On generic quadratic penalty embeddings for nonlinear optimization problems.

Gómez Bofill, Walter

12 September 1997

Abstract

The generic character of the regularity in the sense of Jongen, Jonker and Twilt is studied for a particular class of embeddings, which represents a quadratric penalty procedure. In this paper we state a suitable perturbation result (into the mentioned class), which is the main part for the proof of the genericity. Finally, the results of some numerical experience are mentioned and two selected examples are presented.

1 Introduction.

The use of smooth penalty functions to solve constrained optimization problems in a sequential unconstrained optimization setting has been reported as a numerically inefficient approach for a long time. The principal difficulty appearing is the inevitable ill-conditioning when the penalty parameter is close to the prescribed term (see for example [4, 5, 6]).

Recently, these sequential optimization approaches have been re-examined from a pathfollowing point of view in many papers. The idea is to reformulate the optimality condition of the unconstrained optimization problems in a parametric system of nonlinear equations. The pathfollowing algorithms stated on these systems of parametric equations can give us a new idea of the capacities of the sequential approach.

In the paper [20] two smooth penalty functions (quadratic penalty for equality constraints, and logarithmic barrier functions or quadratic penalty loss for the inequality constraints) are analyzed from the pathfollowing point of view. Two different systems of equations are used for the pathfollowing procedure. As a consequence of reformulating these sequential algorithms as a parametric system of equations the ill-conditioning can be removed. Further papers studying the structure of the solution of the expanded Lagrangian system in the neighbourhood of singularities of different nature ([21],[22]) have been published recently. This analysis follows the ideas of the investigation of general one-parametric optimization problems, and the classification of the possible singularities using tools from bifurcation analysis. In general, this classification results from the violation of the conditions defining the nondegenerated critical points. The study of the stucture of the critical set in the neigbourhood of singularities and the characterization of the structures in terms of the problem data is done for singularities of codimension zero or one ([23],[24]).

For general one parametric optimization problems, two recent, important approaches to the analysis of the global structure of critical points are available. These approaches examine restricted classes of one-parametric optimization problems (we call them regular one-parametric problems) for which the structure of some critical sets is relatively simple to describe. One approach based on piecewise differentiable mappings is due to Kojima and Hirabayashi ([18]) and deals with the global structure of the Karush-Kuhn-Tucker set. The other approach, based on transversal intersection, is due to Jongen, Jonker and Twilt ([14], [15]). In this latter approach the generalized critical points are classified into five types, which are then analyzed in detail. This approach is of fundamental importance, since it treats the generic behaviour of one parametric optimization problems.
With the tools developed by Jongen, Jonker and Twilt different sequential methods for constrained optimization ([2], [3],[7],[9]) have been studied. Results obtained by using pathfollowing procedures with jumps ([10], [8]) for different penalty, exact penalty and multiplier embeddings with the possibilities of their "pure" sequential versions were compared. Convergence analyses were stated under usual assumptions on the sequential approach and under the regularity of the embeddings obtained. This regularity of the stated embeddings in the sense of Jongen, Jonker and Twilt is the key assumption in order to profit from the local description of a finite number of possible singularities. The genericity of the class of regular problems introduced by Jongen, Jonker and Twilt (JJT-regularity) is obtained in a full C^3 differen-

tiability setting. As a consequence this assumption can lose its genericity when dealing with particular classes of one-parametric problems.

In this direction new classifications of singularities for parametric problems with special structure have been developed. As examples let us mention the papers [12], [26] and [19].

Our purpose in this paper is to state the genericity of the JJT-regularity for the particular class of one-parametric problems obtained by use of an embedding representing the quadratic penalty method.

Consider the general optimization problem

$$P \quad \min\{f(x)|\ x \in M\},$$

where

$$M := \{x \in I\!R^n|\ h_i(x) = 0,\ i \in I, \quad g_j(x) \geq 0,\ j \in J\}$$

and $I = \{1, \ldots, m\}$, $J = \{1, \ldots, p\}$.

In order to solve this problem we can use a sequence of unconstrained optimization problems with a quadratic penalty term as in the following one-parametric problem.

$$P(t) \quad \min\{f(x) + \left(\frac{t}{(1-t)}\right)^2 \left[\sum_{i\in I} h_i^2(x) + \sum_{j\in J} (\min\{g_j(x), 0\})^2\right] \mid x \in I\!R^n\}$$

In this case the values of the parameter are taken increasing to 1.

In [7] the close connection of the solution of this problem with a one-parametric constrained one is stated. After some transformations in order to get better properties of the one-parametric problem, the finally proposed embeddings are of the form:

$$\widehat{P}(t) \quad \min\{\hat{f}(x, v, w, t) | (x, v, w) \in \hat{M}(t)\}, \tag{1}$$

where

$$\hat{f}(x, v, w, t) := t f(x) + (1-t)\|x - x^0\|^2 + \|v - v^0\|^2 + \|w - w^0\|^2$$

and

$$\hat{M}(t) := \left\{ (x, v, w) \in I\!R^{n+m+s} \;\middle|\; \begin{array}{c} t h_i(x) + (1-t)(v_i - v_i^0) = 0,\ i \in I \\ t g_j(x) + (1-t)(w_j - w_j^1) \geq 0,\ j \in J \\ p - \|x - x^1\|^2 - \|v - v^0\|^2 - \|w - w^1\|^2 \geq 0 \end{array} \right\}$$

Here the vectors $(x^0, v^0, w^0) \in I\!R^{n+m+s}$, $w^1 \in I\!R^s$, $x^1 \in I\!R^n$, and the positive number $p \in I\!R$ are fixed and such that

$$\begin{aligned} \|x^0 - x^1\|^2 + \|w^0 - w^1\|^2 &< p \\ w_j^1 - w_j^0 &< 0, \ \forall j \in J \end{aligned} \tag{2}$$

We will state the generic character of the JJT-regularity for the parametric problem defined in (1). Many classical concepts and results related to optimization theory and one-parametric optimization will be supposed to be known. Let us mention some of them:

- The concept of generalized critical points, and stationary points.
- The linear independence constraint qualification (LICQ).
- The five types of generalized critical points defining regularity in the sense of Jongen, Jonker and Twilt.
- The strong, or Whitney, topology in the space of all three time differentiable mappings.

For a definition of most of the concepts (and even more) see e.g. the books [16, 17] and the paper [14].

The rest of the report is organized as follows. In the second section the generic property will be presented and its proof will be reduced to three claims. The proof of these claims is the purpose of the third section. Some conclusions and examples are presented in the last and fourth section.

Thoughout the report, we employ the following notation. If $K \subset \{1, \ldots, s\}$ and $d \in I\!R^s$, we denote by d_K the subvector composed from the components d_i, $i \in K$. For a differentiable function $f : I\!R^s \to I\!R$, $Df(x)$ denotes the row vector of partial derivatives.

2 Genericity result.

We begin with recalling the definition of a JJT-regular parametric optimization problem in a slightly modified form (see [14] or [10]).

Definition 1

Let us consider a general parametric optimization problem $P(t)$ defined by the functions $(f(y,t), H(y,t), G(y,t)) \in C^3(\mathbb{R}^{n_1+1}, \mathbb{R}^{n_2+1})$. Let S be a subset of $\mathbb{R}^{n_1+1}$. We call the parametric problem $P(t)$ JJT-regular on S if each generalized critical point of $P(t)$ contained in S belongs to one of the types 1, 2, 3, 4 or 5 (see [14]). This property will be denoted as follows: $P(t) \in \mathcal{F}|_S$.

If S has the structure $S_1 \times \mathbb{R}$, with $S_1 \subset \mathbb{R}^{n_1}$ (or $\mathbb{R}^{n_1} \times S_2$, with $S_2 \subset \mathbb{R}$) we will use the notation $P(t) \in \mathcal{F}|_{S_1}$ (or $P(t) \in \mathcal{F}|_{S_2}$) instead of $P(t) \in \mathcal{F}|_{S_1 \times \mathbb{R}}$ (or $P(t) \in \mathcal{F}|_{\mathbb{R}^{n_1} \times S_2}$). Finally, $P(t) \in \mathcal{F}$ stands for the case of $S = \mathbb{R}^{n_1+1}$, as usually.

