method JUMP II* is an extension of JUMP II since JUMP II* additionally includes possible jumps from gc points of Types 2, 3, 4 which do not belong to $cl\Sigma_{loc}$ (cf. Situations 3-6 and [14] for the theoretical background). Therefore, by using JUMP II* there is a chance to detect more connected components of $\Sigma_{gc} \cap (\mathbb{R}^n \times [t_A, t_B])$ and,

in particular, to find a solution (gc point or stationary point or local minimizer) for a given parameter value \bar{t} , e.g. for $\bar{t} = 1$ (as it is the goal in the framework of the embedding approach). In the following we give a résumé for possible jumps in JUMP II and JUMP II*.

	Jumps in JUMP II	Jumps in JUMP II*
$(\bar{x}, \bar{t}) \in \Sigma_{gc}^1$	for $(\bar{x}, \bar{t}) \in \Sigma_{gc}^1 \setminus \Sigma_{loc}$	for $(\bar{x}, \bar{t}) \in \Sigma_{gc}^1 \setminus \Sigma_{loc}$
$(\bar{x}, \bar{t}) \in \Sigma_{gc}^2$	for $(\bar{x}, \bar{t}) \in cl\Sigma_{loc}$ when (\bar{x}, \bar{t}) is a turning point in Σ_{stat}	for $(\bar{x}, \bar{t}) \in \Sigma_{stat}$ when (\bar{x}, \bar{t}) is a turning point in Σ_{stat}
$(\bar{x}, \bar{t}) \in \Sigma_{gc}^3$	for $(\bar{x}, \bar{t}) \in cl\Sigma_{loc}$	for $(\bar{x}, \bar{t}) \in \Sigma_{gc}^3$
$(\bar{x}, \bar{t}) \in \Sigma_{gc}^4$	for $(\bar{x}, \bar{t}) \in cl\Sigma_{loc}$ with $\delta = -1$ (cf. Remark 4.2)	for $(\bar{x}, \bar{t}) \in \Sigma_{gc}^4$ and with $\delta = -1$ when $(\bar{x}, \bar{t}) \in cl\Sigma_{loc}$

Remark 4.5 (JUMP II* is a heuristic method) By Remark 4.4, there exist the following 5 cases where JUMP II* does not provide a jump at $(\bar{x}, \bar{t}) \in \Sigma_{gc}$:

- Case I: $(\bar{x}, \bar{t}) \in \Sigma_{gc}^1 \cap \Sigma_{loc}$
- Case II: $(\bar{x}, \bar{t}) \in \Sigma_{gc}^2$ and (\bar{x}, \bar{t}) is not a turning point in Σ_{stat}
- Case III: $(\bar{x}, \bar{t}) \in \Sigma_{gc}^4 \cap cl\Sigma_{loc}$ and $\delta = 1$ (cf. Remark 4.2)
- Case IV: $(\bar{x}, \bar{t}) \in \Sigma_{gc}^5$ and (MFCQ) does not hold at (\bar{x}, \bar{t})
- Case V: $(\bar{x}, \bar{t}) \in \Sigma_{gc}^5$ and (MFCQ) holds at (\bar{x}, \bar{t})

We can see from the topological properties of Σ_{gc} (for more details, see [16]) and from Figure 2.1 that for the Cases I (cf. Figure 2.1), II (cf. Figure 2.1 (a), (c)) and V (cf. Figure 2.1 (k)) there is no need for a jump at (\bar{x}, \bar{t}) since locally there exist continuations within Σ_{gc} for increasing and decreasing values of the parameter t . On the other hand, we have a different situation for the Cases III and IV; here, *the connected component of the feasible set $M(t)$ shrinks locally to the point $\{(\bar{x}, \bar{t})\}$ and becomes empty beyond \bar{t} (for $t > \bar{t}$ in Figure 2.1 (g), (i), (l) and for $t < \bar{t}$ in Figure 2.1 (k), (j), (m))*. Obviously, in these situations there does not exist a feasible (descent) direction and one cannot jump to another connected component. According to Theorem 2.1, the JTT-regularity is a generic property and, hence, the Cases III and IV cannot be excluded by appropriate arbitrarily small perturbations. In this sense, JUMP II* is (only) a heuristic method since for the Cases III and IV we cannot continue the path-following procedure (in both directions with respect to the parameter t). Note that for the Cases III and IV the (MFCQ) does not hold at (\bar{x}, \bar{t}) . Therefore, the condition (C4) in Theorem 2.3 (and (C4*) in Corollary 3.1 as well) are very strong conditions since they exclude the Cases III and IV.

