On the Role of the Mangasarian-Fromovitz Constraint Qualification for Penalty-, Exact Penalty- and Lagrange Multiplier Methods*

J. Guddat (Berlin), F. Guerra (Puebla), D. Nowack (Berlin)

Abstract

In this paper we consider three embeddings (one-parametric optimization problems) motivated by penalty, exact penalty and Lagrange multiplier methods. We give an answer to the question under which conditions these methods are successful with an arbitrarily chosen starting point. Using the theory of one-parametric optimization (the local structure of the set of stationary points and of the set of generalized critical points, singularities, structural stability, pathfollowing and jumps) the so-called Mangasarian-Fromovitz condition and its extension play an important role. The analysis shows us that the class of optimization problems for which we can surely find a stationary point using a pathfollowing procedure for the modified penalty and exact penalty embedding is much larger than the class where the Lagrange multiplier embedding is successful. For the first class, the objective may be a "really non-convex" function, but for the second one we are restricted to convex optimization problems. This fact was a surprise at least for the authors.

1 Introduction

We consider a nonlinear optimization problem with the following standard structure:

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

where

$$M := \{x \in I\!R^n \mid h_i(x) = 0,\ i \in I,\ g_j(x) \leq 0,\ j \in J\}, \tag{1.0}$$

$I := \{1, \ldots, m\},\ m < n,\ J := \{1, \ldots, s\}$, and $f, h_i, g_j \in C^2(I\!R^n, I\!R),\ i \in I,\ j \in J.$

For some of the results presented we need a higher degree of differentiability. We recall the well-known concept of embedding (cf. e.g. [5, 8, 13]). Consider a one-parametric optimization problem

*This research was supported by the Deutsche Forschungsgemeinschaft under Grants Gu304/1-4

$$\tilde{P}(t) \quad \min\{f(y,t) \mid y \in \tilde{M}(t)\},\ t \in [0,1], \tag{1.1}$$

where

$$\tilde{M}(t) := \{y \in I\!R^{\bar{n}} \mid h_i(y,t) = 0,\ i \in I,\ g_j(y,t) \le 0,\ j \in \bar{J}\}$$

$n \le \bar{n}$, $\bar{J}$ is a finite index set with $J \subseteq \bar{J}$, with at least the following properties:

(A1) A stationary point of $\tilde{P}(0)$ is known (and the corresponding Lagrange multipliers are known or easy to compute);

(A2) $\tilde{P}(t)$ has a global minimizer for every $t \in [0,1]$;

(A3) $\tilde{P}(1)$ is equivalent to (P) in a certain sense (to be specified below).

In this paper we consider three embeddings (cf. [6, 2, 3]) motivated by penalty, exact penalty and Lagrange multiplier methods and we ask the following question: Are there conditions for finding a discretization of the interval $[0,1]$:

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

and corresponding stationary points $y(t_i)$ of $\tilde{P}(t_i)$, $i = 1,\ldots,N$ starting with an arbitrarily chosen stationary point $y(t_0)$. We will see that the so-called Mangasarian-Fromovitz Constraint Qualification and its extension play an important role. First we introduce the following modified penalty embedding

$$P_1(t) : \min\{F_1(x,v,w,t) \mid (x,v,w) \in M_1(t)\},\ t \in [0,1]$$

where

$$\begin{aligned} F_1(x,v,w,t) &:= F_{(A,B,C,x^0,v^0,w^0)}(x,v,w,t) = \\ &= tf(x) + (1-t)(x-x^0)^T A(x-x^0) + (v-v^0)^T B(v-v^0) + (w-w^0)^T C(w-w^0) \end{aligned}$$

$$\begin{aligned} M_1(t) := M_{(x^0,v^0,w^0,w^1,b,p)} = \{(x,v,w) \in I\!R^n \times I\!R^m \times I\!R^s \mid\ & th_i(x) + (1-t)(v_i - v_i^0) = 0, i \in I, \\ & tg_j(x) + (1-t)(w_j - w_j^1) \le 0, j \in J \\ & \|x - x^0\|^2 + b^T x - p \le 0 \\ & \|v - v^0\|^2 + \|w - w^0\|^2 - q \le 0\}. \end{aligned}$$

In distinction to $\tilde{P}_4(t)$ in [6] we consider two points w^0 and w^1 in a suitable manner. Then the starting point (x^0, v^0, w^0) has better properties (cf. Theorem 1.1 below).

In [6] it is explained how we obtain such a kind of model for the penalty method. Let us consider this here only briefly. The model for the original quadratic penalty method reads

$$\min\{f(x) + \Big(\frac{t}{1-t}\Big)^2 \Big(\sum_{i\in I} (h_i(x)\Big)^2 + \sum_{j\in J} (\max\{g_j(x), 0\})^2 \mid x \in I\!R^n\}, t \in [0,1).$$

We observe that the penalty parameter $c(t) := \dfrac{t}{1-t}$ tends to $+\infty$ if t tends to 1. This one-parametric optimization problem has the following disadvantages: The problem is not defined for $t = 1$, the objective function is exactly once continuously differentiable (i.e., the results of parametric optimization presented in [8, 9, 10, 11, 12, 13, 15, 16, 7] – a short summary is given in Chapter 2 – are not applicable); we do not know any starting point for $t = 0$. It is easy to see that these disadvantages will not appear for $P_1(t)$. Moreover, there are further important properties of $P_1(t)$ (cf. Theorem 1.1). The term $(1-t)(x-x^0)^T A(x-x^0) + (v-v^0)^T B(v-v^0) + (w-w^0)^T C(w-w^0)$ is a regularization term (cf. Theorem 3.5(i)) which does not have any influence on the solution of (P) ((P) is equivalent to $P_1(1)$ (cf. Theorem 1.1(iii)).

Now we assume that

(B1) the feasible set M of the original problem (P) is non-empty and compact. We denote $E(p) := \{x \in I\!R^n \mid \|x - x^0\|^2 + b^T x - p \leq 0\}$. Then, for a fixed x^0 there exists a $\tilde{p} > 0$ such that $M \subseteq E(p)$ for all $p \geq \tilde{p}$. Furthermore, there exists a $\tilde{q} > 0$ such that $M_1(t) \neq \emptyset$ for all $t \in [0,1]$ and all $p \geq \tilde{p}$ and $q > \tilde{q}$.

We assume that

(B2) $p > \tilde{p}$ and $q > \tilde{q}$.