As indicated in the introduction, we consider the embedding defined in (1).This parametric problem $\hat{P}(t)$ is constructed using the function data

$$(f, H, G) := (f(x), h_i(x), i \in I, g_j(x), j \in J) \in C^3(\mathbb{R}^n, \mathbb{R}^{m+s+1})$$

This construction can be interpreted as a mapping $\Phi : C^3(\mathbb{R}^n, \mathbb{R}^{m+s+1}) \longmapsto C^3(\mathbb{R}^{n+m+s+1}, \mathbb{R}^{m+s+2})$ defined as follows:

$$\Phi(f, H, G) = \begin{pmatrix} tf(x) + (1-t)\|x - x^0\|^2 + \|v - v^0\|^2 - \|w - w^0\|^2 \\ th_1(x) + (1-t)(v_1 - v_1^0) \\ \vdots \\ th_m(x) + (1-t)(v_m - v_m^0) \\ tg_1(x) + (1-t)(w_1 - w_1^1) \\ \vdots \\ tg_s(x) + (1-t)(w_s - w_s^1) \\ p - \|x - x^1\|^2 - \|v - v^0\|^2 - \|w - w^1\|^2 \end{pmatrix} \tag{3}$$

In this section we analyze the assumption $\Phi(f, H, G) \in \mathcal{F}$. Then we have to study the properties of the set $\Phi^{-1}(\mathcal{F})$. It is well known that $\mathcal{F}$ is an open and dense subset of the space $C^3(\mathbb{R}^{n+m+s+1}, \mathbb{R}^{m+s+2})$ endowed with the strong (or Whitney) topology. In this sense it is a generic assumption to suppose that a general one-parametric optimization problem belongs to the set $\mathcal{F}$. It is important to note that, from the genericity of $\mathcal{F}$, one cannot directly conclude any topological property of the set $\Phi^{-1}(\mathcal{F})$ with respect to the strong topology. In order to conclude such properties the mapping Φ should be analyzed in detail.

Theorem 1
The set $\Phi^{-1}(\mathcal{F}|_{[0,1]}) \subset C^3(I\!\!R^n, I\!\!R^{m+s+1})$ *is open and dense with respect to the strong topology.*

The result can be obtained by means of a local$\longrightarrow$global construction from the differential topology if we prove the local stability of the condition $\Phi(f, H, G) \in \mathcal{F}|_{[0,1]}$, and the possibility to achieve this condition locally by means of perturbations with polynomials up to degree two.

The local stability can be easily reduced by continuity and compacity arguments to the local stability of the five types of generalized critical points definning the JJT-regularity.

The property to be proved concerning perturbations can be formulated in the following form:

Proposition 1
Let $(\bar{f}, \overline{h}_1, \ldots, \overline{h}_m, \bar{g}_1, \ldots, \bar{g}_s) \in C^3(I\!\!R^n, I\!\!R^{m+s+1})$ *be fixed, then each measurable subset of the set*

$$\left\{ (A, b, c^1, \ldots, c^{m+s}, d) \;\middle|\; \Phi \begin{pmatrix} \bar{f} + 0.5x^T Ax + b^T x \\ \bar{h}_1 + c_1^T x + d_1 \\ \vdots \\ \bar{h}_m + c_m^T x + d_m \\ \bar{g}_1 + c_{m+1}^T x + d_{m+1} \\ \vdots \\ \bar{g}_s + c_{m+s}^T x + d_{m+s} \end{pmatrix} \notin \mathcal{F}|_{[0,1]} \right\} \qquad (4)$$

has Lebesgue measure zero.

The vector $A \in I\!\!R^{0.5n(n+1)}$ is considered here as a symmetric matrix, the vector b and the vectors c_i and c_j belong to the space $I\!\!R^n$, and the vector d belongs to the space $I\!\!R^{m+s}$. Therefore in (4) a subset of the space $I\!\!R^{0.5n(n+1)+n+n(m+s)+m+s}$ is defined and the Lebesgue measure affirmation of the above theorem is understood in this space.

With $\mathcal{A}$ will be denoted the whole vector of perturbations and with $\mathcal{B}$ the vector of perturbation corresponding to the constraints, then $\mathcal{A} = (A, b, \mathcal{B}) = (A, b, c^1, \ldots, c^{m+s}, d)$

For a fixed $\mathcal{A} = (A, b, \mathcal{B})$ we will use $\Phi(\bar{f}, \overline{H}, \overline{G}, \mathcal{A})$ or $\overline{\Phi}(x, v, w, t, \mathcal{A})$ for referring to the problem with data:

$$\Phi \begin{pmatrix} \bar{f} + 0.5x^T Ax + b^T x \\ \bar{h}_1 + c_1^T x + d_1 \\ \vdots \\ \bar{h}_m + c_m^T x + d_m \\ \bar{g}_1 + c_{m+1}^T x + d_{m+1} \\ \vdots \\ \bar{g}_s + c_{m+s}^T x + d_{m+s} \end{pmatrix}$$

and $\hat{M}(\overline{H}, \overline{G}, \mathcal{B}, t)$ to its parameter-dependent feasible set.

Following the same lines as for the analogous Theorem in [25] the proof of our Proposition 1 is reduced to the following three claims.

Claim 1
For almost all $\mathcal{A} \in I\!R^{0.5n(n+1)+n+n(m+s)+m+s}$ each generalized critical point satisfying LICQ, for the problem defined with the data $\Phi(\bar{f}, \overline{H}, \overline{G}, \mathcal{A})$, is either a point of Type 1, Type 2, or Type 3.

Claim 2
For almost all $\mathcal{B} \in I\!R^{n(m+s)+m+s}$ the following two conditions hold:

1. *The subset*

$$\Big\{(x, v, w, t) \in \hat{M}(\overline{H}, \overline{G}, \mathcal{B}, t) | \textit{LICQ fails to hold}\Big\} \tag{5}$$

   is a zero-dimensional manifold.

2. *At each point $(\overline{x}, \overline{v}, \overline{w}, \overline{t})$ belonging to the set defined in (5) the following two conditions hold true:*

   (a) The gradients (derivatives with respect to x, v, w and t) of the active constraints (equalities and active inequalities) are linearly independent.

   (b) Each vanishing linear combination of the partial derivatives of the active constraints has components corresponding to active inequalities different from zero.

Claim 3
Let $\mathcal{B} \in I\!R^{n(m+s)+m+s}$ be a fixed vector satisfying the conditions of the Claim 2, then for almost all $(A, b) \in I\!R^{0.5n(n+1)+n}$, each point belonging to the set defined in (5) *is of Type 4.*

The main tool used for the proof of these claims is the Parametrized Sard's Theorem (cf.[1]). The idea of the proofs consists in finding mappings that depend on the parameters of the perturbations and on other variables satisfying two conditions (these other variables containing the variables of the one-parametric problems and the multipliers):

1. Zero is a regular value of the mapping.
2. If we fix the perturbation parameters and consider the mapping depending only on the other variables, then the regularity of the value zero for the restricted mapping implies conditions contained in the definitions of the five types of generalized critical points.

The main difference between the claims stated above and the analogous statements in [25] is the reduced number of perturbation parameters arising in our one-parametric problem. Note, for example, that there are no perturbations with respect to the variables v and w , and that the compatification constraints are not perturbed, either. In this sense the main difficulty consists in proving that zero remains a regular value.

3 Proof of the claims

We fix a tuple of functions $(\bar{f}, \bar{H}, \bar{G}) \in C^3(I\!R^n, I\!R^{m+s+1})$ for the whole section. Let us introduce some practical notations before getting into details. The new variable y is defined as $y = (x, v, w) \in I\!R^{n+m+s}$ and $u \in I\!R^{m+s+1}$ will be used to denote the multipliers associated with the problem $\overline{\Phi}(y, t, \mathcal{A})$.