Remark 4.6 The heuristic method JUMP II* (as well as PATH III and JUMP II) is implemented as a part of the program package PAFO [9] which was developed at the Institute of Mathematics at the Humboldt University, Berlin. According to JUMP II*, PAFO combines pathfollowing methods with standard NLP-solvers (jumps) in order to describe numerically *more than one* connected component of the solution set of a one-parametric optimization problem. Note that, in general, a standard pathfollowing method describes *only one* connected component. In the subsequent section we present several numerical illustrations for the heuristic method JUMP II* obtained by the application of the program package PAFO.

5 Computational results

In this section we present three numerical examples which illustrate the application of JUMP II* to the modified standard embedding $P^m(t)$. All examples contain an equality constraint. As already mentioned in Remark 4.6, we used the program package PAFO [9] for the realization of these computational tests.

Before showing the numerical results we emphasize again the advantages of our approach:

- The modified standard embedding only contains *inequality* constraints (although the original problem (P) may contain *equality* and inequality constraints) which is important for the numerical application of feasible descent direction methods. Furthermore, we have that $M^m(t) \neq \emptyset$, $t \in [0, 1]$.
- The use of jumps from one connected component to another one implies that more than one connected component of the solution set can be followed and described numerically.
- The new method proposed in the current paper allows a larger set of starting points for the jumps. The known methods only work with starting points which are endpoints of branches of local minimizers. By using JUMP II*, different other types of bifurcation and turning points of the solution set can be chosen as starting points.

Example 5.1 We consider an example from [6] for which the standard embedding (cf. Example 1.1) does not work since the condition (3.1) fails to hold. We consider in \mathbb{R}^2 the problem

$$(P) \quad \min \left\{ x_1 + x_2 \mid \frac{1}{4}x_1^4 + \frac{2}{3}x_1^3 - \frac{2}{3}x_1^2 = -4, \left(x_1 + \frac{3}{2}\right)^2 \leq \frac{9}{4}, -x_1 - x_2 \leq 5 \right\}$$

as well as the starting point

$$x^0 = \begin{pmatrix} 2.40858 \\ 0 \end{pmatrix} \text{ and } q = 1000$$

for the corresponding modified standard embedding $P^m(t)$. By applying JUMP II*, we obtain the solution

$$x^* = \begin{pmatrix} -1.43522 \\ 6.43516 \end{pmatrix}.$$

Figures 5.1 and 5.2 illustrate the solution curves of the components x_1 and x_2 , respectively, where the stationary points are presented by a full line and other gc points by a dotted one. Note that there does not exist a curve in Σ_{stat} connecting $(x^0, 1)$ and $(x^*, 1)$.

