MATHEMATICAL PROGRAMS WITH EQUILIBRIUM CONSTRAINTS: SOLUTION TECHNIQUES FROM PARAMETRIC OPTIMIZATION

Gemayqzel Bouza Allende

Composition of the Graduation Committee:

Prof.	Dr.	R.	Boucherie	Universiteit Twente
Prof.	Dr.	J.	Guddat	Humboldt Unversität
Prof.	Dr.	H.Th.	Jongen	RWTH Aachen
	Dr	W.	Kern	Universiteit Twente
Prof.	Dr.	J.J.	Rückmann	Universidad de Las Americas
	Dr.	O.	Stein	RWTH Aachen
	Dr.	G.	Still	Universiteit Twente
Prof.	Dr.	G.J.	Woeginger	TU Eindhoven&Universiteit Twente

MATHEMATICAL PROGRAMS WITH EQUILIBRIUM CONSTRAINTS: SOLUTION TECHNIQUES FROM PARAMETRIC OPTIMIZATION

PROEFSCHRIFT

ter verkrijging van
de graad van doctor aan de Universiteit Twente,
op gezag van de rector magnificus,
prof. dr. W.H.M. Zijm,
volgens besluit van het College voor Promoties
in het openbaar te verdedigen
op donderdag 1 juni 2006 om 13.15 uur

door

Gemayqzel Bouza Allende

geboren op 23 september, 1977
La Habana, Cuba

Dit proeschrift is goedgekeurd door

Prof.	Dr.	J.	Guddat	promotor
Prof.	Dr.	G.J.	Woeginger	promotor
	Dr.	G.	Still	assitent-promotor

Acknowledgements

First, I want to thank my supervisors Prof. Dr. Jürgen Guddat and Dr. Georg Still. Prof. Dr. Guddat introduced me in the world of Parametric Optimization and kindly accepted to supervise my diploma and master's thesis in Cuba. As supervisor he has always supported my work and has always tried to keep me in touch with the new results in the area. Scientific discussions with him were always an fruitful experience, since they made me know how research can be done. These years of work supervised by Dr. Still have been really pleasant. His ideas were really insightful. Despite all the fears you may have in the beginning of your Ph.D., he makes you think that Math can be simple. Even when discussing on tough topics, he makes you feel confident, relaxed and motivated. These facts helped me at work, because, from the very beginning, I did not feel embarrassed when he corrected my mistakes. Of course there were bad moments, but then I had also the personal support of both, looking for solutions, improving results and redaction and making things lighter.

I also want to thank Prof. Dr. Kees Hoede, who kindly volunteered to read the manuscript, although he knew it will be a hard task. His suggestions allowed me to improve the initial version of the thesis.

Although I did not spend to much time working directly in the University of Twente, I enjoyed my stays at the DWMP department. I want to thank Dini Heres, who always helped me with different procedures in a very efficient way.

I also want to thank my (former) professors, in Cuba. I am in debt with them for the knowledge they gave me during my student years. In this period, I had also the support (sometimes electronic) of Dr. Luis Ramiro Piñeiro, Dr. Vivian Sistachs and the members of the groups of Numeric Analysis and Optimization of U.H. (I will not mention the names because the list will be huge). Sometimes the words of Loretta, Roxana, Maribel Freyre, Dr. Josefina Martinez, Francis Fuster and friends from ITC-CF and BS-UT, here in Enschede, helped me to go ahead despite difficulties.

Finally I want to thank my family, who patiently heard (or read) all my difficulties and give me an special strength to continue, even when difficulties appeared. Specially thanks to my grandmother and my parents for putting love over their own wishes, advising me well and accepting my decisions.

Abstract

Equilibrium constrained problems form a special class of mathematical programs where the decision variables satisfy a finite number of constraints together with an equilibrium condition. Optimization problems with a variational inequality constraint, bilevel problems and semi-infinite programs can be seen as particular cases of equilibrium constrained problems. Such models appear in many practical applications.

Equilibrium constraint problems can be written in bilevel form with possibly a finite number of extra inequality constraints. This opens the way to solve these programs by applying the so-called Karush-Kuhn-Tucker approach. Here the lower level problem of the bilevel program is replaced by the Karush-Kuhn-Tucker condition, leading to a mathematical program with complementarity constraints (MPCC). Unfortunately, MPCC problems cannot be solved by classical algorithms since they do not satisfy the standard regularity conditions. To solve MPCCs one has tried to conceive appropriate modifications of standard methods. For example sequential quadratic programming, penalty algorithms, regularization and smoothing approaches.

The aim of this thesis is twofold. First, as a basis, MPCC problems will be investigated from a structural and generical viewpoint. We concentrate on a special parametric smoothing approach to solve these programs. The convergence behavior of this method is studied in detail. Although the smoothing approach is widely used, our results on existence of solutions and on the rate of convergence are new. We also derive (for the first time) genericity results for the set of minimizers (generalized critical points) for one-parametric MPCC.