It is also necesary to refer separately the $m+s+2$ components of $\overline{\Phi}(y, t, \mathcal{A})$. Let us then introduce a subindex in the following form:

$$\begin{aligned}\overline{\Phi}_0(y, t, (A, b)) \;\; = \;\; & t\overline{f}(x) + (1-t)\|x - x^0\|^2 + \|v - v^0\|^2 - \\ & -\|w - w^0\|^2 + t(0.5x^T Ax + b^T x)\end{aligned}$$

$$
\begin{aligned}
\overline{\Phi}_i(y,t,\mathcal{B}) &= t\overline{h}_i(x) + (1-t)(v_i - v_i^0) + t(c_i^T x + d_i),\ i \in I \\
\overline{\Phi}_{m+j}(y,t,\mathcal{B}) &= t\overline{g}_j(x) + (1-t)(w_j - w_j^1) + t(c_{m+j}^T x + d_{m+j}),\ j \in J \\
\overline{\Phi}_{m+s+1}(y,t,\mathcal{B}) &= p - \|x - x^1\|^2 - \|v - v^0\|^2 - \|w - w^1\|^2
\end{aligned}
$$

For an index subset $K \subseteq \{0,\ldots,m+s+1\}$ we will use $\overline{\Phi}_K(x,v,w,t,\mathcal{A}) \in C^3(I\!R^{n+m+s+1}, I\!R^{|K|})$ for referring to the mapping formed with index in K.

Proof of Claim 1 :

In this proof we define a mapping according to three fixed index sets. Namely

- $K_1 \subseteq \{m+1,\ldots,m+s+1\}$, interpreted as the index set of active inequality constraints for the problem $\overline{\Phi}(y,t,\mathcal{A})$. We suppose the most general case, i.e. $\{m+s+1\} \subset K_1$.
- $K_2 \subseteq \{1,\ldots,\gamma\}$ where $\gamma = n+m+s+m+|K_1|$ and $|.|$ denotes the cardinality.
- $K_3 \subseteq K_1$, to be interpreted as the vanishing multipliers associated to active inequality constraints.

Let us denote by $\overline{K_1}$ the index set obtained by adding the equality constraint indices to K_1 , then $\overline{K_1} = \{1,\ldots,m\} \cup K_1$. We define now the following mapping (see [18] for the meaning of this system).

$$
H(y,t,u,\mathcal{A}) = \begin{pmatrix} D_y\overline{\Phi}_0(y,t,(A,b)) - \sum_{k\in\overline{K_1}} u_k D_y \overline{\Phi}_k(y,t,\mathcal{B}) \\ \\ \overline{\Phi}_k(y,t,\mathcal{B}), \quad k \in \overline{K_1} \\ \\ u_k + \overline{\Phi}_k(y,t,\mathcal{B}), \quad k \in \{1,\ldots,m+s+1\} \setminus \overline{K_1} \end{pmatrix}
$$

In the following we will denote the set of symmetric matrices of size $p \times p$ by $M^s(p)$ and identify the elements of this space with the vectors of the space $I\!R^{0.5p(p+1)}$. Let us define now the following map, giving its values on the space $M^s(\gamma)$.

$$
M(y,t,u,\mathcal{A}) = \begin{pmatrix} D_y^2[\overline{\Phi}_0(y,t,(A,b)) - \sum_{k\in\overline{K_1}} u_k \overline{\Phi}_k(y,t,\mathcal{B})] & D_y^T \overline{\Phi}_{\overline{K_1}}(y,t,\mathcal{B}) \\ \\ D_y \overline{\Phi}_{\overline{K_1}}(y,t,\mathcal{B}) & 0 \end{pmatrix}
$$

Let $\mathcal{M}(K_2) \subset M^s(\gamma)$ be the set of those symmetric matrices with rank $|K_2|$ whose columns with indices in K_2 are linearly independent. It is known (cf. [27]), that $\mathcal{M}(K_2)$ is a smooth manifold with codimension $0.5(\gamma - |K_2|)(\gamma - |K_2| + 1)$. We denote by $\Theta \in C^3(\mathbb{R}^{0.5\gamma(\gamma+1)}, \mathbb{R}^{0.5(\gamma-|K_2|)(\gamma-|K_2|+1)})$ a smooth mapping defining the manifold $\mathcal{M}(K_2)$ locally. An important observation about these mappings Θ is that their Jacobian contains an identity matrix corresponding to the partial derivatives with respect to the components of the matrices not belonging to columns of index in K_2.

We can now introduce the mapping Ψ_1 that plays the role mentioned in the previous section.

$$\Psi_1(y, t, u, z, \mathcal{A}) = \begin{pmatrix} H(y, t, u, \mathcal{A}) \\ M(y, t, u, \mathcal{A}) - z \\ \Theta(z) \\ u_{K_3} \end{pmatrix}$$

Here the variable z belongs to the space $\mathbb{R}^{0.5\gamma(\gamma+1)}$. The principal aim of this proof consists in the study of the Jacobian of Ψ_1, which is presented in the following matrix 6.

$$\begin{array}{c|cccccc|}
 & D_{(y,t)} & D_u & D_z & D_A & D_b & D_d \\
\hline
 & & & & \otimes & tI_n & 0 \\
 & & & & 0 & 0 & 0 \\
H & \otimes & \otimes & 0 & 0 & 0 & tI_{m+s} \\
 & & & & 0 & 0 & 0 \\
M-z & \otimes & \otimes & -I_{0.5\gamma(\gamma+1)} & \begin{array}{c} tI_{0.5n(n+1)} \\ 0 \end{array} & 0 & 0 \\
\Theta(z) & 0 & 0 & I_{0.5(\gamma-|K_2|)(\gamma-|K_2|+1)}|\times & 0 & 0 & 0 \\
u_{K_3} & 0 & I_{|K_3|}|0 & 0 & 0 & 0 & 0 \\
\hline
\end{array} \tag{6}$$

We will prove that, under some restrictions, this matrix has full row rank if $\Psi_1(y, t, u, z, \mathcal{A}) = 0$. These restriction are:

1. t is not 0 or 1.

2. The matrix $D_y \overline{\Phi}_{\overline{K_1}}(y, t, \mathcal{B})$ has full row rank (can be interpreted as LICQ).

From the structure of the matrix (6) it is possible to note that, if K_2 contains the set of indices $\{n+1, \ldots, \gamma\}$, then $0.5n(n+1)$ is greater than or equal to $0.5(\gamma - |K_2|)(\gamma - |K_2| + 1)$ and the desired rank condition can be reduced to a rank condition over a partial Jacobian of some components from $H(y, t, u, \mathcal{A})$ (to be explained later).

Since $\Psi_1(y, t, u, z, \mathcal{A}) = 0$, the structure of K_2 mentioned above is reduced to an analysis of the last $2m + s + |K_1|$ columns of $M(y, t, u, \mathcal{A})$ (see the following matrix (7))

$$
\begin{array}{c}
\begin{array}{cccc}
\quad D_x \quad & \quad D_{(v,w)} \quad & \quad k \in \overline{K_1} \setminus \{m+s+1\} \quad & \quad m+s+1 \quad
\end{array} \\
\begin{bmatrix}
\otimes & 0 & D_x^T \overline{\Phi}_k & -2(x-x^1) \\
0 & 2I_{m+s} & \begin{array}{c} (1-t)I \\ 0 \end{array} & \begin{array}{c} -2(v-v^0) \\ -2(w-w^1) \end{array} \\
D_x \overline{\Phi}_k & (1-t)I \;\; | \;\; 0 & 0 & 0 \\
-2(x-x^1)^T & -2(v,w)^T - 2(v^0,w^1)^T & 0 & 0
\end{bmatrix}
\end{array}
\tag{7}
$$

If the vector $(w - w^1)$ has all its components with indices from K_1 equal to zero, then the rank condition follows from the restrictions imposed by use of a suitable part of the matrix $2I_{m+s}$. In the other case we are done, because the sub-matrix presented in (8) is nonsingular.

$$
\begin{pmatrix}
2I_{m+s} & \begin{array}{c} (1-t)I \\ 0 \end{array} & \begin{array}{c} -2(v-v^0) \\ -2(w-w^1) \end{array} \\
(1-t)I \;\; | \;\; 0 & 0 & \\
-2(v,w)^T - 2(v^0,w^1)^T & 0 &
\end{pmatrix}
\tag{8}
$$

Finally, we can conclude the full rank of the last $2m + s + |K_1|$ columns of $M(y, t, u, \mathcal{A})$.