Now it is easy to see:

Theorem 1.1 *$P_1(t)$ has the following properties:*

(i) *Choosing $w^0, w^1 \in I\!R^s$ with $w_j^0 < w_j^1, j \in J$, $x^0, b \in I\!R^n$, $v^0 \in I\!R^m$, A, B and C positive definite, $p > 0$ and $q > 0$ sufficiently large, then (x^0, v^0, w^0) is the global minimizer and the only stationary point for $P_1(0)$. Furthermore, $(x^0, v^0, w^0, 0)$ is a non-degenerate stationary point (non-degenerate cf. the conditions (2.4)).*

(ii) *Assume that (B1) and (B2) are satisfied. Then $\psi_{\text{glob}}(t) \neq \emptyset$ for all $t \in [0,1]$, where $\psi_{\text{glob}}(t)$ is the set of all global minimizers for $P_1(t)$.*

(iii) *$P_1(t)$ is equivalent to (P) in the following sense:*

  (a) *If (x, v, w) is a stationary point for $P_1(1)$, then x is a stationary point for (P).*

  (b) *If x is a stationary point for (P), then there exist vectors $v \in I\!R^m, w \in I\!R^s$, such that (x, v, w) is a stationary point for $P_1(1)$.* □

Secondly, we consider the so-called exact penalty methods proposed e.g. in [14] for problems without equality constraints (i.e., $I = \emptyset$). We propose the following embedding (cf. $M_5(t)$ in [2])

$$P_2(t): \ \min\{F_2(x, w, t) \mid (x, w) \in M_2(t)\}, \ t \in [0,1]$$

where

$$\begin{aligned} F_2(x,w,t) \ &:= \ F_{(A,B,x^0,w^0,w^1,d)}(x,w,t) = tf(x) + (1-t)\sum_{j \in J} \mu_j(x)(w_j - w_j^0) + \\ &\quad +(1-t)\ (x-x^0)^T A(x-x^0) + (w-w^0)^T[\operatorname{diag}(\alpha_j - g_j(x))]^{-1} B(w-w^0) \end{aligned}$$

$\alpha^T = (\alpha_1, \ldots, \alpha_s)$ with $\alpha_j > g_j(x^0),\ j = 1, \ldots, s,$
$\mu(x) = -[Dg(x)Dg(x)^T + \gamma^2 G(x) \cdot G(x)]^{-1} Dg(x) Df(x)^T$
$g(x) = (g_1(x), \ldots, g_s(x)), Dg(x)$ is the Jacobian $, \gamma \neq 0,$
$G(x) =$ diag $(g_j(x)),$

$$M_2(t) := M_{(x^0,w^0,w^1,b,p,q)}(t) := \{(x,w) \in I\!R^n \times I\!R^s \mid tg_j(x) - (1-t)(\omega_j - w_j^1) \le 0, j \in J, \\ \|x - x^0\|^2 + b^T x - p \le 0, \\ \|w - w^0\|^2 - q \le 0\}$$

Then we have

Theorem 1.2 *$P_2(t)$ has the following properties*

(i) *Choosing $w^0, w^1 \in I\!R^s$ with $w_j^1 < w_j^0$, $j \in J$, $x^0, b \in I\!R^n$, $\alpha \in I\!R^s$ with $\alpha_j > g_j(x^0)$, $j \in J$, $p > 0$ and $q > 0$ sufficiently large, the matrices A and B positive definite, then (x^0, w^0) is the global minimizer and the only stationary point for $P_2(0)$. Furthermore, (x^0, w^0) is a non-degenerate stationary point.*

(ii) *Analogously to Theorem* 1.1(ii).

(iii) *Analogously to Theorem* 1.1(iii). □

Thirdly, we introduce the following Lagrange multiplier embedding (cf. [3]), for the original method cf. e.g. [1].

$P_3(t):\ \min\{F_3(x,v,w,\lambda,\mu,t) \mid (x,v,w,\lambda,\mu) \in M_3(t)\}, t \in [0,1],$

where

$$F_3(x,v,w,\lambda,\mu,t) := F_{(A,B,C,D,E,x^0,\lambda^0,\mu^0,v^0,w^0)}(x,v,w,\lambda,\mu) = \\ = t[f(x) + \sum_{i \in I} \lambda_i h_i(x,t) + \sum_{j \in J} \mu_j g_j(x,t)] + (1-t)[(x-x^0)^T A (x-x^0) - \\ -(\lambda-\lambda^0)^T B(\lambda-\lambda^0) - (\mu-\mu^0)^T C(\mu-\mu^0)] + (v-v^0)^T D(v-v^0) + \\ +(w-w^0)^T E(w-w^0)$$

$$M_3(t) := M_{(x^0,v^0,w^0,w^1,\lambda^0,\mu^0,b,c,d,p,q)}(t) = \{(x,\lambda,\mu,v,w) \in I\!R^{n+2m+2s} \mid \\ th_i(x) + (1-t)(v_i - v_i^0) = 0, i \in I, \\ tg_j(x) + (1-t)(w_j - w_j^1) \le 0, j \in J, \\ -\mu_j \le 0, j \in J \\ \|x - x^0\|^2 + b^T x - p \le 0 \\ \|\lambda - \lambda^0\|^2 + c^T\lambda + \|\mu - \mu^0\|^2 + d^T\mu + \|v - v^0\|^2 + \|w - w^0\|^2 - q \le 0\},$$

$$h_i(x,t) := th_i(x) + (1-t)(v_i - v_i^0),\ i \in I,$$
$$g_j(x,t) := tg_j(x) + (1-t)(w_j - w_j^1),\ j \in J.$$

We note that the Lagrange multiplier method follows a curve of $\min\limits_{x,v,w}$-$\max\limits_{\lambda,\mu}$-points that are saddle points in the set of stationary points for $P_3(t)$.

Then, it is easy to see:

Theorem 1.3 *$P_3(t)$ has the following properties*

(i) *Choosing $w^0, w^1 \in I\!R^s$ with $w_j^0 < w_j^1, j \in J$, $\mu^0 \in I\!R^s$ with $\mu_j^0 > 0, j \in J$, $x^0 \in I\!R^n$, $v^0, \lambda^0, c \in I\!R^m$, $\mu^0, d \in I\!R^s$, the matrices A, B, C, D, E positive definite, $p > 0$, $q > 0$ sufficiently large, then the point $(x^0, v^0, w^0, \lambda^0, \mu^0)$ is a global $\underset{x,v,w}{\min}$-$\underset{\lambda,\mu}{\max}$-point for $P_3(0)$. Furthermore, $(x^0, v^0, w^0, \lambda^0, \mu^0)$ is a non-degenerate stationary point.*

(ii) *Analogously to Theorem* 1.1(ii).

(iii) *Analogously to Theorem* 1.1(iii). □

Now we have the following important

Remark 1.4 We observe that there is the following strict difference with respect to the starting situation between the penalty and the exact penalty embedding on the one hand, and the Lagrange multiplier embedding on the other hand: For the first two embeddings the known starting points are the only stationary points. For the Lagrange multiplier method at least one global minimizer $(x^1, w^1, \lambda^1, \mu^1)$ will appear, i.e., we have at least two stationary points for $P_i(0), i = 1, 2$. Therefore, we will show that the class of optimization problems for which the penalty and the exact penalty method are surely successful is much larger than the class where the Lagrange multiplier embedding is successful. □