In a second part we will consider the MPCC problem obtained by applying the KKT-approach to equilibrium constrained programs and bilevel problems. We will analyze the generic structure of the resulting MPCC programs and adapt the related smoothing method to these particular cases. All corresponding results are new.

Contents

Chapter 1

Introduction

1.1 Introduction

An Equilibrium Constrained optimization problem (EC) is a mathematical program such that a part of the variables should satisfy an equilibrium condition. In its simplest form, an equilibrium condition is given by the critical point equation $\nabla\phi(y) = 0$. So the simplest prototype of an EC is:

$$\begin{aligned} &\min f(x,y) \\ \text{s.t. } \nabla_y \phi(x,y) &= 0, \end{aligned} \tag{1.1.1}$$

where $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, $\phi : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ and $\nabla_y \phi$ denotes the partial derivatives of ϕ with respect to y. For the case that $\phi(x, y)$ is convex in y, it can be expressed equivalently as a so-called Bilevel problem (BL):

$$\begin{aligned} &\min f(x,y) \\ \text{s.t.} \quad & y \text{ solves } \min_{y \in \mathbb{R}^m} \phi(x,y). \end{aligned}$$

More generally, we are led to consider BL problems of the following form:

$$\begin{aligned} &\min f(x,y) \\ \text{s.t.} \qquad & (x,y) \in C, \\ & y \text{ solves } Q(x), \\ & Q(x): \quad \min_{y \in Y(x)} \phi(x,y), \end{aligned} \tag{1.1.2}$$

where $f, \phi : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ and $Y(x)$ is the feasible set the *lower level problems* $Q(x)$ depending on x.

Another classical example of BL arises when considering an equilibrium point in a 0-sum game with two players. Let us assume that player i may choose strategies y_i from the set $Y_i \subset \mathbb{R}^m$, $i = 1, 2$, and that the utility functions are $g(y_1, y_2)$ for player 1 and $-g(y_1, y_2)$ for player 2, if player i chooses strategy

$y_i,\ i = 1, 2$. The Nash equilibrium points $(\bar{y}_1, \bar{y}_2)$ are saddle points of $g(y_1, y_2)$, *i.e.*,

$$g(\bar{y}_1, y_2) \geq g(\bar{y}_1, \bar{y}_2) \geq g(y_1, \bar{y}_2), \quad \text{for all } y_1 \in Y_1, y_2 \in Y_2,$$

or

$$\bar{y}_1 \text{ solves } \max_{y_1 \in Y_1} g(y_1, \bar{y}_2) \quad \text{and} \quad \bar{y}_2 \text{ solves } \min_{y_2 \in Y_2} g(\bar{y}_1, y_2).$$

If we want to minimize a function $f(x, y_1, y_2)$ under the condition that (y_1, y_2) are Nash equilibrium points of the previous game, the model becomes a bilevel problem of the type:

$$\begin{aligned} & \min f(x, \bar{y}_1, \bar{y}_2) \\ \text{s.t.} \quad & \bar{y}_1 \text{ solves } \max_{y_1 \in Y_1} g(y_1, \bar{y}_2), \\ & \bar{y}_2 \text{ solves } \min_{y_2 \in Y_2} g(\bar{y}_1, y_2) \end{aligned}$$

with functions $f : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R},\ g : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$.
More generally we could consider bilevel problems of the form:

$$\begin{aligned} & \min f(x, \bar{y}_1, \dots, \bar{y}_k) \\ \text{s.t.} \quad & (x, \bar{y}_1,\ \bar{y}_k) \in C, \\ & \bar{y}_i \text{ solves } Q_i(x), \quad \text{where} \\ & Q_i(x) : \quad \min_{y_i \in Y_i(x)} \phi_i(x, \bar{y}_1, \dots, \bar{y}_{i-1}, y_i, \bar{y}_{i+1}, \dots, \bar{y}_k),\ i = 1, \dots, k. \end{aligned}$$

For simplicity, in this thesis we will only consider the case $k = 1$, *i.e.* bilevel problems of type (1.1.2).

Under smoothness and convexity conditions, a minimizer y of the problem $Q(x)$ in (1.1.2) satisfies the inequality $\nabla_y \phi(x, y)^T (z - y) \geq 0,\ \forall z \in Y(x)$. More generally we examine the so-called variational inequalities

$$VI : \quad \text{Find } y \in Y(x) \text{ such that } \Phi(x, y)^T (z - y) \geq 0,\ \forall z \in Y(x),$$

and we are led to optimization problems of the form

$$\begin{aligned} P_{VI} : \quad & \min f(x, y) \\ \text{s.t.} \quad & (x, y) \in C, \\ & y \in Y(x), \\ & \Phi(x, y)^T (z - y) \geq 0,\ \forall z \in Y(x), \end{aligned}$$

or more generally to

$$\begin{aligned} P_{EC} : \quad & \min f(x, y) \\ \text{s.t.} \quad & (x, y) \in C, \\ & y \in Y(x), \\ & \phi(x, y, z) \geq 0,\ \forall z \in Y(x), \end{aligned}$$

where $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R},\ \phi : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$.