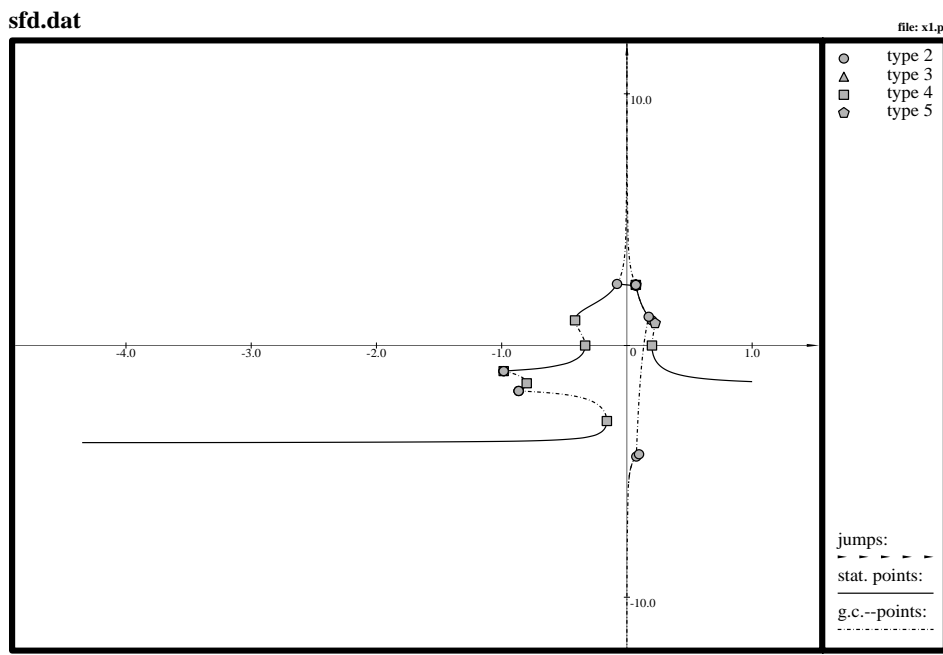


Figure 5.1

Example 5.2 Consider the problem

$$(P) \quad \min\{x_1 + x_2 \mid x_1^4 + \frac{1}{2}x_2^2 = 16, \sin(5x_1) \leq 0\}.$$

Now, our goal is to obtain as many solutions of (P) as possible. JUMP II* provides the following two jumps:

$$t = 0.92490 : \begin{pmatrix} 1.27289 \\ -5.15627 \end{pmatrix} \rightarrow \begin{pmatrix} 0.61206 \\ -5.63663 \end{pmatrix}$$

and

$$t = 0.88168 : \begin{pmatrix} 1.28356 \\ -5.18071 \end{pmatrix} \rightarrow \begin{pmatrix} 0.60140 \\ -5.62923 \end{pmatrix}.$$

At $t = 1$ we obtain the following 4 solutions of (P) :

x_1	x_2
1.88496	-2.59835
1.25663	-5.19737
0.62833	-5.62923
-0.00001	-5.65685

The Figures 5.3 and 5.4 are organized in an analogous manner as the Figures 5.1 and 5.2; they illustrate the solution curves corresponding to Example 5.2.

Recall that the construction of the modified standard embedding yields that $P^m(t)$ is not JJT-regular with respect to an interval which contains the parameter value $t = 1$ (cf. Section 3). However, although $P^m(t)$ is only JJT-regular with respect to $[a, 0)$, $a \in \mathbb{R}$, the numerical experience shows that JUMP II* attains $t = 1$.