2 Theoretical background

We present a reduced version of Chapter 2 in [2].
We consider the one-parametric problem with some additional information that is important for our investigation:

$$P(t) \qquad \min\{f(x,t) | x \in M(t)\}, \tag{2.1}$$

where $t \in I\!R$, $M(t) = \{x \in I\!R^n \mid h_i(x,t) = 0, \ i \in I, \ g_j(x,t) \le 0, \ j \in J\}$, and $f, h_i, g_j \in C^k(I\!R^n \times I\!R, I\!R), \ i \in I, j \in J, k \ge 2$.
Furthermore, we introduce the following notations:

$$\begin{aligned}
\Sigma_{gc} &:= \{(x,t) \in I\!R^n \times I\!R \mid x \text{ is a g.c. point of } P(t)\},\\
\Sigma_{\text{stat}} &:= \{(x,t) \in I\!R^n \times I\!R \mid x \text{ is a stationary point of } P(t)\},\\
\Sigma_{\text{loc}} &:= \{(x,t) \in I\!R^n \times I\!R \mid x \text{ is a local minimizer of } P(t)\},\\
H &:= (h_1, \ldots, h_m)^T, \quad G := (g_1, \ldots, g_s)^T.
\end{aligned}$$

The Linear Independence Constraint Qualification (briefly LICQ) is satisfied at $\bar{x} \in M(\bar{t})$ if the vectors $D_x h_i(\bar{x}, \bar{t}), i \in I, D_x g_j(\bar{x}, \bar{t}), j \in J_0(\bar{x}, \bar{t})$, are linearly independent ($J_0(x,t) := \{j \in J \mid g_j(x,t) = 0\}$).
The Mangasarian-Fromovitz Constraint Qualification (briefly MFCQ) is satisfied at $\bar{x} \in M(\bar{t})$ if:

(MF1) $D_x h_i(\bar{x}, \bar{t}), i \in I$, are linearly independent, (MF2) there exists a vector $\xi \in R^n$ with

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

The KKT-system for a given problem $P(t)$ is fulfilled at a point $(\bar{x},\bar{t})$ if there exists a point $\bar{y} \in \mathbb{R}^{m+s}$ such that $\mathcal{H}(\bar{x},\bar{y},\bar{t}) = 0$, where $\mathcal{H} : \mathbb{R}^{n+m+s+1} \to \mathbb{R}^{n+m+s}$ is defined by

$$\mathcal{H}(x,y,t) = \left\{ \begin{array}{l} D_x f(x,t) + \sum\limits_{i\in I} y_i D_x h_i(x,t) + \sum\limits_{j\in J} y^+_{m+j} D_x g_j(x,t) \\ -h_i(x,t), i \in I \\ y^-_{m+j} - g_j(x,t), j \in J \end{array} \right\} \tag{2.2}$$

(for $\alpha \in \mathbb{R}$ let $\alpha^+ = \max\{\alpha, 0\}$ and $\alpha^- = \min\{\alpha, 0\}$). Obviously, the so-called Kojima-mapping $\mathcal{H}$ in (2.2) is piecewise continuously differentiable. In [13] the classical definition of a regular value of a continuously differentiable function is generalized for piecewise continuously differentiable functions. Furthermore, it is shown that if $0 \in \mathbb{R}^{n+m+s}$ is a regular value of $\mathcal{H}$, then the set $\mathcal{H}^{-1}(0)$ is a piecewise one-dimensional C^1-manifold.

Next, we cite our short characterization from [2] of the class $\mathcal{F}$ introduced by Jongen, Jonker and Twilt ([11, 12]). In [12] the local structure of Σ_{gc} is completely described if (f, H, G) belongs to a C^3_s-open and dense subset $\mathcal{F}$ of $C^3(\mathbb{R}^n \times \mathbb{R}, \mathbb{R})^{1+m+s}$, where C^3_s denotes the strong (or Whitney-) C^3-topology.
If $(f, H, G) \in \mathcal{F}$, then Σ_{gc} can be divided into 5 types.

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

$$\Big(D_x f + \sum_{i\in I} \lambda_i D_x h_i + \sum_{j\in J_0(\bar{z})} \mu_j D_x g_j\Big)|_{z=\bar{z}} = 0, \tag{2.3}$$

LICQ is satisfied at $\bar{x} \in \mathbf{M}(\bar{t})$, (2.4a)

$\mu_j \neq 0, \quad j \in J_0(\bar{z})$, (2.4b)

$D^2_x L(\bar{x})|_{T(\bar{z})}$ is nonsingular, (2.4c)

where $D^2_x L$ is the Hessian of the Lagrangian

$$L = f + \sum_{i\in I} \lambda_i h_i + \sum_{j\in J_0(z)} \mu_j g_j,$$

and the uniquely determined numbers λ_i, μ_j are taken from (2.3). 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. $D^2_x L(z)|_{T(z)}$ represents $V^T D^2_x L V$, where V is a matrix whose columns form a basis of $T(z)$.
A point of Type 1 is called a nondegenerate critical point. The set Σ_{gc} is the closure of the set of all points of Type 1, the points of the Types 2–5 constitute a discrete subset of Σ_{gc}. The points of the Types 2–5 represent three basic degeneracies:

Type 2 – violation of (2.4b)

The full curve stands for the curve of stationary points $z = (x,t)$, and the dotted curve represents the curve of g.c. points

Figure 2.1

Type 3 – violation of (2.4c)

Type 4 – violation of (2.4a) and $|I| + |J_0(\bar{z})| - 1 < n$

Type 5 – violation of (2.4a) and $|I| + |J_0(\bar{z})| = n + 1$.

For each of these five types Figure 2.1 illustrates the local structure of Σ_{gc}. Let $\Sigma_{\mathrm{gc}}^{\nu}$, $\nu \in \{1, \ldots, 5\}$ be the set of g.c. points of Type ν. Figure 2.2 illustrates the local structure of $\mathcal{F}$ in Σ_{loc} and Σ_{stat}. The class $\mathcal{F}$ is defined by

The full curve stands for a curve of local minimizers and the dotted curve in Fig. 2.2(c), (d), (e), (f) represents a curve of stationary points. The dotted curve in Fig. 2.2(g), (h) stands for a curve of stationary points in case $J_0(\bar{x}, \bar{t}) = \emptyset$.

Figure 2.2

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

The following theorem provides a special perturbation of (f, H, G) with additional parameters that can be chosen arbitrarily small such that the perturbed function vector belongs to the class $\mathcal{F}$.

Theorem 2.1 ([8]). *Let* $(f, H, G) \in C^3(\mathbb{R}^n \times R, R)^{1+m+s}$. *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*

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

Here "almost all" means: each measurable subset of

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

has the Lebesgue-measure zero. $\square$

Remark 2.2 Considering $\Sigma_{\rm stat}$ we note that the condition $(f, H, G) \in \mathcal{F}$ implies that zero is a regular value of the Kojima-mapping $\mathcal{H}$. $\square$

Definition 2.3 Let $K \subseteq I\!R \cup \{\pm\infty\}$.

(i) The problem $P(t)$ is called regular in the sense of Jongen-Jonker-Twilt – briefly JJT-regular – (with respect to K) if $(f, H, G) \in \mathcal{F}\Big((I\!R^n \times K) \cap \Sigma_{\text{gc}} \subset \bigcup_{\nu=1}^{5} \Sigma^{\nu}_{\text{gc}}\Big)$.

(ii) The problem $P(t)$ is called regular in the sense of Kojima-Hirabayashi – briefly KH-regular – (with respect to K) if $0 \in I\!R^{n+m+s}$ is a regular value of $\mathcal{H}$ $(\mathcal{H}|_{I\!R^n \times I\!R^m \times I\!R^s \times K})$.