Remark 1.1.1 *P_{EC} can be regarded as a particular case of a Generalized Semi-infinite Optimization Problem*

$$GSIP: \; \min \{f(x,y) \mid (x,y) \in C, \; \phi(x,y,z) \geq 0, \; \forall z \in Y(x)\}. \tag{1.1.3}$$

Note that P_{EC} contains the additional condition $y \in Y(x)$.

The problems considered so far will be called equilibrium constrained problems.

We emphasize that in general it is difficult to solve these problems. Due to the two level structure, even to check feasibility, we have to compute a (global) minimizer of a general optimization problem or a solution of a variational inequality problem.

In this thesis we will deal with the analytic and generic structure of problems P_{BL}, P_{VI}, and P_{EC} and we are also interested in numerical solution methods. For numerical purposes it is natural to assume that the involved sets $Y(x)$ and C are described analytically. Throughout the thesis we will assume that these sets are defined as:

$$\begin{aligned} Y(x) &= \{y \in \mathbb{R}^m \mid \; v_i(x,y) \geq 0, \quad i = 1, \dots, l\}, & (1.1.4)\\ C &= \{(x,y) \in \mathbb{R}^n \times \mathbb{R}^m \mid \; g_j(x,y) \geq 0, \;\; j = 1, \dots, q\}, & (1.1.5) \end{aligned}$$

with given functions v_i and g_j, $i = 1, \dots, l$, $j = 1, \dots, q$.

So we will consider bilevel problems of the form

$$\begin{array}{llc} P_{BL}: & & \min\limits_{x,y} f(x,y) \\ & \text{s.t.} & g_j(x,y) \geq 0, \; j = 1, \dots, q, \\ & & y \text{ solves } Q(x), \\ & & Q(x): \qquad \min\limits_{y} \; \phi(x,y) \\ & & \qquad \text{s.t.} \qquad y \in Y(x) \end{array} \tag{1.1.6}$$

with sets $Y(x)$ defined as in (1.1.4) and functions $f, \phi, v_i, g_j : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, $i = 1, \dots, l$, $j = 1, \dots, q$. The parametric problem $Q(x)$ will be called the lower level problem.

Problems with equilibrium constraints are examined in the form

$$\begin{array}{llrcll} P_{EC}: & & & \min\limits_{x,y} f(x,y) & & \\ & \text{s.t.} & g_j(x,y) & \geq & 0, & j = 1, \dots, q, \\ & & y & \in & Y(x), & \\ & & \phi(x,y,z) & \geq & 0, & \forall z \in Y(x), \end{array} \tag{1.1.7}$$

where $f, v_i, g_j : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, $i = 1, \dots, l$, $j = 1, \dots, q$, $\phi : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ and $Y(x)$ is defined as in (1.1.4).

As a particular case, in which the function ϕ satisfies the condition $\phi(x, y, y) = 0, \ \forall (x, y)$, we consider the problem with variational inequalities constraints

$$
\begin{array}{llrcll}
P_{VI}: & & \multicolumn{2}{c}{\min\limits_{x,y} f(x,y)} & & \\
& \text{s.t.} & g_j(x,y) & \geq & 0, & j = 1, \dots, q, \\
& & y & \in & Y(x), & \\
& & \Phi(x,y)^T(z-y) & \geq & 0, & \forall z \in Y(x).
\end{array}
\tag{1.1.8}
$$

Another related type of optimization problems are the so-called Mathematical Programs with Complementarity Constraints (MPCC)

$$
\begin{array}{llrcll}
P_{CC}: & & \multicolumn{2}{c}{\min f(x)} & & \\
& \text{s.t.} & g_j(x) & \geq & 0, & j = 1, \dots, q, \\
& & r_i(x) & \geq & 0, & i = 1, \dots, l, \\
& & s_i(x) & \geq & 0, & i = 1, \dots, l, \\
& & r_i(x)s_i(x) & = & 0, & i = 1, \dots, l.
\end{array}
\tag{1.1.9}
$$

As we shall see later on, we will obtain problems of this type if we apply the Karush Kuhn Tucker (KKT) approach to the problems P_{BL}, P_{VI} and P_{EC}. As a solution method for problems (1.1.9), in the present thesis we will investigate the so-called one-parametric smoothing approach. In this approach, we consider the perturbation of P_{CC} (1.1.9)

$$
\begin{array}{llrcll}
P_{\tau}: & & \multicolumn{2}{c}{\min f(x)} & & \\
& \text{s.t.} & g_j(x) & \geq & 0, & j = 1, \dots, q, \\
& & r_i(x) & \geq & 0, & i = 1, \dots, l, \\
& & s_i(x) & \geq & 0, & i = 1, \dots, l, \\
& & r_i(x)s_i(x) & = & \tau, & i = 1, \dots, l,
\end{array}
\tag{1.1.10}
$$

where $\tau > 0$ is the perturbation parameter. Then P_{τ} is solved for $\tau \to 0^+$.