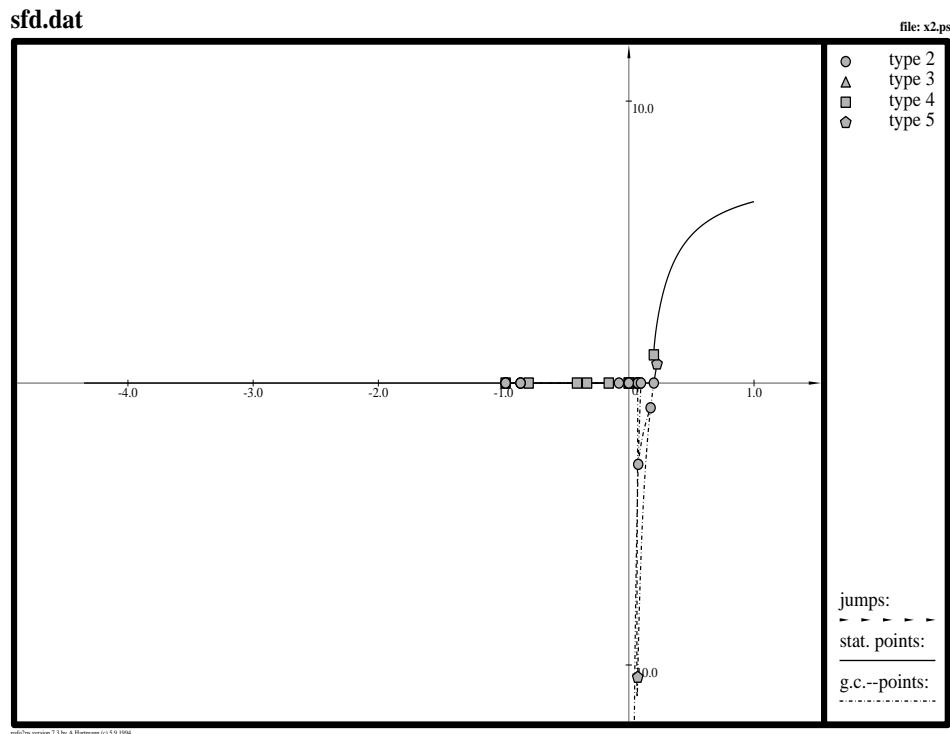


Figure 5.2

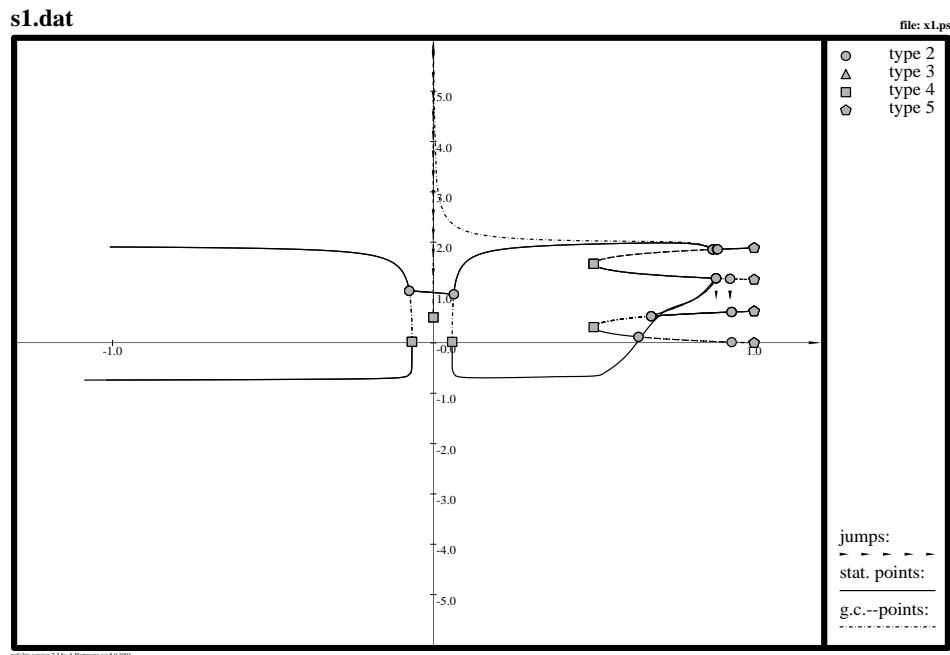


Figure 5.3

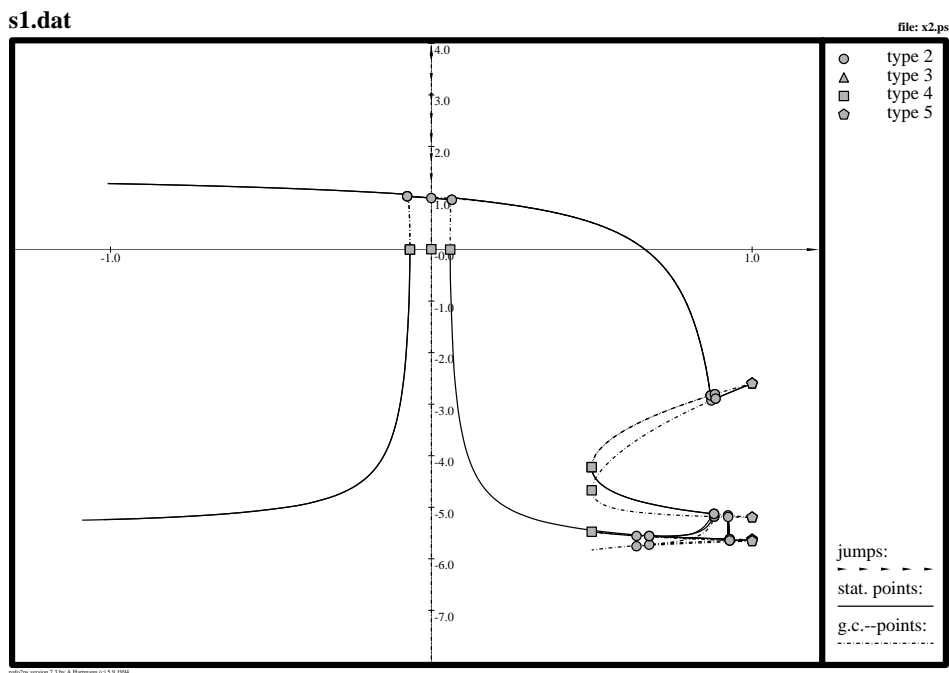


Figure 5.4

Example 5.3 Let (P) be defined by $n = 3$,

$$f(x) = x_1x_2 + x_3^3,$$

$$h_1(x) = x_1x_2x_3 - x_3^5 - x_2x_3^2 - x_1^2x_3 + 88,$$

$$g_1(x) = 0.02650736x_1^8 - 0.211505207x_1^7 + 0.26753848x_1^6 + 1.34579642x_1^5 \\ - 2.34222067x_1^4 - 2.65020635x_1^3 + 3.45664738x_1^2 + 0.01447716x_1 + 2,$$

$$g_2(x) = x_2^2 - 10 \cos(2\pi x_2) + 5,$$

and choose the starting point $x^0 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ as well as $q = 25$.

As in Example 5.2 our goal is to obtain as many solutions of (P) as possible and therefore we follow all curves of gc points. We get the following solution points of (P) :

x_1	x_2	x_3
1.59999	0.84686	2.40352
2.80002	-0.16616	2.33565
1.60000	-0.16616	2.41499
3.87835	-0.16616	2.22054
1.60000	1.14147	2.40022
1.60000	1.91715	2.39163
1.60000	0.16616	2.41120
3.87286	0.16616	2.23010
3.35655	-0.16616	2.28199
3.87835	-0.16616	2.22054
3.35656	0.16616	2.28751

The Figures 5.5, 5.6 and 5.7 present the solution curves of the components x_1 , x_2 and x_3 , respectively.