Now we come back to the matrix (6). Let $\vartheta = (\vartheta^H, \vartheta^M, \vartheta^\Theta, \vartheta^{K_3})$ be the coefficients of a vanishing linear combination of the rows of this matrix. We want to show that $\vartheta = 0$. The coefficients corresponding to the map H can be divided into the following parts $\vartheta^H = (\vartheta^x, \vartheta^{(v,w)}, \vartheta^{\{1,\ldots,m+s\}}, \vartheta^{m+s+1})$, where ϑ^{m+s+1} corresponds to the compactification constraint and $\vartheta^{\{1,\ldots,m+s\}}$ to the other equality and inequality constraints.

From the identity tI_n (resp. tI_{m+s}) found in (6) on the row corresponding to derivatives with respect to the parameter b (resp. d) it follows immediately that $\vartheta^x = 0$ (resp. $\vartheta^{\{1,\ldots,m+s\}} = 0$). Now, these equalities, the proved fact about the last columns of $M(y,t,u,\mathcal{A})$ and the structure of (6) lead to $\vartheta^M = 0$ and $\vartheta^\Theta = 0$.

In order to prove that $(\vartheta^{(v,w)}, \vartheta^{m+s+1}, \vartheta^{K_3}) = 0$ let us analyze the rest of the jacobian given in (6). The part to be annalysed is presented in the following matrix (9).

$$\begin{array}{cccc} D_x & D_{(v,w)} & \{1,\ldots,m\} & K_1 \end{array}$$
$$\left[\begin{array}{cccc} & & (1-t)I_m \quad 0 & -2(v-v^0) \\ 0 & 2I_{m+s} & 0 \quad \begin{array}{c}(1-t)I_{K_1-1}\\ 0\end{array} & -2(w-w^1) \\ -2(x-x^1)^T & -2(v-v^0)^T|-2(w-w^1)^T & 0 & 0 \\ 0 & 0 & 0 & 0 \mid I_{|K_3|} \mid 0 \end{array}\right] \tag{9}$$

If we can prove that $\vartheta^{m+s+1} = 0$ then we are done. This identity follows immediately in case that $(x,v) \neq (x^1,v^0)$ (observe the structure of the columns associated to D_v and $\{1,\ldots,m\}$). We suppose then that $(x,v) = (x^1,v^0)$. Since the compactification constraint is supposed to be active, it holds that $\|w-w^1\|^2 = p \neq 0$.

If the index $m+s+1$ does not belong to the set K_3(In other words, the last column of (9) contains only zeros in the rows corresponding to ϑ^{K_3}), then the nonsingularity of the submatrix

$$\begin{pmatrix} 2I_{m+s} & \begin{array}{c}-2(v-v^0)\\ -2(w-w^1)\end{array} \\ -2(v-v^0)^T|-2(w-w^1)^T & 0 \end{pmatrix}$$

implies that $(\vartheta^{(v,w)}, \vartheta^{m+s+1}) = 0$ and we are done.

In case of $\{m+s+1\} \subset K_3$ ($u_{m+s+1} = 0$), the equation

$$D_y\overline{\Phi}_0(y,t,(A,b)) - \sum_{k\in\overline{K_1}} u_k D_y \overline{\Phi}_k(y,t,\mathcal{B}) = 0$$

(contained in $H(y,t,u,\mathcal{A})=0$) considered only in the derivatives D_w is reduced to the following equalities on $j \in J$.

$$2(w_j - w_j^0) = \begin{cases} (1-t)u_{m+j}, & m+j \in K_1 \\ 0, & m+j \notin K_1 \end{cases} \tag{10}$$

On the other hand, the inequalities

$$\begin{array}{rcl} \|w - w^1\|^2 & = & p \\ \|w^0 - w^1\|^2 & < & p \end{array}$$

imply the existence of an index (without lost of generality 1) in J such that w_1 is different from w_1^1 and from w_1^0. We obtain from the equalities (10), that $m+1 \in K_1$, $u_{m+1} \notin K_3$ and $w_1 - w_1^0 \neq 0$. Returning to the matrix (9) we note that, it follows from $u_{m+1} \notin K_3$ that $\vartheta^{w_1} = 0$ and, finally, $\vartheta^{m+s+1}(w_1 - w_1^1) = 0$, which implies $\vartheta^{m+s+1} = 0$ and we are done.

It has been proved that at each zero of the mapping $\Psi_1(y,t,u,z,\mathcal{A})$, such that t is not 0 or 1, and $D_y\overline{\Phi}_{\overline{K_1}}(y,t,\mathcal{B})$ has full row rank, the Jacobian of Ψ_1 has full row rank, too. This condition is analogous to the transversality used in [25], and the conclusion about the types of generalized critical points is obtained with the same arguments. This observation concludes the proof of the first claim. $\square$

We go into the proof of the second claim and the notations introduced in the proof of the first one will be used implicitly.

Proof of the Claim 2 :

The proof will be done again for the case of $t \neq 1$ and $t \neq 0$, as a consequence the compactification constraint must be active at each point where the LICQ fails to hold, because the gradients corresponding to the other constraints are linearly independent. The mapping to be used in this proof depends on two index sets as in the claim before. Let us then fix again K_1 and K_3 as in the above proof. Since the compactification constraint must be active we suppose, for simplicity of notation that $\{m+s+1\} \notin K_1$.In this proof we have to introduce a variable to denoting the coefficients of a linear combination of the active constraints. Let us take the variable λ having indices in $\overline{K_1}$(notation as in the proof of the first claim), and then $\lambda \in I\!R^{|\overline{K_1}|}$. The mapping to be used in this proof is defined as follows:

$$\Psi_2(y,t,\lambda,\mathcal{B}) = \begin{pmatrix} \sum_{k\in\overline{K_1}} \lambda_k D_y\overline{\Phi}_k(y,t,\mathcal{B}) + D_y\overline{\Phi}_{m+s+1}(y,t,\mathcal{B}) \\ \overline{\Phi}_k(y,t,\mathcal{B}), \quad k\in\overline{K_1} \\ \overline{\Phi}_{m+s+1}(y,t,\mathcal{B}) \\ \lambda_{K_3} \end{pmatrix}$$

Since $D_y\overline{\Phi}_{m+s+1}(y,t,\mathcal{B}) = -2(x-x^1, v-v^0, w-w^1) \neq 0$ at a zero $(y,t,\lambda,\mathcal{B})$ of Ψ_2 this mapping must be different from the zero vector by $\lambda \in \mathbb{R}^{|\overline{K_1}|}$. Let us suppose w.l.o.g that $\lambda_1 \neq 0$, which implies immediately $(v_1 - v_1^0) \neq 0$ and $1 \notin K_3$. We consider in the matrix (11) the Jacobian of this mapping in those points where $\Psi_2(y,t,u,\mathcal{B}) = 0$, $t \neq 1$ and $t \neq 0$.

$$\begin{array}{cccccc} D_x & D_{(v,w)} & D_{\lambda_1} & D_{\lambda_{K_3}} & D_{c_1} & D_{d_{\overline{K_1}}} \end{array}$$
$$\left[\begin{array}{cccccc} \otimes & 0 & D_x^T\overline{\Phi}_1 & \otimes & \lambda_1 t I_n & 0 \\ 0 & -2I_{m+s} & \begin{matrix}(1-t)\\0\end{matrix} & 0 & 0 & 0 \\ D_x\overline{\Phi}_{\overline{K_1}} & (1-t)I_{\overline{K_1}}|0 & 0 & 0 & \otimes & tI_{K_1} \\ -2(x-x^1)^T & -2(v-v^0)^T|-2(w-w^1)^T & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I_{|K_3|} & 0 & 0 \end{array}\right] \qquad (11)$$

In order to note that this matrix has full row rank we suppose that the vector $\vartheta = (\vartheta^x, \vartheta^{(v,w)}, \vartheta^{\overline{K_1}}, \vartheta^{m+s+1}, \vartheta^{K_3})$ contains the coefficients of a vanishing linear combination from the rows of (11) (Note that the subdivision of ϑ concides with that of the rows). We should prove that $\vartheta = 0$.