1.2 Relations between the problems

We have already pointed out some connections between the problems P_{VI}, P_{EC}, P_{BL} and P_{CC}. The aim of this section is to analyze these relations further. We assume obvious differentiability conditions on the problem functions.

Let us consider the relations between P_{EC} and P_{BL} (see also Stein and Still [59]). By using the fact:

$$
\phi(x, y, z) \geq 0, \ \forall z \in Y(x) \Leftrightarrow z \in arg \min_{u \in Y(x)} \ \phi(x, y, u) \ \text{ and } \phi(x, y, z) \geq 0,
$$

the problem P_{EC} turns into the bilevel problem:

$$\begin{array}{rl} & \min\limits_{x,y,z} f(x,y) \\ \text{s.t.} & g_j(x,y) \geq 0, \quad j = 1,\dots,q, \\ & y \in Y(x), \\ & \phi(x,y,z) \geq 0, \\ & z \text{ solves } Q(x,y), \\ & Q(x,y): \qquad \begin{array}{rl} & \min \phi(x,y,u) \\ \text{s.t.} & u \in Y(x). \end{array} \end{array} \tag{1.2.1}$$

In the particular case of P_{EC} where $\phi(x,y,y) = 0, \;\; \forall y$ (*e.g.* the problems P_{VI}) we can eliminate the variable z as follows. In view of $\phi(x,y,y) = 0$, the condition $\phi(x,y,z) \geq 0, \; \forall z \in Y(x)$, is equivalent with the fact that y is a global solution of $Q(x,y)$. So, in this case, P_{EC} is equivalent with:

$$\begin{array}{rl} & \min\limits_{x,y} f(x,y) \\ \text{s.t.} & g_j(x,y) \geq 0, \; j = 1,\dots,q, \\ & y \in Y(x), \\ & y \text{ solves } Q(x,y), \\ & Q(x,y): \qquad \begin{array}{rl} & \min\limits_{z} \phi(x,y,z) \\ \text{s.t.} & z \in Y(x). \end{array} \end{array} \tag{1.2.2}$$

Note that this problem has a more complicated structure than P_{BL}, since it contains a sort of fixed point condition. Indeed, the lower level problem $Q(x,y)$ depends on y, which at the same time should solve $Q(x,y)$. For a comparison from the structural and generical viewpoint between linear P_{BL} and P_{EC}, we refer to Birbil, Bouza, Frenk and Still [7] and Still [61].

We now apply the KKT approach to the above bilevel problems. The idea is to replace the minimum condition for the lower level problem $Q(x)$, or $Q(x,y)$, by the KKT optimality condition. This will lead to problems of type P_{CC}.

We begin with the standard bilevel problem P_{BL} in (1.1.6). Let some constraint qualification such as MFCQ (see Definition 2.2.2) hold for the lower level problem $Q(x)$. Then a (local) minimizer of $Q(x)$ will necessarily satisfy the KKT conditions, *i.e.*, the feasible points (x,y) of P_{BL} will fulfill the following constraints with some multiplier vector $\lambda \in \mathbb{R}^l$:

$$\begin{aligned} \nabla_y \phi(x,y) - \sum_{i=1}^{l} \lambda_i \nabla_y v_i(x,y) &= 0, \\ v_i(x,y) &\geq 0, \quad i = 1,\dots,l, \\ \lambda_i &\geq 0, \quad i = 1,\dots,l, \\ v_i(x,y)\lambda_i &= 0, \quad i = 1,\dots,l. \end{aligned} \tag{1.2.3}$$

Consequently, as a relaxation of P_{BL}, we obtain:

$$\begin{array}{lrcl}
 & \min\limits_{x,y,\lambda} f(x,y) & & \\
\text{s.t.} & g_j(x,y) & \geq & 0, \quad j=1,\dots,q, \\
 & \nabla_y \phi(x,y) - \sum\limits_{i=1}^{l} \lambda_i \nabla_y v_i(x,y) & = & 0, \\
 & v_i(x,y) & \geq & 0, \quad i=1,\dots,l, \\
 & \lambda_i & \geq & 0, \quad i=1,\dots,l, \\
 & \lambda_i v_i(x,y) & = & 0, \quad i=1,\dots,l.
\end{array} \tag{1.2.4}$$

If we apply the same arguments to the bilevel form of P_{EC} (see (1.2.1)) we are led to the problem:

$$\begin{array}{lrcl}
 & \min\limits_{x,y,z,\lambda} f(x,y) & & \\
\text{s.t.} & g_j(x,y) & \geq & 0, \quad j=1,\dots,q, \\
 & v_i(x,y) & \geq & 0, \quad i=1,\dots,l, \\
 & \nabla_z \phi(x,y,z) - \sum\limits_{i=1}^{l} \lambda_i \nabla_z v_i(x,z) & = & 0, \\
 & v_i(x,z) & \geq & 0, \quad i=1,\dots,l, \\
 & \lambda_i & \geq & 0, \quad i=1,\dots,l, \\
 & \lambda_i v_i(x,z) & = & 0, \quad i=1,\dots,l, \\
 & \phi(x,y,z) & \geq & 0.
\end{array} \tag{1.2.5}$$