Theorem 2.4 ([13]). *Let* $(f, H, G) \in C^3(I\!R^n \times I\!R, I\!R)^{1+m+s}$. *Then, for almost all* $(b, c, d) \in I\!R^n \times I\!R^m \times I\!R^s$, *the problem*

$$P_{(b,c,d)}(t) \quad \min\{f(x,t) + b^T x \;\Big|\; \begin{array}{l} h_i(x,t) + c_j = 0, \;\; i \in I, \\ g_j(x,t) + d_j \leq 0, \;\; j \in J \end{array} \Big\}$$

is KH-regular. $\square$

Now, we present some facts about pathfollowing methods (for more details see [8]). For this, we assume (A1),(A2) and the JJT-regularity of the considered problem $P(t)$.
The algorithm PATH III – the corresponding computer program is called PAFO – (cf. [8, 7, 16]) computes a numerical description of a compact connected component in Σ_{gc} and Σ_{stat}, respectively.

In the last part of Chapter 2 we present two theorems that are essential for our analysis.

Theorem 2.5 ([5]). *We assume*

(C1) *$M(t)$ is non-empty and there exists a compact set C with $M(t) \subseteq C$ for all $t \in [0,1]$.*

(C2) *$P(t)$ is KH-regular with respect to* $[0,1]$.

(C3) *There exists a $t_1 > 0$ and a continuous function $x : [0, t_1) \to I\!R^n$ such that $x(t)$ is the unique stationary point for $P(t)$ for $t \in [0, t_1)$.*

(C4) *MFCQ is satisfied for all $x \in M(t)$ for all $t \in [0,1]$.*

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

Applying Remark 2.2 we obtain

Corollary 2.6 *We assume* (C1), (C3), (C4) *and*

(D2) *$P(t)$ is JJT-regular with respect to* $[0,1]$.

Then there exists a PC^2-path $K(x^0, 0)$ in Σ_{stat} connecting $(x^0, 0)$ with some point $(x^*, 1)$. Furthermore $(x,t) \in K(x^0, 0)$ belongs to $\bigcup_{\nu \in \{1,2,3,5\}} \Sigma^{\nu}_{\text{gc}}$. $\square$

Remark 2.7 Now we have a look at Fig. 2.2. Since the MFCQ is satisfied, points of Type 5 in (j) and (k) are excluded. $\square$.

Finally we present a consequence of a general structural stability result given in [9]:

Theorem 2.8 ([9]). *We assume* (C1) *and* (C4). *Then $M(t_1)$ is homeomorphic with $M(t_2)$ for all $t_1, t_2 \in [0,1]$.* $\square$

3 On connecting curves

Considering Theorem 2.5, Corollary 2.6 and Theorem 2.8 we first ask for a condition on the original problem (P) ensuring that the MFCQ is satisfied for all $(x, v, w) \in M_1(t)$, $(x, w) \in M_2(t)$, and $(x, v, w, \lambda, \mu) \in M_3(t)$ for all $t \in [0, 1]$. We discuss the so-called Enlarged Mangasarian-Fromovitz Constraint Qualification (briefly EnMFCQ) introduced in 1985 in [5], and, simultaneously in [14] in a modified version, later on used in [8], [6], [2] and in further papers of di Pillo, Grippo et al.

Definition 3.1 Assume (B1) and $p \geq \tilde{p}$ (cf. (B2)). The EnMFCQ is satisfied if, for each $x \in E(p)$:

(a) $D_x h_i(x)$, $i \in I$, are linearly independent,

(b) there exists a vector $\xi \in I\!R^n$ with

$$\begin{aligned} h_i(x) + Dh_i(x)\xi = 0, &\quad i \in I, \\ g_j(x) + Dg_j(x)\xi < 0, &\quad j \in J_+(x) := \{j \in J \mid g_j(x) \geq 0\}, \\ (2x - 2x^0 + b)^T \xi < 0, &\quad \text{if } \|x - x^0\|^2 + b^T x = p. \end{aligned} \qquad \square$$

The answer to the question above was published for $M_i(t)$, $i \in \{1, 2\}$ in [6] and [2].

Theorem 3.2 *Assume* (B1), (B2), *and EnMFCQ. Then we have*

(i) $M_i(t) \neq \emptyset$, $M_i(t) \subseteq C$, $i \in \{1, 2, 3\}$ *for all* $t \in [0, 1]$, *where* C *is a compact set.*

(ii) *The MFCQ is satisfied for all* $(x, v, w) \in M_1(t)$, $(x, w) \in M_2(t)$ *and* $(x, v, w, \lambda, \mu) \in M_3(t)$ *for all* $t \in [0, 1]$.

Proof. (i) obvious.
(ii) 1) W.r.t. $M_1(t)$: The proof runs analog the main part of the proof of Theorem 4.2 in [5].
(ii) 2) W.r.t. $M_2(t)$: The proof is given in [2].
(ii) 3) W.r.t. $M_3(t)$: For $y^T := (x, v, w, \lambda, \mu) \in I\!R^{n+2m+2s}$ we consider

$$\begin{aligned} M_3(t) := \{y \in I\!R^{n+2m+2s} \mid\ & H_i(y, t) := t h_i(x) + (1 - t)(v - v_i^0) = 0,\ i \in I, \\ & (G_1)_j(y, t) := t g_j(x) + (1 - t)(w_j - w_j^1) \leq 0,\ j \in J, \\ & (G_2)_j(y, t) := -\mu_j \leq 0,\ j \in J, \\ & (G_3)(y, t) := \|\lambda - \lambda^0\|^2 + \|\mu - \mu^0\|^2 + \|v - v^0\|^2 + \|w - w^0\|^2 - q \leq 0, \\ & (G_4)(y, t) := \|x - x^0\|^2 + b^T x - p \leq 0\},\ \ t \in [0, 1]. \end{aligned}$$

We will discuss the three cases of $t = 0$, $t \in (0, 1)$ and $t = 1$ separately. For $y \in M_3(t)$ we define

$$\begin{aligned} (J_1)_0(y, t) &= \{j \in J \mid (G_1)_j(y, t) = 0\}, \\ (J_2)_0(y, t) &= \{j \in J \mid (G_2)_j(y, t) = 0\}. \end{aligned}$$

Obviously, $D_yH_i(y,t)$, $i \in I$, are linearly independent for all $t \in [0,1]$.

Case I: $t = 0$.

Let $y \in M(0)$. If $\|x - x^0\|^2 + b^T x = p$ choose $\xi \in I\!R^n$ from Definition 3.1(b). Otherwise take $\xi \in I\!R^n$ arbitrarily. We need a vector $(\eta^\lambda, \eta^\mu, \eta^v, \eta^w) \in I\!R^m \times I\!R^s \times I\!R^m \times I\!R^s$ such that $(\xi, \eta^\lambda, \eta^\mu, \eta^v, \eta^w)$ is a Mangasarian-Fromovitz-vector (briefly MF-vector) for $M_3(0)$. Take $\eta^\lambda = -2(\lambda - \lambda^0)$, $\eta_j^\mu = 1$, $j \in (J_2)_0(y,0)$, $\eta_j^\mu = 2(\mu_j - \mu_j^0)$, $j \in J \setminus (J_2)_0$, $\eta^v = 0$, $\eta_j^w = -1$, $j \in (J_1)_0(y,0)$, $\eta_j^w = -2(w_j - w_j^0)$, $j \in J \setminus (J_1)_0$.