Since t and λ_1 are different from zero, the matrices tI_{K_1}, $\lambda_1 t I_n$ and $I_{|K_3|}$ lead to the equalities $\vartheta^x = 0$, $\vartheta^{\overline{K_1}} = 0$ and $\vartheta^{K_3} = 0$. Now the nonzero scalar $(1-t)$ in the column corresponding to the partial derivative with respect to λ_1 implies that the coefficient ϑ^{v_1} is zero, too. Analyzing the column corresponding to the partial derivative with respect to v_1 and taking into account that $(v_1 - v_1^0) \neq 0$ we can conclude that $\vartheta^{m+s+1} = 0$. The same can be obtained for the other components of $\vartheta^{(v,w)}$ (for ϑ^{v_1} this is already known) by use of a suitable part of the matrix $-2I_{m+s}$.

Applying the parametrized Sard's Theorem, the same full rank condition about the Jacobian evaluated in the zeros can be ensured, assuming the parameter $\mathcal{B}$ to be fixed in a set with a complement of measure zero. These leads to the statements of this lemma with similar arguments as used in the proof of Lemma 1 in [25] (note, for example, that the transversal condition used there is rewritten here in the full row rank of the Jacobians). With this reference we conclude the proof of the claim. □

Let us now deal with the last claim.

Proof of Claim 3 :

In this proof we work with a fixed value of the parameter $\mathcal{B}$ satisfying the thesis of the above claim. For simplicity we suppose that $\mathcal{B} = 0$ and omit it in the notations. We should prove that for almost all values of the remaining parameters each point belonging to the set defined in (5) is of Type 4 (Type 5 is not possible because the number of active constraints can not be a greater than the number of variables in this parametric problem). Since the set given in (5) is a countable one, it is sufficient to prove the statement for a fixed point $(\overline{y}, \overline{t})$ belonging to the set defined by (5). We prove the claim again for the case that t is different from 0 and 1. In particular it follows that the compactification constraint must be active at $(\overline{y}, \overline{t})$. Following the notation used in the above proof let us denote by K_1 the index set of active inequality constraints at $(\overline{y}, \overline{t})$. Let $\overline{\lambda}$ be a vector such that

$$\sum_{k \in \overline{K_1}} \overline{\lambda}_k D_y \overline{\Phi}_k(\overline{y}, \overline{t}) + D_y \overline{\Phi}_{m+s+1}(\overline{y}, \overline{t}) = 0$$

From the results of the above claim it is not difficult to conclude the following facts:

- The vector $\overline{\lambda} \in I\!\!R^{|\overline{K_1}|}$ is uniquely determined and all its components are different from zero.
- $D_t L(\overline{y}, \overline{t}) \neq 0$ and the matrix $W^T D_y^2 L(\overline{y}, \overline{t}) W$ is regular, where by definition:
$$L(y, t) := \sum_{k \in \overline{K_1}} \overline{\lambda}_k \overline{\Phi}_k(y, t) + \overline{\Phi}_{m+s+1}(y, t)$$
and the columns of the matrix W form a basis for the orthogonal of the linear space generated by the vectors $\{D_y \overline{\Phi}_k(\overline{y}, \overline{t}), \quad k \in \overline{K_1}\}$.

The result of this claim can be reduced to the fulfilment of the following two inequalities for almost all values of the parameters (A, b):

1. $D_y \overline{\Phi}_0(\overline{y}, \overline{t}, (A, b)) W \neq 0$.

2. $D_y\overline{\Phi}_0(\overline{y},\overline{t},(A,b))W(W^TD_y^2L(\overline{y},\overline{t})W)^{-1}W^TD_y^T\overline{\Phi}_0(\overline{y},\overline{t},(A,b)) \neq 0.$

In order to prove these inequalities let us calculate a matrix W. Since the columns of W form a basis for the orthogonal of the matrix

$$\begin{array}{c}
\begin{array}{ccc} \qquad D_x \qquad & \qquad D_v \qquad & \qquad D_w \qquad \end{array} \\
\left[\begin{array}{ccc}
D_x\overline{\Phi}_I(\overline{y},\overline{t}) & (1-t)I_m & 0 \\
D_x\overline{\Phi}_{K_1}(\overline{y},\overline{t}) & 0 & (1-t)I_{|K_1|} \mid 0
\end{array}\right]
\end{array}$$

the matrix W can be taken with the structure

$$W = \begin{pmatrix} -(1-t)I_n & 0 \\ D_x\overline{\Phi}_{\overline{K_1}}(\overline{y},\overline{t}) & 0 \\ 0 & I_{s-|K_1|} \end{pmatrix}.$$

Now the first inequality is rewritten as

$$D_y\overline{\Phi}_0(\overline{y},\overline{t},(A,b))W = \begin{pmatrix} t(t-1)(\overline{x}^TA+b^T)+\xi_1^T & \xi_2^T \end{pmatrix} \neq 0$$

where $\xi_1 \in I\!R^n$ and $\xi_1 \in I\!R^{s-|K_1|}$ are vectors not depending on the parameters (A,b). Since ξ_1 is fixed, the first part of $D_y\overline{\Phi}_0(\overline{y},\overline{t},(A,b))W$ is different from zero for almost all (A,b). (By fixing A the vector b is also fixed automatically).

For the second inequality we need a formula for $(W^TD_y^2L(\overline{y},\overline{t})W)^{-1}$. By definition this matrix is obtained by multiplying the following terms:

$$W^T \begin{pmatrix} \Delta_1 & 0 & 0 \\ 0 & -2I_{m+|K_1|} & 0 \\ 0 & 0 & -2I_{s-|K_1|} \end{pmatrix} W.$$

After a short calculation we obtain that

$$(W^TD_y^2L(\overline{y},\overline{t})W)^{-1} = \begin{pmatrix} \Delta_2 & 0 \\ 0 & -2I_{s-|K_1|} \end{pmatrix},$$

where the expressions Δ stand for terms being not interesting for the analysis (in particular, these terms do not depend on the parameters (A,b)). It must be noted that the term Δ_2 is a nonsingular matrix of size $n \times n$. Finally, the desired inequality can be written in the form

$$(t(t-1)(\overline{x}^TA+b^T)+\xi_1^T)\Delta_2(t(t-1)(A\overline{x}+b)+\xi_1) - 2||\xi_2||^2 \neq 0.$$

The nonsingularity of Δ_2 and the independence of (A, b) of the terms Δ_2, ξ_1 and ξ_2, implies the above inequality to hold for almost all values of the mentioned parameters.

This observation concludes the proof of this claim. $\Box$

Since these claims have been proved in case that t is different from 0 or 1, it should be mentioned how the proof for these values can be completed.

For the value $t = 0$ we have no parameter of perturbation in our optimization problem. With the following lemma we aim at proving that the regularity in $t = 0$ follows from the special way we selected (x^0, w^0) and (x^1, w^1) (see (1)).

Lemma 1

Each generalized critical point of the problem

$$\begin{array}{lc} min & ||x-x^0||^2 + ||v-v^0||^2 + ||w-w^0||^2 \\ s.t. & \\ & v - v^0 = 0 \\ & w - w^1 \geq 0 \\ & -||x-x^1||^2 - ||v-v^0||^2 - ||w-w^1||^2 \geq -p \end{array} \tag{12}$$

is nondegenerated.