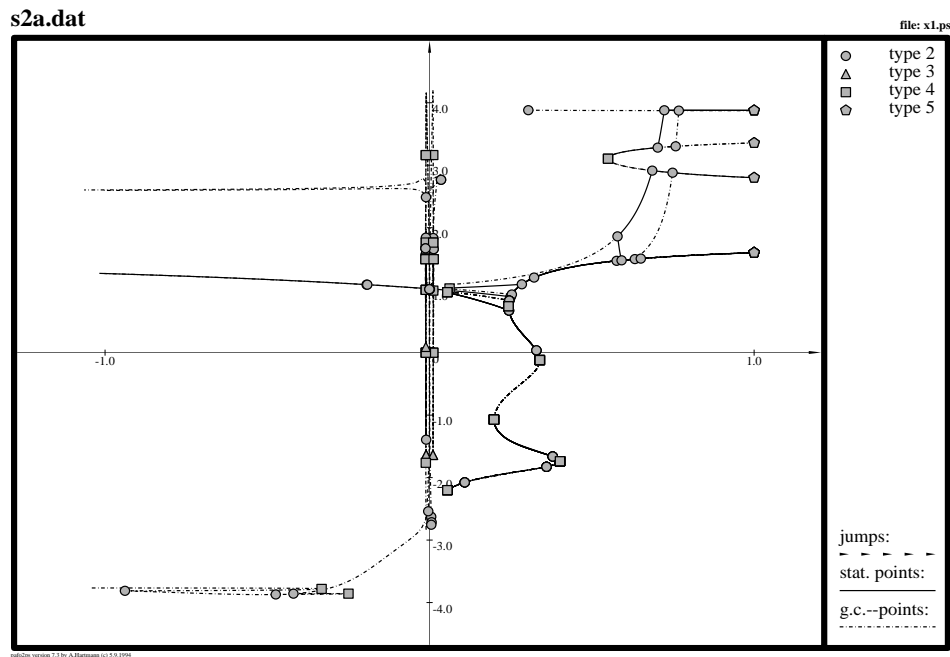


Figure 5.5

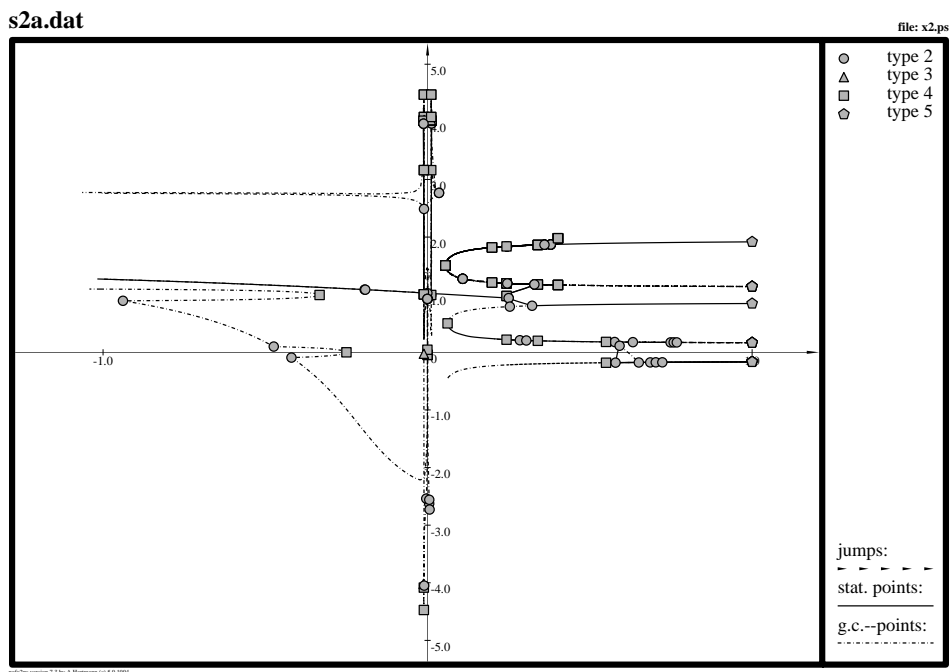


Figure 5.6

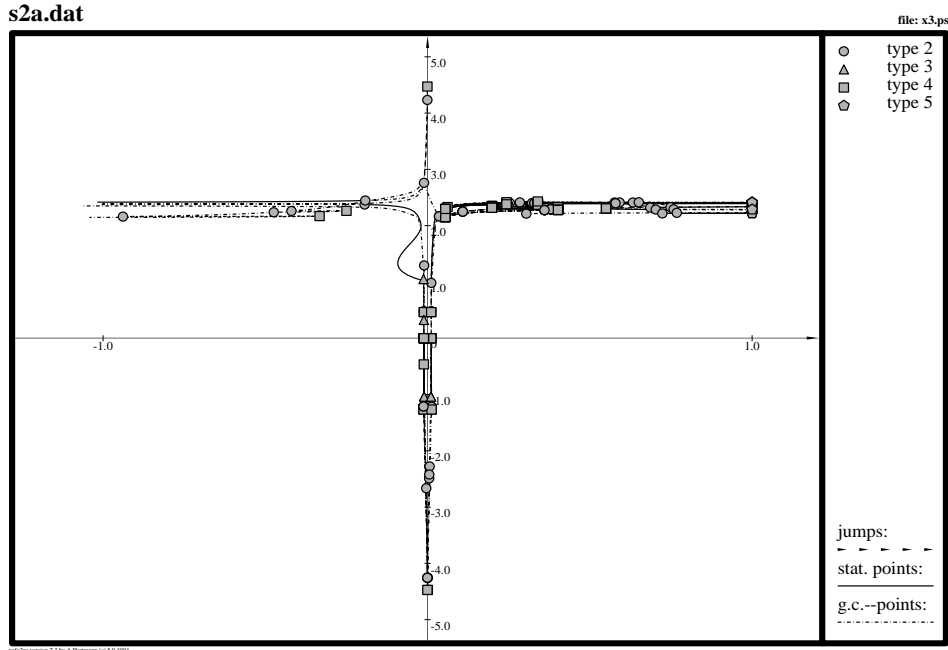


Figure 5.7

References

- [1] Allgower, E., Georg, K.: Introduction to Numerical Continuation Methods, Springer, Berlin, Heidelberg, New York, 1990.
- [2] Billups, S. C.: A homotopy-based algorithm for mixed complementarity problems, SIAM J. Optimization 12, 583-605, 2002.
- [3] Billups, S. C., Watson, L. T.: A probability-one homotopy algorithm for nonsmooth equations and mixed complementarity problems, SIAM J. Optimization 12, 606-626, 2002.
- [4] Dentcheva, D., Gollmer, R., Guddat, J., Rückmann, J.-J.: Pathfollowing methods in nonlinear optimization II: Exact penalty embedding, in: Florenzano, M. et al. (eds.), Approximation and Optimization in the Caribbean, Peter Lang, Frankfurt, 200-230, 1995.
- [5] Dentcheva, D., Guddat, J., Rückmann, J.-J., Wendler, K.: Pathfollowing methods in nonlinear optimization III: Lagrange multiplier embedding, Z. Oper. Res. 41, 127-152, 1995.
- [6] Fandom, R.: On modifications of the standard embedding in nonlinear optimization, in Guddat, J. et al. (eds.): Parametric Optimization and Related Topics IV, Peter Lang, Frankfurt, 59-84, 1997.

- [7] Gfrerer, H., Guddat, J., Wacker, H.J., Zulehner, W.: Path-following methods for Kuhn-Tucker curves by an active index set strategy, in: Baghi, A. and Jongen, H.Th. (eds.), *Systems and Optimization*, Springer, Berlin, Heidelberg, New York, 111-132, 1985.
- [8] Gollmer, R., Guddat, J., Guerra, F., Nowack, D., Rückmann, J.-J.: Pathfollowing methods in nonlinear optimization I: Penalty embedding, in: Guddat, J. et al. (eds.), *Parametric Optimization and Related Topics*, Peter Lang, Frankfurt, 163-214, 1993.
- [9] Gollmer, R., Kausmann, U., Nowack, D., Wendler, K., Bacallao Estrada, J.: Program package PAFO, Humboldt-Universität Berlin, Institut für Mathematik, 1995-2004.
- [10] Gómez Bofill, W., Guddat, J., Jongen, H.Th., Rückmann, J.-J., Solano, C.: *Curvas critica y saltos en optimizacion no lineal*, electronically published: <http://www.emis.de/monographs/curvas/index.html>, 2000.