Case II: $t \in (0,1)$.

Let $y \in M(t)$. Choose $\xi \in I\!R^n$ from Definition 3.1(b). Take $\eta^\lambda = -2(\lambda - \lambda^0)$, $\eta_j^\mu = 1$, $j \in (J_2)_0$, $\eta_j^\mu = -2(\mu_j - \mu_j^0)$, $j \in J \setminus (J_2)_0$,

$\eta_i^v = \frac{t}{t-1} Dh_i(x)\xi$, $i \in I$,

$\eta_j^w \in \left(0, \frac{t}{t-1} Dg_j(x)\xi\right)$, $j \in (J_1)_0 \cap J_+(x)$,

$\eta_j^w < 0$, $j \in (J_1)_0 \setminus J_+(x)$ if $Dg_j(x)\xi < 0$,

$\eta_j^w < \frac{t}{t-1} Dg_j(x)\xi$, $j \in (J_1)_0 \setminus J_+(x)$ if $Dg_j(x)\xi \geq 0$,

$\eta_j^w = -2(w_j - w_j^0)$, $j \in J \setminus (J_1)_0$.

Then $(\xi, \eta^\lambda, \eta^\mu, \eta^v, \eta^w)$ is an MF-vector for $M(t)$.

Case III: $t = 1$.

Let $y \in M(1)$. Choose $\xi \in I\!R^n$ from Definition 3.1(b). Take $\eta^\lambda = -2(\lambda - \lambda^0)$, $\eta_j^\mu = 1$, $j \in (J_2)_0(y,1)$,

$\eta_j^\mu = -2(\mu_j - \mu_j^0)$, $j \in J \setminus (J_2)_0(y,1)$, $\eta^v = -2(v - v^0)$, $\eta^w = -2(w - w^0)$.

Then $(\xi, \eta^\lambda, \eta^\mu, \eta^v, \eta^w)$ is an MF-vector for $M(t)$. □

Using Theorem 2.8 we obtain

Corollary 3.3 *Let the EnMFCQ be fulfilled. Then, for $i \in \{1,2,3\}$, $M_i(t_1)$ is homeomorphic with $M_i(t_2)$ for all $t_1, t_2 \in [0,1]$.* □.

For $i \in \{1,2\}$ this result is already included in [6] and [2].

Remark 3.4 From Corollary 3.3 it follows, in particular, that $M_i(0)$ is homeomorphic with $M_i(1)$. Using the special structure of $M_i(0)$, the set $M_i(0)$ is a convex set for $i \in \{1,2,3\}$. This means that $M_i(1)$ is homeomorphic with a convex set and therefore, an unconnected feasible set M of the original problem is excluded for the considered embeddings. This show us how restrictive the EnMFCQ is, but restrictive to the constraint functions only. □

Now, we give an answer to the question whether there exists a "connecting curve between $t = 0$ and $t = 1$" for the three embeddings using Theorem 2.5. We remind the reader of Remark 1.4 which states that the assumption (C3) is not satisfied for the embedding $P_3(t)$. Hence, Theorem 2.5 is not applicable.

Theorem 3.5 *Assume*

(a) $f, h_i, g_j \in C^3(\mathbb{R}^n, \mathbb{R})$, $i \in I$, $j \in J$.

(b) (B1), (B2), *and EnMFCQ.*

(c) $P_1(t)$ *(and* $P_2(t)$*, respectively) is JJT-regular with respect to* $(0, 1]$.

(d) *Choosing* A, B, C (A, B) *positive definite and* $w^0, w^1 \in \mathbb{R}^s$ *with* $w_j^0 < w_j^1$, $j \in J$, $(w_j^1 < w_j^0$, $j \in J)$, $x^0, b \in \mathbb{R}^n$, $v^0 \in \mathbb{R}^m$, $p, q \in \mathbb{R}$ *with* $p > \tilde{p}$ *and* $q > \tilde{q}$ $(x^0, b \in \mathbb{R}^n$, $\alpha \in \mathbb{R}^s$ *with* $\alpha_j > g_j(x^0)$, $j \in J$, $p, q \in \mathbb{R}$ *with* $p > \tilde{p}$ *and* $q > \tilde{q})$.

Then there exists a P^2*-path in* Σ_{stat} *connecting* $(x^0, v^0, w^0, 0)$ $((x^0, w^0, 0))$ *with some point* $(\hat{x}, \hat{v}, \hat{w}, 1)$ $((\hat{x}, \hat{w}, 1))$*, where* $\hat{x}$ *is a stationary point of* (P) *and* $\Sigma_{\text{stat}}|_{[0,1]} = \bigcup_{\nu=1}^{3} \Sigma_{\text{gc}}^{\nu} \cap \Sigma_{\text{stat}}$ *for* $n > 1$ *(for* $n = 1$ *a point of Type 5 may appear).*

Remarks to the proof. We check the assumptions of Corollary 2.6: For (C1) see Theorem 1.1. Furthermore, (x^0, v^0, w^0) $((x^0, w^0))$ is a point of Type 1 and the only stationary point for $P_1(0)$ $(P_2(0))$. Then (C3) is satisfied and (D2) holds, too. Using Theorem 3.2 MFCQ is satisfied for all $(x, v, w) \in M_1(t)$ $((x, w) \in M_2(t))$ for all $t \in [0, 1]$, points of Type 4 cannot appear. Using the special structure of the two-parametric optimization problems, points of Type 5 cannot appear for $n > 1$. $\square$

In the following we discuss the assumption (c) of Theorem 3.5. Let $\mathcal{A} \subset \mathbb{R}^{\frac{1}{2}n(n+1)}(\mathcal{B} \subset \mathbb{R}^{\frac{1}{2}m(m+1)}, \mathcal{C} \subset \mathbb{R}^{\frac{1}{2}s(s+1)})$ be the set of all symmetric nonsingular $(n, n)((m, m), (s, s))$ matrices. Obviously, $\mathcal{A}(\mathcal{B}, \mathcal{C})$ is an open and dense subset of $\mathbb{R}^{\frac{1}{2}n(n+1)}(\mathbb{R}^{\frac{1}{2}m(m+1)}, \mathbb{R}^{\frac{1}{2}s(s+1)})$ and $\mathbb{R}^{\frac{1}{2}n(n+1)} \setminus \mathcal{A}$ $(\mathbb{R}^{\frac{1}{2}m(m+1)} \setminus \mathcal{B}, \mathbb{R}^{\frac{1}{2}s(s+1)} \setminus \mathcal{C})$ has the Lebesgue measure 0.