Proof:

It is immediately noted that the LICQ holds at each feasible point of the problem defined in (12). For the sake of simplicity let us suppose that there are no equality constraints.

Let us begin proving the strict complementarity.

If a linear constraint is active, then the corresponding partial derivative of the quadratic constraint (equal to $-2(w_j - w_j^1)$) vanishes, and the inequality $w_j^0 \neq w_j^1$ implies that the corresponding multiplier is not zero. On the other hand, if the multiplier corresponding to the quadratic constraint vanishes, it follows that $x = x^0$, and for all $j \in J$, w_j is equal to w_j^0 or to w_j^1 for all $j \in J$. The latter fact together with the inequalities

$$||x^0 - x^1||^2 + ||w - w^1||^2 \leq ||x^0 - x^1||^2 + ||w^0 - w^1||^2 < p$$

implies that the last constraint cannot be active and, thus, the strict complementarity is proved.

The nonsingularity of the Hessian from the Lagrangian restricted to the tangent space is obtained by proving that the Hessian is always strictly definite. Since the linear constraints have no influence on this Hessian, it is only important to consider the matrices corresponding to the objective function and to the latter constraint. Both matrices represent the identity, but with different sign. Now the definiteness of the Hessian can be concluded if the multiplier corresponding to the quadratic constraint is different from -1. This follows immediately from the inequality $x^0 \neq x^1$. $\square$

For $t = 1$ we consider two cases separately. First, using the parametrized Sard's Theorem it can be easily proved that it holds for almost all values of the parameter $\mathcal{A}$: At each feasible point of the set

$$\left\{ (x, v, w) \in I\!R^{n+m+s} \;\middle|\; \begin{array}{c} \overline{h}_i(x) + c_i^T x + d_i = 0,\ i \in I \\ \overline{g}_j(x) + c_{m+j}^T x + d_{m+j} \geq 0, \quad j \in J \\ -||x - x^1||^2 - ||v - v^0||^2 - ||w - w^1||^2 \geq -p \end{array} \right\}$$

the LICQ is fulfilled.

As second step it should be proved that for almost all values of the parameter $\mathcal{A}$, each generalized critical point of the problem

$$\begin{array}{ll} \min & \overline{f}(x) + 0.5 x^T A x + b^T x + ||v - v^0||^2 + ||w - w^0||^2 \\ \text{s.t.} & \\ & \overline{h}_i(x) + c_i^T x + d_i = 0,\ i \in I \\ & \overline{g}_j(x) + c_{m+j}^T x + d_{m+j} \geq 0, \quad j \in J \\ & -||x - x^1||^2 - ||v - v^0||^2 - ||w - w^1||^2 \geq -p, \end{array} \tag{13}$$

where the LICQ is fulfilled, is a nondegenerated one.

The idea is to consider a mapping very similar to Ψ_2. The modifications to be done on Ψ_2 are the following:

1. Eliminate the variable t, since the problem given in (13) is not parametric.

2. Consider a new multiplier associated to the compactification inequality constraint.

3. Add the gradient of the objective function to the sum of gradients in Ψ_2

After these modifications, similar arguments (using also that $w^0 \neq w^1$) lead to the use of the parametrized Sard's Theorem and, after a simple analysis, to the mentioned result (note that a zero of the mentioned mapping is associated to each generalized critical point with the LICQ).

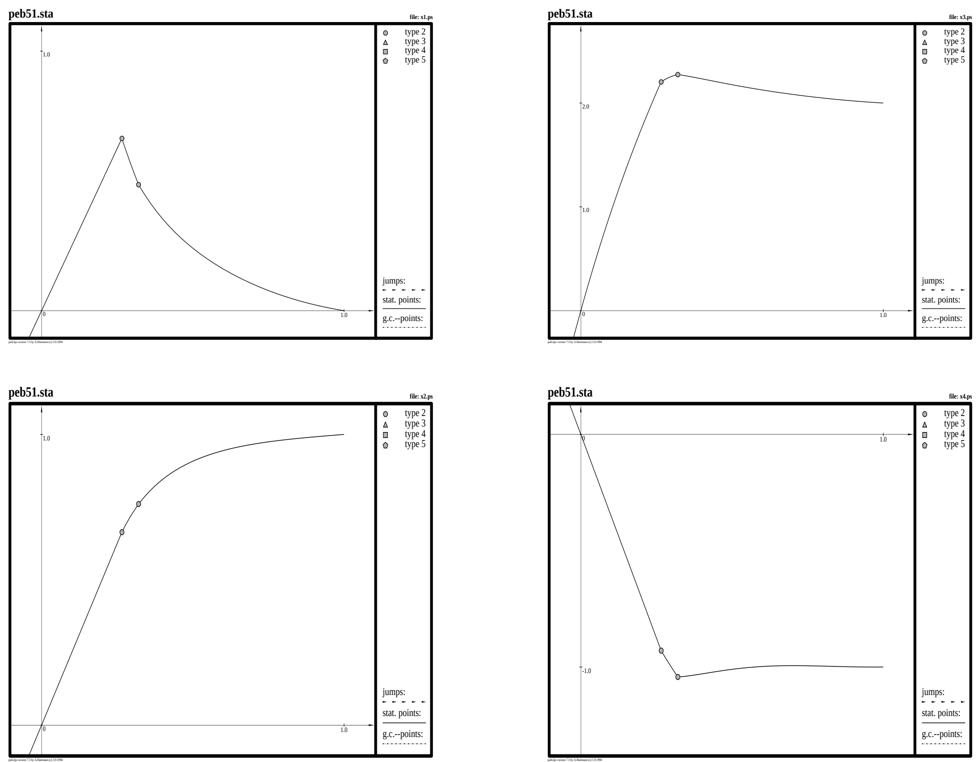

Figure 1: The four x-components of the solution curve of problem 46.

The proof for $t = 1$ is completed combinning both partial results and the fact that the intersection of a finite number of sets of measure zero is again a set of measure zero.

Le us still mention that the statement obtained for the latter case $t = 1$ provides a measurable set, and a suitable combination with the statement for t different from 1 can now be realized.

This observation concludes the proof of the genericity result.

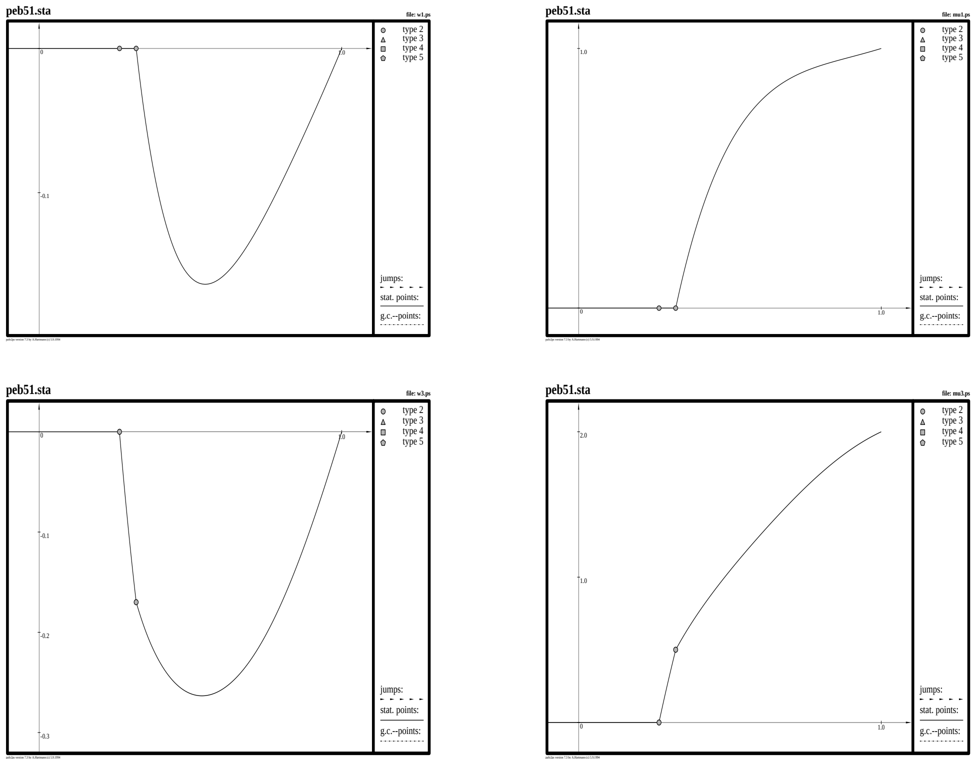

Figure 2: The component (w_1, w_3) (left) and the corresponding multiplier (u_1, u_3) in the solution curve of problem 46.