The KKT approach for the special case $\phi(x,y,y)=0,\ \forall y \in Y(x)$, *i.e.*, for (1.2.2) gives:

$$\begin{array}{lrcl}
 & \min\limits_{x,y,\lambda} f(x,y) & & \\
\text{s.t.} & g_j(x,y) & \geq & 0, \quad j=1,\dots,q, \\
 & v_i(x,y) & \geq & 0, \quad i=1,\dots,l, \\
 & \nabla_z \phi(x,y,y) - \sum\limits_{i=1}^{l} \lambda_i \nabla_y v_i(x,y) & = & 0, \\
 & \lambda_i & \geq & 0, \quad i=1,\dots,l, \\
 & \lambda_i v_i(x,y) & = & 0, \quad i=1,\dots,l.
\end{array} \tag{1.2.6}$$

Obviously, only under some constraint qualification on the feasible set $Y(x)$ in the lower level, we can guarantee that these schemes are relaxations of the feasible set for the original problems P_{BL} and P_{EC}. Unfortunately as we will see later on, the condition MFCQ cannot be expected to hold generically for the lower level problem.

If $Q(x)$ is a convex problem satisfying MFCQ, then the original problem P_{BL} and problem (1.2.4) are equivalent. Under similar assumptions on $Q(x,y)$, the problem P_{EC} in (1.2.1) is equivalent with the KKT relaxation (1.2.5) and, in case $\phi(x, y, y) = 0,\ \forall y \in Y(x), x \in \mathbb{R}^n$, the equivalence holds between P_{EC} in (1.2.2) and problem (1.2.6).

So, this KKT approach opens the way for solving P_{BL} and P_{EC} via programs with complementarity constraints of type P_{CC} (1.1.9).

1.3 Summary of the results

In the present subsection we give a summary of the thesis. We try to sketch the results also in comparison with earlier investigations. Throughout the exposition all earlier results (lemmas, theorems, propositions) are indicated by giving (at least) one reference. All results where such a reference does not appear, are (at least for a substantial part) new.

In essence, the main aim of the thesis is as follows:

- MPCC problems (*cf.* (1.1.9)) and the parametric smoothing approach for solving these programs (*cf.* (1.1.10)) are investigated from a structural and generical viewpoint.

- We apply the KKT approach to different types of equilibrium constrained problems (VI, BL and EC-problems). Thereby the equilibrium constrained programs are transformed into a problem of MPCC type with special structure, *cf.* Section 1.2. We analyze the generic properties of the resulting MPCC programs and study the behavior of the related parametric smoothing method.

The investigations on the analytic and generic structure of mathematical programs, form the basis for the development of any general purpose solution method. In fact the generic structure reveals the typical properties of a problem.

More detailed, the thesis is organized as follows. In Section 1.4, some applications of equilibrium constrained problems are presented.

The investigations of the whole exposition are based on the deep genericity results for (non-parametric and parametric) finite programming problems developed during the last two decades, starting with the work of Jongen, Jonker and Twilt in [27]-[28]. Further results appear in Guddat, Guerra and Jongen in [22], Gómez, Guddat, Jongen, Rückmann and Solano in [21], and Jongen, Jonker and Twilt in [29]. The basic concepts and results are outlined in Chapter 2.

Chapter 3 is devoted to Variational Inequality problems (VI). In Section 3.2, the KKT approach to solve VI is described. Section 3.3 summarizes the genericity results of Gómez [19] for this approach, applied to one-parametric VI problems.

In Section 3.4, we extend two different types of parametric embeddings for nonlinear programs (the standard embedding and the penalty embedding, see Gómez, Guddat, Jongen, Rückmann and Solano [21]) to the VI case. We analyze these solution methods from a structural and generical perspective. As new results, we can mention Proposition 3.2.1 (partially new), Example 3.2.2 and the genericity analysis of the one-parametric embeddings in Propositions 3.4.1-3.4.4.

Chapter 4 deals with general MPCC-problems of the form P_{CC} in (1.1.9). This is a topic of intensive recent research, see *e.g.* Scheel and Scholtes [49], Scholtes [51], Lin and Fukushima [37], Hu and Ralph [24] and the references in these contributions.