Theorem 3.6 *Assume that* $f, h_i, g_j \in C^3(\mathbb{R}^n, \mathbb{R})$, $i \in I$, $j \in J$, *and that* $P_i(1)$ *is JJT-regular (i.e.,* $P_i(t)$ *is JJT-regular with respect to* $K = \{1\}$*),* $i \in \{1, 2\}$.

(i) *(with respect to* $P_1(t)$*). Then,* $P_1(t)$ *is JJT-regular with respect to* $[0, 1]$ *for almost all* $w^0, w^1 \in \mathbb{R}^s$ *with* $w_j^0 < w_j^1$, $j \in J$, $x^0, b \in \mathbb{R}^n$, $v^0 \in \mathbb{R}^m$, $p, q \in \mathbb{R}$ *with* $p > \tilde{p}$ *and* $q > \tilde{q}$, $A \in \mathcal{A}$, $B \in \mathcal{B}$, $C \in \mathcal{C}$.

(ii) *(with respect to* $P_2(t)$*). Then,* $P_2(t)$ *is JJT-regular with respect to* $[0, 1]$ *for almost* $w^0, w^1 \in \mathbb{R}^s$ *with* $w_j^1 < w_j^0$, $j \in J$, $x^0, b \in \mathbb{R}^n$, $\alpha \in \mathbb{R}^s$ *with* $\alpha_j > g_j(x^0)$, $j \in J$, $p, q \in \mathbb{R}$ *with* $p > \tilde{p}$ *and* $q > \tilde{q}$, $A \in \mathcal{A}$, $B \in \mathcal{B}$.

Remark to the proof. $P_i(t)$, $i \in \{1, 2\}$, is JJT-regular with $t \in [0, 1)$ using the special perturbations and we follow the explanation with respect to the proof of Theorem 5.1 in [6], cf. also [15]. The JJT-regularity for $P_i(1)$ is necessary for the extension to the closed interval $[0, 1]$. $\square$

Now we consider $P_1(t)$ with $A = I^n$, $B = I^m$, $C = I^s$, $b = 0$ and $P_2(t)$ with $A = I^n$, $B = I^s$, $b = 0$ and answer the same questions.

Corollary 3.7 *Assume*

(a) $f, h_i, g_j \in C^2(I\!R^n, I\!R)$, $i \in I$, $j \in J$.

(b) (B1), (B2), *and EnMFCQ.*

(c) $P_1(t)$ *with* $A = I^n$, $B = I^m$, $C = I^s$, $b = 0$ ($P_2(t)$ *with* $A = I^n$, $B = I^s$, $b = 0$, *respectively) is KH-regular with respect to* $(0, 1]$.

(d) *Choosing* $w^0, w^1 \in I\!R^s$ *with* $w_j^0 < w_j^1$, $j \in J$ $\quad (w_j^1 < w_j^0,\ j \in J)$, $x^0 \in I\!R^n$, $v^0 \in I\!R^m$, $p, q \in I\!R$ *with* $p > \tilde{p}$ *and* $q > \tilde{q}$ $\ (x^0 \in I\!R^n$, $\alpha \in I\!R^s$ *with* $\alpha_j > g_j(x^0)$, $j \in J$, $p, q \in I\!R$ *with* $p > \tilde{p}$ *and* $q > \tilde{q})$, *there exists a* PC^1*-path in* $\Sigma_{\rm stat}$ *connecting* $(x^0, v^0, w^0, 0)$ $((x^0, w^0, 0))$ *with some point* $(\hat{x}, \hat{v}, \hat{w}, 1)$ $\ ((\hat{x}, \hat{w}, 1))$, *where* $\hat{x}$ *is a stationary point of* (P). □

Corollary 3.8 *Assume that* $f, h_i, g_j \in C^j(I\!R^n, I\!R)$, $i \in I$, $j \in J$, *and that* $P_i(1)$ *is KH-regular (i.e.,* $P_i(t)$ *is KH-regular with respect to* $K = \{1\}$), $i \in \{1, 2\}$.

(i) (*with respect to* $P_1(t)$ *with* $A = I^n$, $B = I^m$, $C = I^s$, $b = 0$). *Then* $P_1(t)$ *is KH-regular with respect to* $[0, 1]$ *for almost all* $w^0, w^1 \in I\!R^s$ *with* $w_j^0 < w_j^1$, $j \in J$, $x^0 \in I\!R^n$, $v^0 \in I\!R^m$, $p, q \in I\!R$ *with* $p > \tilde{p}$ *and* $q > \tilde{q}$.

(ii) (*with respect to* $P_2(t)$ *with* $A = I^n$, $B = I^s$, $b = 0$). *Then* $P_2(t)$ *is KH-regular with respect to* $[0, 1]$ *for almost all* $w^0, w^1 \in I\!R^s$ *with* $w_j^1 < w_j^0$, $j \in J$, $x^0 \in I\!R^n$, $\alpha \in I\!R^s$ *with* $\alpha_j > g_j(x^0)$, $j \in J$, $p, q \in I\!R$ *with* $p > \tilde{p}$ *and* $q > \tilde{q}$. □

We observe that the EnMFCQ is a sufficient condition for a "connecting curve between $t = 0$ and $t = 1$". Now we ask for a necessary and sufficient condition. We know that for almost all inputs x^0, w^0, w^1, etc. (cf. Theorem 3.5) the starting point (x^0, v^0, w^0) for $P_1(0)$ and (x^0, w^0) for $P_2(0)$, respectively, lies on a uniquely determined connected component $K(x^0, v^0, w^0, 0)$ and $K(x^0, w^0, 0)$, respectively (briefly $K(y^0, 0)$ for both embeddings). Furthermore, we know that $K(y^0, 0)|_{[0,1]}$ is the only connected component in $\Sigma_{\rm stat}$ crossing the hyperplane $\{(y, t) \mid t = 0\}$. We introduce the following condition (cf. [3] for $P_2(t)$):

(F1) The MFCQ is satisfied for all $y \in M_i(t)$, $i \in \{1, 2\}$, with $(y, t) \in {\rm cl}\, K(y^0, 0)|_{[0,1]}$ (${\rm cl}\, K$ means the closure of K).

Theorem 3.9 (with respect to $P_1(t)$ (and $P_2(t)$, respectively)).
Assume (a), (c) *and* (d) *from Theorem* 3.5. *Then there exists a* PC^2*-path in* $\Sigma_{\rm stat}$ *connecting* $(x^0, v^0, w^0, 0)$ $\ ((x^0, w^0, 0))$ *with some point* $(\hat{x}, \hat{v}; \hat{w}, 1)$ $\ ((\hat{x}, \hat{w}, 1))$, *where* $\hat{x}$ *is a stationary point of* (P) *if and only if* (F1) *is satisfied.*

Remark to the proof. Use the same concept of proof as in Theorem 2.5.

Remark 3.10 This weakest condition for the existence of a "connecting curve between $t = 0$ and $t = 1$" is of theoretical nature. We cannot expect to check the condition for a given problem. Once more, the MFCQ plays an essential role. □

4 Illustrative examples

$(P) \quad \min\{f(x) \mid g(x) \leq 0\}$

Example 4.1

$$\begin{aligned} f(x) &= 0.0265073509x^8 - 0.211505207x^7 + 0.25753848x^6 + 1.34579642x^5 \\ &\quad -2.34222067x^4 - 2.65029635x^3 + 3.45664738x^2 + 0.91447716x + 5, \\ g(x) &= 10(x-2.5)^2 - 5. \end{aligned}$$

Fig. 4.1 shows $f(x)$ and $g(x)$.