4 Examples and numerical results.

In order to give a better idea of the feasibility of this continuation method and its robusticity we have selected some standard problems from the literature and computed the corresponding paths by means of the embedding studied in this paper. Our purpose with these calculations is not to compare the power of the continuation methods with other algorithms. For the pathfollowing in our the system we have used PAFO, which was develop at the Humboldt-University Berlin. We are more interested in testing the possibility of calculating the actual solution of optimization problems by using pathfollowing procedures. Therefore, we mainly report about the success or failure of the quadratic penalty embedding for the calculation of a solution.

The problems selected were taken from two recent papers testing algo-

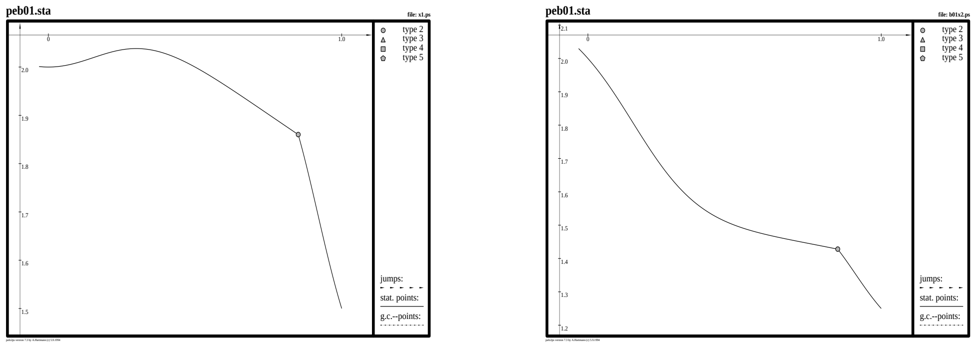

Figure 3: The x-components of the solution curves for example 1.

rithms for solving nonlinear optimization problems. From [20] we tested all problems except for the numbers 55, 13 and 118 (in the numeration of the book [13]). For all these examples the calculation of the given solution (using the startpoint proposed in the book) runs without any difficulties. In fact, singularities of Type 4 are not presented. From the paper [11] we tested almost all examples of nonlinear programming problems presented (some of them having too large a dimension are not yet solvable by PAFO). For the other group the studied embedding has also been successful. In the case of problem 1, a new solution which improves the solution proposed for the example is generated inclusively. Furthermore, we have also used the proposed starting points for these examples. Note that in our continuation method any kind of first approximation step, by the resolution of a fixed optimization problem, in order to aproximate the penalty curve is used.

We present some figures representing the graphics of the obtained solution curves for the problems mentioned above. The figures 1, 2 and 3 represent the solution curves obtained for the problem 46 of [13] (example of Rosen-Zusuki) and figure 3 corresponds to the calculations of the nonlinear problem 1 in [11].

As we present the results of the example 1 of [11] only to show that we obtain a new solution, the corresponding figure 3 contains information on the original variables x only.

Finally, we have implemented the following example.

$$f(x) \quad = \quad (x_1+4)^2 + 0.5x_2^2$$

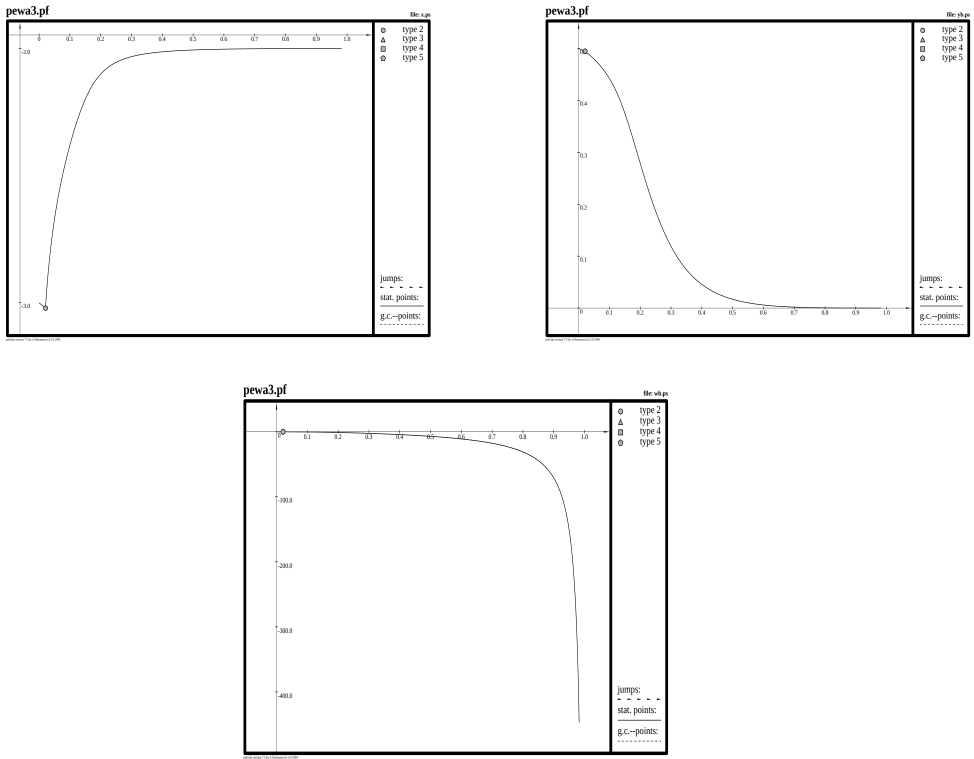

Figure 4: The curves $x_1(t)$ (left/top), $x_2(t)$ (right/top) and $w(t)$ (bottom) for example 14 without compactification.

$$\begin{aligned} & s.t \\ 0 \quad & \geq \quad 3x_1^4 + 8x_1^3 - 6x_1^2 - 24x_1 + x_2^2 \end{aligned} \tag{14}$$

Starting point: $x^0 = (-3, 0.5)$

In this example, the LICQ is not fulfilled in the points generated by using the pure approach of any sequential method. This is due to the selected startpoint. With the given startpoint the solution curve obtained converges to a point where the LICQ and the MFCQ are not fulfilled. In our continuation approach this will be notified by the appearance of a generalized critical point of Type 4, where it is imposible to jump.