- [11] Gómez Bofill, W.: On generic quadratic penalty embeddings for nonlinear optimization problems, *Optimization* 50, 279-295, 2001.
- [12] Guddat, J., Guerra, F., Jongen, H. Th.: *Parametric Optimization: Singularities, Pathfollowing and Jumps*, Teubner, Stuttgart and John Wiley, Chichester, 1990.
- [13] Guddat J., Guerra F., Nowack D.: On the role of the Mangasarian- Fromovitz constraint qualification for penalty, exact penalty and Lagrange multiplier methods, in: A. V. Fiacco (ed.), *Mathematical Programming with Data Perturbations*, Marcel Dekker, New York, 159-185, 1997.
- [14] Guddat, J., Rückmann, J.-J.: One-parametric optimization: jumps in the set of generalized critical points, *Control and Cybernetics* 23, 1/2, 139-151, 1994.
- [15] Hirsch, M. W.: *Differential Topology*, Springer, Berlin, Heidelberg, New York, 1976.
- [16] Jongen, H. Th., Jonker, P., Twilt, F.: Critical sets in parametric optimization, *Math. Programming* 34, 333-353, 1986.
- [17] Jongen, H. Th., Jonker, P., Twilt, F.: *Nonlinear Optimization in Finite Dimension*, Kluwer Academic Publishers, Boston, 2000.
- [18] Kojima, M., Hirabayashi, R.: Continuous deformation of nonlinear programs, *Math. Progr. Study* 21, 150-198, 1984.
- [19] Lundberg, B. N., Poore, A. B., Yang, B.: Smooth penalty functions and continuation methods for constrained optimization, in: Allgower, E. L., Georg, K. (eds.), *Lectures in Applied Mathematics Series 26*, AMS, Providence, 389-412, 1990.

- [20] Mangasarian, O. L., Fromovitz, S.: The Fritz John necessary optimality conditions in the presence of equality and inequality constraints, *J. Math. Anal. Appl.* 17, 37-47, 1967.
- [21] Rheinboldt, W. C.: *Numerical Analysis of Parametric Nonlinear Equations*, John Wiley, New York, 1986.
- [22] Rückmann, J.-J., Tammer, K.: On linear-quadratic perturbations in one-parametric non-linear optimization, *Systems Science* 18, 37-48, 1992.
- [23] Wacker, HJ. (ed.): *Continuation Methods*, Academic Press, New York, 1978.