Figure 4.1

Of course, the EnMFCQ is satisfied.

We have chosen the penalty embedding $P_1(t)$ with $A = I^n$, $C = I^s$, $b = 0$, $p = 25$, $q = 36$, $x^0 = -2$, $w^0 = 1$, $w^1 = 0$. Figures 4.2 and 4.3 show the curve of stationary points.
Note that all curves were computed by PAFO ([7, 16]). We see that the classical penalty method (with increasing t) is successful.

Fig. 4.2

Fig. 4.3

We have chosen the exact penalty embedding $P_2(t)$ with $A = I^n$, $B = I^s$, $\alpha = 200$, $p = 25$, $q = 36$, $x^0 = -2$, $w^0 = 1$, $w^1 = 0$. Fig. 4.4 and 4.5 show the curve of stationary points. We see that the classical exact penalty method (with increasing t) is successful.

Fig. 4.4

Fig. 4.5

We have chosen the Lagrange multiplier embedding $P_3(t)$ with $A = I^n$, $C = E = I^s$, $p = 25$, $q = 36$, $x^0 = -2$, $w^0 = 1$, $w^1 = 2$, $\mu^0 = 1$. The starting point (x^0, w^0, μ^0) is a global $\underset{x,w}{\min}$-$\underset{\mu}{\max}$-point for $P_3(0)$. The corresponding curve (cf. Fig. 4.6 – 4.8) does not attain $\{(x, w, \mu, t) \mid t = 1\}$ and returns to the local minimizer $(\bar{x}, \bar{w}, \bar{\mu}) = (-2, 1, 0)$ by using the pathfollowing procedure only. The example shows that the Lagrange multiplier method is not necessarily successful under EnMFCQ. We see that there is another connected component in Σ_{stat} where $(\tilde{x}, \tilde{w}, \tilde{\mu}, 0) = (-2, 1, 7, 0)$ belongs to. $(\tilde{x}, \tilde{w}, \tilde{\mu})$ is the global minimizer for $P_3(0)$. A jump from the first connected component to the second one is possible (cf. Remark 5.5). We observe that the pathfollowing procedure applied to $P_3(t)$ and the second connected component has nothing to do with the classical Lagrange multiplier method (following a curve of $\underset{x,v,w}{\min}$-$\underset{\lambda,\mu}{\max}$-points of $P_3(t)$).

Fig. 4.6

Fig. 4.7

Fig. 4.8

Choose the same embedding $P_3(t)$ and the following example

Example 4.2

$$\begin{aligned} f(x) &= 0.0265073509x^8 - 0.211505207x^7 + 0.25753848x^6 + 1.34579642x^5 \\ &\quad -2.34222067x^4 - 2.65029635x^3 + 3.45664738x^2 + 0.91447716x + 5, \\ g(x) &= 0.1x^6 - 0.3x^5 - 0.25x^4 + x^3 - 2x + 4, \\ p &= 25,\ q = 36,\ x^0 = -2,\ w^0 = 1,\ w^1 = 2,\ \mu^0 = 1. \end{aligned}$$

Fig. 4.9 shows $f(x)$ and $g(x)$.

Fig. 4.9

We see that the EnMFCQ is satisfied.

Fig. 4.10 shows the curve of stationary points. We have two points of Type 2 and four points of Type 3, and we attain $\{(x, w, \mu, t) \mid t = 1\}$.

Fig. 4.10

Example 4.2 illustrates that the Lagrange multiplier embedding could also be successful. Since we have four points of Type 3, we have to change the direction of t four times (increasing – decreasing etc.). Of course, the character of the stationary point will be changed at each turning point, i.e., in particular after the first turning point we are outside the classical Lagrange multiplier method from two points of view: t decreases and we follow a curve of stationary points of $P_3(t)$ that are not $\min\limits_{x,w}$-$\max\limits_{\mu}$-points.

5 Concluding remarks

Remark 5.1 We present arguments to the following thesis: "Penalty and exact penalty methods are successful for a larger class of optimization problems than the Lagrange multiplier method, more precisely: The Lagrange multiplier method is definitely successful for convex optimization problems only, provided there exists "a connecting curve between $t = 0$ and $t = 1$" for the penalty and exact penalty embedding under EnMFCQ (a condition to the constraint functions only) and certain regularity conditions."

(i) We observe (cf. Remark 1.4) that the starting point (x^0, v^0, w^0) for the penalty embedding and (x^0, w^0) for the exact penalty embedding, respectively, is the only stationary point for $P_1(0)$ and $P_2(0)$, respectively, and it is non-degenerate, i.e., the condition (C3) of Theorem 2.5 is satisfied. For the Lagrange multiplier embedding, the starting point $(x^0, w^0, \lambda^0, \mu^0)$ is a $\min\limits_{x,w}$-$\max\limits_{\lambda,\mu}$-point and a stationary point for $P_3(0)$, but not the only one, because there exists at least a global minimizer with respect to

(x, w, λ, μ) that is a stationary point, too. Considering Example 4.1 we have at least 3 stationary points for $P_3(0)$: The min-max-point $(-2, 1, 1)$ as the known starting point and a local minimizer $(-2, 1, 0)$, both are connected by a curve of stationary points (cf. Figures 4.6 – 4.8). Besides, there exists the global minimizer $(-2, 1, 7)$ for $P_3(0)$, which is a stationary point, too. Example 4.1 satisfies the EnMFCQ because the constraint function $g(x)$ is strictly convex. Figures 4.6 – 4.8 show that the curve connecting $(-2, 1, 1)$ and $(-2, 1, 0)$ returns to $\{(x, w, \mu, t) \mid t = 0\}$ at a turning point of Type 2 before the curve attains $\{(x, w, \mu, t) \mid t = 1\}$. The penalty and exact penalty method is successful (Example 4.1 illustrates Theorem 3.5).

(ii) The class for which the pathfollowing procedure definitely computes a stationary point for (P) using the penalty embedding $P_1(t)$, $t \in [0, 1]$, and the exact penalty embedding $P_2(t)$, $t \in [0, 1]$, is restricted by the EnMFCQ. The EnMFCQ is a condition to constraint functions only. The objective can be "really non-convex" (including functions with local minimizers and local maximizers that are non-degenerate). Such functions are excluded from the constraint functions.

(iii) Now we give an answer why the Lagrange multiplier method is surely successful for convex problems only. Suppose that (P) is a non-convex optimization problem, then we cannot exclude the appearance of turning points (Type 2 and Type 3). That means, the same situation as in Example 4.1 (Figures 4.6 – 4.8) is possible.