Firstly we discuss the structure of the feasible set in Section 4.2 (see *e.g.* Luo, Pang and Ralph [42]) and give some well-known necessary optimality conditions in Section 4.3 (see Flegel and Kanzow [14]). Section 4.4 derives other necessary and sufficient conditions for minimizers of order one and two of MPCC, based on the disjunctive structure, and sketches basic genericity results. Partially these results are due to Scholtes and Stöhr [53]. The optimality conditions for minimizer of order one in Theorem 4.4.2 and 4.4.3 are new (in the MPCC context) and also the only part of Theorem 4.4.4. Section 4.5 is concerned with the convergence behavior of the parametric smoothing approach P_τ for $\tau \to 0$ in (1.1.10). This approach has been studied in Fukushima and Pang [16] from another point of view. It has been investigated under which conditions the KKT solutions of P_τ converges to a B-stationary point of P_{CC}. In Stein and Still [60], such results were obtained for a similar (interior point) approach for solving semi-infinite programming problems. In this thesis we derive new convergence results for the whole feasible set and for (local) solutions $x(\tau)$ of P_τ for $\tau \to 0$, see Lemmas 4.5.2, 4.5.3 and Theorem 4.5.1. It is shown that under natural assumptions the rate of convergence is $\mathcal{O}(\sqrt{\tau})$. We also give some illustrative Examples 4.5.2-4.5.3.

In Section 4.6 we prove that generically P_τ is regular in the sense of Jongen, Jonker, Twilt (JJT), see Definition 2.4.7. These investigations are entirely new, see as main results Propositions 4.6.1-4.6.2. Section 4.7 deals with one-parametric MPCC problems. In the first part we discuss the one-parametric and non-parametric (feasibility) problem for the special case $n = l$. The related new results are given in Propositions 4.7.1-4.7.4. The second part is devoted to the general one-parametric MPPC problem. In Hu and Ralph [24], this problem has been analyzed only locally around non-degenerate minimizers. In this thesis, we develop the global theory and, based on the results of Jongen, Jonker and Twilt in [27]-[28] for nonlinear programs, we are able to describe the characteristics of generic one parametric MPCC programs and their singularities. The corresponding new results are contained in Lemma 4.7.3, Theorem 4.7.1. The analysis of the behavior of the set of stationary points around each singularity is also new.

Chapter 5 is devoted to bilevel problems, see (1.1.2). These problems have been extensively studied during the last two decades. We refer the reader *e.g.* to Shimizu and Aiyoshi [55], Bard [3], Luo, Pang and Ralph [42], Dempe [11] and the references therein. It is well-known that in general the structure of BL does not allow a local reduction to common finite programs. So special solution methods have to be conceived.

In this thesis we consider the KKT approach to solve BL which consists of a transformation of BL into a special structured MPCC problem (*cf.* (1.2.4)). This approach has been discussed earlier see *e.g.* Shimizu and Aiyoshi [55] for the general case and Stein [58], in connection with semi-infinite problems.

The new contribution of the present thesis is to analyze for the first time the MPCC form of BL programs from a generical viewpoint. This allows to apply the results obtained for MPCC problems in Chapter 4. All results of Section 5.2 and 5.3 are new.

In Theorem 5.2.1 of Section 5.2, we show that generically MPCC-LICQ (see Definition 4.3.1) holds at all feasible points of the MPCC form of BL. However a counterexample reveals that, in contrast to general MPCC problems, for the MPCC form of BL programs the condition MPCC-SC in Definition 4.3.3, is not generically fulfilled at minimizers. It follows that the structural difficulties of the original BL problem partially reappear in the MPCC formulation. Roughly speaking this may happen at solutions of BL where a local reduction to a finite standard problem is not possible.

Theorem 5.2.2 describes the precise generic structure around a critical point. Corollaries 5.2.3-5.2.5 reveal the relation between (the solutions of) the original BL and the MPCC form. An example describes a possible bad behavior of the KKT approach. It appears that even in the case of convex lower level problems it may happen that $(\overline{x}, \overline{y}, \overline{\lambda})$ is a minimizer of the KKT form but that the feasible point $(\overline{x}, \overline{y})$ is not a minimizer of the original problem. This bad behavior is stable w.r.t. small smooth perturbations. Based on the results of Section 5.2, in Section 5.3 we discuss the parametric approach for solving the MPCC formulation of BL. This leads to convergence results in Proposition 5.3.1 Also here a stable counterexample (Example 5.3.1) shows that it is not always possible to approach each minimizer $(\overline{x}, \overline{y}, \overline{\lambda})$ of the KKT formulation by the parametric smoothing procedure.

A part of our investigation was not yet completely finished at the time when the preliminary version of the thesis was send to the members of the Doctorate Committee. In the meantime this part has been completed and we would like to add it to the thesis as a supplement (Chapter 6).

In Chapter 6 equilibrium constrained problems EC (see (1.1.7)) are examined. Also this class of mathematical programs is a topic of intensive research in the past and at present. We refer the reader *e.g.* to the books Outrata, Kocvara and Zowe [43] and Luo, Pang and Ralph [42].