We follow the curve of generalized critical points (for decreasing parameter t and negative multipliers) and finally, after the appearance of new degenerated critical points of Type 4, the value $t = 1$ is attained and a solution of

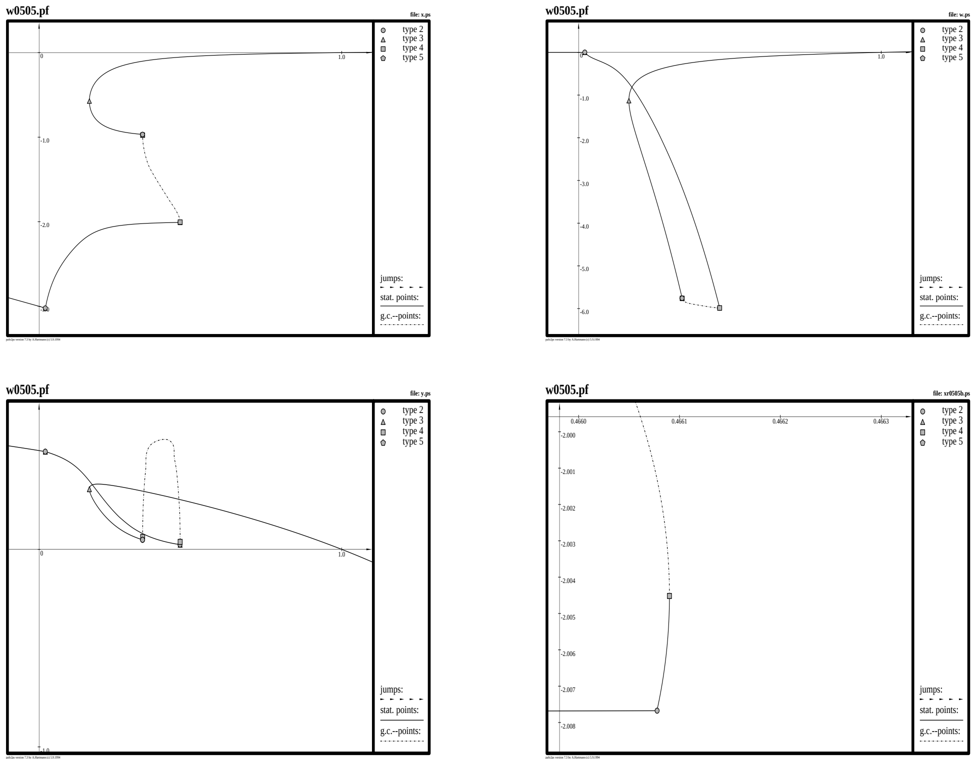

Figure 5: The curves $x_1(t)$ (left/top), $x_2(t)$ (left/bottom) and $w(t)$ (right/top) for example 14 with compactification.

the example is obtained. This example gives an idea of the utility of the continuation approach in order to obtain convergence without the assumption of the MFCQ, but with the generic assumption of regularity of the obtained embedding.

For a better understanding we have calculated this example twice, first without the use of a compactification constraint, and, second, with this constraint. Figure 4 shows some components for the first calculation and the sake of convergence is noted. In this case the "pure" penalty method would converge to the unfeasible point where the LICQ is violated. Observe how the introduced variable w tends to infinity when we approximate the point $(-2, 0)$, where the LICQ is violated.

In figure 5 the curves of the second calculation are presented. An enlargement of the $x(t)$ component in the neighbourhood of $x_1 = -2$ shows one of the points of Typ 4 in the right bottom.

Acknowledgement

We thank Prof. Dr. Jürgen Guddat for his support, and Dr. Dieter Nowack for his valuable help by the numerical experiments.

References

[1] Abraham R., Robbin J., Transversal Mappings and Flows, Benjamin, New York, 1967.

[2] Dentcheva, D., Gollmer, R., Guddat, J. and Rückmann, J.-J., Pathfollowing methods in nonlinear optimization II: Exact penalty methods. In Florenzano, M. et al. (eds.), Approximation and Optimization II (Proceedings of the 2nd International Conference Approximation and Optimization in the Caribbean, Havana, Cuba, 1993). Ser. Approximation and Optimization, Peter Lang Verlag, Frankfurt a.M., Berlin, Bern, New York, Paris, Wien 1995, 200-30

[3] Dentcheva, D., Guddat, J., Rückmann, J.-J., and Wendler, K., Pathfollowing methods in nonlinear optimization III: Lagrange Multiplier Embedding, in ZOR 1994, 127-152.

[4] Fletcher R., Practical methods of optimization, Vol 2, John Wiley, New York, 1981.

[5] Fletcher R., Penalty functions, in: Mathematical Programming: The State of the Art, A. Bachem, M. Grötschel and B. Korte, eds., Springer-Verlag, New York, 1983.

[6] Gill, P.E., Murray W. and Wright, M.H., Practical optimization, Academic Press, New York, 1981.

[7] Gollmer, R., Guddat,J., Guerra, F., Nowack, D. and Rückmann, J.-J.: Pathfollowing methods in nonlinear optimization I: penalty embedding. In: J. Guddat et al. (eds.), Parametric Optimization and Related Topics III, Peter Lang Verlag, Frankfurt a.M., Berlin, Bern, New York, Paris, Wien, (1993), 163-214

[8] Gómez W., Guddat J., Jongen H.Th., Rückmann J.J. and Solano C., Curvas críticas y saltos en optimización no lineal. Textbook in Spanish (in Preparation).

[9] Guddat J., Guerra F. and Nowack D., On the role of the Mangasarian-Fromovitz constraints qualification for penalty, exact penalty and Lagrange multiplier methods. Special volume: Mathematical Programming with Data Perturbations, Marcel Dekker, Inc, New York, 1997.

[10] Guddat J., Guerra F. and Jongen, H.Th., Parametric Optimization: singularities, pathfollowing and jumps. John Wiley, Chichester and Teubner, Stuttgart 1990.

[11] Harker P.T. and Xiao B., A nonsmooth Newton method for variational inequalities,II: numerical results, Mathematical Programming 65 (1994 195-216.

[12] Henn, M., Jonker, P. and Twilt, F., On the critical sets of one-parameter quadratic optimization problems, Durier, Roland et al. (eds.), Recent developments in optimization. 7th French-German conference on optimization, Dijon, France, June 27-July 2, 1994. Berlin: Springer-Verlag. Lect. Notes Econ. Math. Syst. 429, 183-197 (1995).

[13] Hock W. and Schittkowski K., Test examples for nonlinear programming code, Springer-Verlag, New York, 1981.

[14] Jongen, H.Th., Jonker, P. and Twilt, F., Critical sets in parametric optimization, Math. Program. 34 (1986), 333-353.

[15] Jongen, H.Th., Jonker, P. and Twilt, F., One-parameter families of optimization problems: equality constraints, in J. Opt. Theory Appl. 48 (1986), 141-161.

[16] Jongen, H.Th., Jonker, P. and Twilt, F., Nonlinear optimization in $\mathbb{R}^n$. I: Morse Theory, Chebychev Approximation, Peter Lang Verlag, Frankfurt a. M., 1986.

[17] Jongen, H.Th., Jonker, P. and Twilt, F., Nonlinear optimization in $\mathbb{R}^n$. II: transversality, flows, parametric aspects, Peter Lang Verlag, Frankfurt a. M., 1986.

[18] Kojima, M. and Hirabayashi, R., Continuation and deformation of nonlinear programs, Math. Program. Study 21, 1984, 150-198.

[19] Pateva, D., On the singularities in linear one-parametric optimization problems, Optimization 22, 1991, 193-220.

[20] Poore, A.B. and Al-Hassan, Q., The expanded Lagrangian system for constrained optimization problems, in SIAM J. Control Opt., Vol. 26 No. 2 (1988), pp. 417-27

[21] Poore, A.B. and Al-Hassan, Q., Bifurcation analysis for singularities on a tangent space for quadratic penalty-barrier and multiplier methods for solving constrained optimization problems, part I, in J. Math. Anal. Appl. 197 (1996), pp. 658-78

[22] Poore, A.B. and Al-Hassan, Q., Analysis of bifurcation due to loss of linear indepence and strict complementarity for penalty methods for solving constrained optimization problems, II, in J. Math. Anal. Appl. 201 (1996), pp. 756-85

[23] Poore, A.B. and Thiart C.A., Bifurcation problems in nonlinear parametric programming, in Mathematical Programming 39 (1987), pp.189-205

[24] Poore, A.B. and Thiart C.A., A bifurcation analysis of the nonlinear parametric programming problem, in Mathematical Programming 47 (1990), pp.117-141

[25] Rückmann, J. and Tammer K., On linear-quadratic perturbations in one-parametric nonlinear optimization. Systems Science 18 (1992) 1, 37-48.

[26] Tammer, K., Parametric complementarity problems. Preprint 96-12, Humboldt-University Berlin.

[27] Lu, Y. C., Singularity Theory and an Introduction to Catastrophe Theory, Universitext, Springerverlag, Berlin, Heidelberg, New York, 1976.