□

Remark 5.2 The classical penalty, exact penalty and Lagrange multiplier method (with increasing t) are not necessarily successful if there exists a "curve connecting $t = 0$ and $t = 1$". We know that turning points (Type 2 and Type 3) may appear. Then the original methods (for increasing t) stop at such a point. We have to go back. We do not have any problems using PAFO, but we leave the original method. In particular, the character of the stationary point will be changed, for instance: We follow a curve of local minimizers and a point of Type 3 appears, then we have to go back and follow a curve of saddle-points. Here, we refer to Remark 3.1 in [6].
Considering the Lagrange multiplier method which follows a curve of $\min_{x,v,w}\text{-}\max_{\lambda,\mu}$-points (cf. Example 4.2 and Fig. 4.10 where a turning point appears) the character of the stationary point will be changed. We have left the classical Lagrange multiplier method from two points of view: t decreases and we follow a curve of stationary points that are not $\min_{x,v,w}\text{-}\max_{\lambda,\mu}$-points. □

Remark 5.3 Assuming that we are not successful with the pathfollowing procedure only, then we refer to our investigation in [2], [3], [6], [8], [10] using pathfollowing methods in the set Σ_{gc} instead of the set Σ_{stat}. In case we are not successful either, we propose pathfollowing procedures with jumps from one connected component to another one in the set Σ_{loc} and Σ_{gc}. We refer to the algorithms JUMP I and JUMP II in [8]. A summary is presented in Chapter 2 in [6] and [3]. JUMP I follows a curve of local minimizers and jumps to another connected component by finding a direction of descent if points of Type 2 or Type 3 appear. If we pass a point $(\bar{x}, \bar{t})$ of Type 4, then the local minimizer turns into a local maximizer. Then we have two cases:

Case I: The value of $f(\cdot,\bar{t})$ decreases.

Case II: The value of $f(\cdot,\bar{t})$ increases.

In Case I it is possible to jump to another branch of local minimizers. In Case II the corresponding connected component of the feasible set becomes empty and we do not have a jump. We note that the MFCQ is violated at a point of Type 4, i.e., under EnMFCQ we are able to jump to another connected component in Σ_{loc} of $P_i(t)$, $i \in \{1,2\}$, if a point of Type 2 or 3 appears.
This is another possibility in order to reach the level $t = 1$ for increasing t,but the path-following procedure (with increasing and decreasing t) seems to be more efficient. With respect to the Lagrange multiplier embedding $P_3(t)$ we have to use the jumps described in [10] (for a short description cf. Chapter 2 of [3]). There, jumps are presented at points of Type i, $i \in \{2,3,4\}$, if we follow a curve of stationary points not being local minimizers. We have jumps for all types, even in points of Type 4. These jumps are applicable to the Lagrange multiplier embedding. We can jump to another curve of local minimizers with respect to (x, v, w, λ, μ). Fig. 5.1 illustrates this situation.

Fig. 5.1

— local $\min\limits_{x,v,w}$-$\max\limits_{\lambda,\mu}$-point

– – stationary point that is not a $\min\limits_{x,v,w}$-$\max\limits_{\lambda,\mu}$-point

– · – local minimizer with respect to (x, v, w, λ, μ)

In Fig. 5.1 we reach the level $t = 1$. We observe that we are outside the Lagrange multiplier method if we follow a branch of local minimizers. For more details we refer to Remark 3.1 in [3].
Considering $P_3(t)$ with $n > 1$, which is JJT-regular with respect to $[0,1]$, and assuming the EnMFCQ to be satisfied, we attain the level $t = 1$ by pathfollowing procedures with

jumps. Under EnMFCQ and $n > 1$ we have points of Type 2 and 3 only. Therefore, we also have jumps if a point of Type 2 or Type 3 appears on a branch of local minimizers.

Remark 5.4 Theorem 3.6 represents a necessary and sufficient condition that there exists a "connecting curve between $t = 0$ and $t = 1$" for the penalty embedding $P_1(t)$, $t \in [0, 1]$, and the exact penalty embedding $P_2(t)$, $t \in [0, 1]$, see also Corollary 3.7.

Remark 5.5 Corollary 3.3 and Remark 3.4 show how restrictive the EnMFCQ is with respect to the feasible set M of the original problem (P): An unconnected feasible set M is excluded for the considered embeddings.

References

[1] Bertsekas, D.P. (1982): Constrained optimization and Lagrange multiplier methods. Academic Press New York .

[2] Dentcheva, D., Gollmer, R., Guddat, J., Rückmann, J.-J. (1995): Pathfollowing methods in nonlinear optimization II: Exact penalty methods. In [4], 200-230.

[3] Dentcheva, D., Guddat, J., Rückmann, J.-J., Wendler, K. (1995): Pathfollowing methods in nonlinear optimization III: Lagrange multiplier embedding. ZOR 41, 127-152.

[4] Florenzano, M. et al. (eds.) (1995): Approximation and Optimization in the Caribbean II. In: Ser. approximation and optimization. Verlag Peter Lang, Frankfurt a.M., Berlin, Bern, New York, Paris, Wien.

[5] Gfrerer, H., Guddat, J., Wacker, Hj., Zulehner, W. (1985): Pathfollowing methods for Kuhn-Tucker curves by an active index set strategy. In: Bagchi, A., Jongen, H.Th. (eds.) Systems and optimization. Lecture Notes in Control and Information Sciences 66, Springer-Verlag Berlin, Heidelberg, New York, 111-131.

[6] Gollmer, R., Guddat, J., Guerra, F., Nowack, D., Rückmann, J.-J. (1993): Pathfollowing methods in nonlinear optimization I: Penalty embedding. In: Guddat, J. et al. (eds.) Parametric optimization and related topics III. In: Ser. approximation and optimization. Verlag Peter Lang, Frankfurt a.M., Berlin, Bern, New York, Paris, Wien.

[7] Gollmer, R., Kausmann, U., Nowack, D., Wendler, K. (1995): Computerprogramm PAFO, Humboldt-Universität Berlin, Institut für Mathematik.

[8] Guddat, J., Guerra, F., Jongen, H.Th. (1990): Parametric optimization: Singularities, pathfollowing and jumps. BG Teubner, Stuttgart and John Wiley, Chichester.

[9] Guddat, J., Jongen, H.Th., Rückmann, J.-J. (1986): On stability and stationary points in nonlinear optimization. J. Austral. Math. Soc., Ser. B 28, 36-56.

[10] Guddat, J., Rückmann, J.-J. (1994): One-parametric optimization: Jumps in the set of generalized critical points. Control and Cybernetics 22, 1/2.

[11] Jongen, H.Th., Jonker, P., Twilt, F. (1986): On one-parametric families of optimization problems. Equality constraints. JOTA 48, 141-161.

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

[13] Kojima, M., Hirabayashi, R. (1984): Continuous deformation of nonlinear programs. Math. Progr. Study 21, 150-198.

[14] Di Pillo, G., Grippo, L. (1986): An exact penalty function method with global convergence properties for nonlinear programming. Math. Programming, 36, 1-18.

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

[16] Wendler, K. (1993): Implementation of a pathfollowing procedure for solving nonlinear one-parametric optimization problems. In: Brosowski, B. et al. (eds.) Multicriteria decision. In: Ser. approximation and optimization. Verlag Peter Lang, Frankfurt a.M., Berlin, Bern, New York, Paris, Wien, 139-163.