Here we are mainly interested in the structural and generical properties of such problems which are closely related to generalized semi-infinite programs (GSIP). In fact, EC programs can be seen as GSIP in the unknown (x, y) (see (1.1.3)) with the extra condition $y \in Y(x)$. Many investigations have been conducted on stability and the generic structure of semi-infinite problems. See *e.g.* Jongen and Zwier [34]-[33], Jongen, Twilt and Weber [32], Stein [57], Jongen and Rückmann [30] and Jongen and Rückmann [31] for common semi-infinite problems (SIP) (*i.e.* $Y(x) = Y$, $\forall x$) and Stein [58], Still [61], for GSIP.

In this thesis we are mainly devoted to the study of the generic properties of the KKT approach for solving EC. We continue the investigations started with the paper Birbil, Bouza, Frenk and Still [7] where only a special (linear) case was considered.

The Chapter is organized as follows. In Section 6.2 the topological structure of the feasible set of EC is discussed. As in GSIP it appears that the feasible set need even not to be closed in general. The results here were obtained in [7] (see also Stein [58] for GSIP). In Section 6.3 and 6.4 the MPCC formulation of EC (*cf.* (1.2.5) is analyzed for the first time from a generical viewpoint. A stable counterexample, Example 6.3.1, reveals that in contradiction to BL here the condition MPCC-LICQ (see Definition 4.3.1) is not generically fulfilled for all feasible points. This is due to the extra constraint $y \in Y(x)$ in EC.

So, in Section 6.4 we restrict ourselves to the class of EC problems with convex lower level program $Q(x, y)$ (see (1.2.1)) and satisfying additional regularity assumptions. For this class the MPCC formulation of EC (*cf.* (1.2.5)) is proven to be generically regular in the MPCC sense, see Definition 4.3.3. The corresponding new results are contained in Proposition 6.4.1 and Theorem 6.4.1.

In Section 6.5, we consider the special linear case, *i.e.* EC programs such that the MPCC form is given by only linear functions. Such problems have also been studied in [7] where the generic structure of the feasible set of the original EC program has been described. For this linear class, the structure of the KKT approach for EC appears to be similar to the structure for (linear) bilevel problems. So, by modifying the ideas in the proof of Theorems 5.2.1 and 5.2.2, we can show in Proposition 6.5.1 that MPCC-LICQ is a generic property and in Theorem 6.5.1, we describe the structure around local minimizers in the generic situation. The consequences of these results for the original EC program are shown in Corollary 6.5.2 and Proposition 6.5.2.

Finally based on the obtained genericity results, we describe a new algorithm for solving linear EC problems. The algorithm makes use of the MPCC form and performs descent steps on (faces of) the feasible set of the corresponding program. The finite algorithm (eventually) ends up with a local minimizer of the original EC.

1.4 Applications

In this section we present a few applications of equilibrium constrained problems. We will consider two main fields, namely applications in economics and applications in mathematical physics. They are mainly taken from [43], see [3] for other examples.

1.4.1 Applications in economics

We start with two situations where optimization problems with equilibrium constraints arise, the Cournot competitive equilibrium and the generalized Nash equilibrium. We begin with some notations and present different ways of modeling Nash equilibrium points.

Nash equilibrium

Consider a game with n players. For $i = 1, 2, \ldots, n$ the set of all possible strategies for player i is denoted by $Y_i \subset \mathbb{R}^{m_i}$. If for all j, the player j chooses strategy $y_j \in Y_j$, the payoff for player i is equal to $u_i(y_1, \ldots, y_n)$. We assume u_i to be a concave C^1-function in the variable y_i.

A Nash equilibrium point $\overline{y} = (\overline{y}_1, \ldots, \overline{y}_n)$ satisfies:

$$\overline{y}_i \text{ solves } \max_{y_i \in Y_i} u_i(\overline{y}_1, \ldots, \overline{y}_{i-1}, y_i, \overline{y}_{i+1}, \ldots \overline{y}_n), \quad \forall i.$$

In case that, for any i, the set Y_i is a non-empty, closed and convex set, the previous optimization problems are convex. Consequently, another way of expressing that $\overline{y}$ is a Nash equilibrium is via the following variational inequality. The point $\overline{y}$ is a Nash equilibrium if and only if:

$$\begin{gathered} \overline{y} \in Y, \\ \text{such that } \langle F(\overline{y}), v - \overline{y} \rangle \geq 0, \ \forall v \in Y = Y_1 \times Y_2 \times \ldots \times Y_n \end{gathered} \tag{1.4.1}$$

where

$$F(y) = \begin{pmatrix} -\nabla_{y_1} u_1(y) \\ \vdots \\ -\nabla_{y_n} u_n(y) \end{pmatrix}.$$

Now, as a first application, we will present an optimization problem in which the feasible set is described by a Nash equilibrium.

Cournot equilibrium

In this model, there are n firms producing a certain good. Each firm has to decide how many units of the product it will produce. The decision of firm i is denoted by y_i. The price of the product is $P(T)$, where T denotes the total amount of the good in the market, *i.e.*, $T = \sum_{i=1}^n y_i$. The cost of producing y_i units for the firm i is described by $f_i(y_i)$. Its profit is then given by $u_i(y) = y_i p(T) - f_i(y_i)$. Of course, here we have $Y_i \subset \mathbb{R}_+$.

Suppose the firm 1 places its production in the market, say y_1, first. Knowing this value, the other firms will plan their productions in order to maximize their profits. The first firm must decide the value of y_1 that maximizes its profit under these circumstances. Then it will solve the bilevel problem:

$$
\begin{array}{lcc}
 & \min\limits_{x,y} f_1(x) - xp(x + \sum\limits_{j=2}^{n} y_j) & \\
\text{s.t.} & x \in Y_1, & \\
 & \text{and for } i = 2, \ldots, n & (1.4.2) \\
y_i \text{ solves} & \min\limits_{z} f_i(z) - zp(x + \sum\limits_{j=1}^{i-1} y_i + z + \sum\limits_{j=i+1}^{n} y_i) & \\
 & \text{s.t.} \quad z \in Y_i, &
\end{array}
$$

Of course if $f_i(y_i) - y_i p(x + \sum_{i=1}^n y_i) \in C^1$ are convex functions of y_i and Y_i are convex sets, $i = 1, \ldots, n$, the system (1.4.1) characterizes the equilibrium points as solutions of a variational inequality problem. Then we find the equivalent formulation as a P_{VI}:

$$
\begin{array}{ll}
 & \min\limits_{x,y} f_1(x) - xp(x + \sum\limits_{j=2}^{n} y_j) \\
\text{s.t.} & x \in Y_1, \\
 & y_i \in Y_i \quad i = 2, \ldots, n, \\
 & \langle F(x, y_2, \ldots, y_n), z - (y_2, \ldots, y_n)^T \rangle \geq 0, \ \forall z \in Y_2 \times Y_3 \times \ldots \times Y_n,
\end{array}
$$

where $F : \mathbb{R}^n \to \mathbb{R}^{n-1}$ is given by $F_i(x, y_2, \ldots, y_n) = f'_{i+1}(y_{i+1}) - p(T) - y_{i+1}p'(T)$, $i = 1, \ldots, n-1$ and $T = x + \sum\limits_{i=2}^{n} y_i$.

Generalized Nash equilibrium problem

In this example we consider a game of n players, where the payoff of player i is described by the function u_i, $i = 1, 2, \ldots, n$ and the strategies of player i depend on the decisions of the other players. The set of feasible strategies for player i is

$$
Y_i(y_{-i}) = \left\{ y_i \in \mathbb{R}^{m_i} \;\middle|\; \begin{array}{rcll} g_i^j(y_i, y_{-i}) & \geq & 0, & j = 1, \ldots, q_i, \\ \overline{g}_i^j(y_i) & \geq & 0, & j = 1, \ldots, l_i \end{array} \right\}
$$

if the player j chooses strategy y_j for $j = 1, \ldots i-1, i+1, \ldots n$. Here $y_{-i} = (y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_n)$. Now the problem is to find $\overline{y} = (\overline{y}_1, \ldots, \overline{y}_n)$

such that, for $i = 1, 2, \ldots, n$,

$$\begin{array}{lll} \overline{y}_i \text{ solves} & & \max_z u_i(\overline{y}_1, \ldots, \overline{y}_{i-1}, z, \overline{y}_{i+1}, \ldots, \overline{y}_n) \\ & \text{s.t.} & z \in Y_i(\overline{y}_{-i}). \end{array}$$

If we assume that u_i is concave in the variable y_i and the sets $Y_i(y_{-i})$ are non-empty and convex, the previous model is equivalent to the problem of finding $\overline{y}$:

$$\begin{gathered} \overline{y}_i \in Y_i(\overline{y}_{-i}), \\ \text{such that } \langle -\nabla_{y_i} u_i(\overline{y}), y_i - \overline{y}_i \rangle \geq 0,\ \forall y_i \in Y_i(\overline{y}_{-i}),\ i = 1, \ldots, n \end{gathered}$$

which is equivalent with finding $(\overline{y}, x)$ such that:

$$\begin{gathered} \overline{y} \in Y(x), \\ \langle F(\overline{y}), y - \overline{y} \rangle \geq 0,\ \forall y \in Y(x), \\ x = \overline{y}, \end{gathered} \tag{1.4.3}$$

where $F(y) = \begin{bmatrix} -\nabla_{y_1} u_1(y) \\ \vdots \\ -\nabla_{y_n} u_n(y) \end{bmatrix}$ and $Y(x) = Y(x_{-1}) \times Y(x_{-2}) \times \ldots \times Y(x_{-n})$.

If we write the constraint $x = y$ as $\min \|x - y\|^2$, we have the following formulation of the generalized Nash equilibrium problem as a P_{VI}:

$$\begin{array}{ll} & \min_{x,y} \|x - y\|^2 \\ \text{s.t.} & y \in Y(x), \\ & \langle F(y), z - y \rangle \geq 0,\ \forall z \in Y(x). \end{array} \tag{1.4.4}$$

Of course $\overline{y}$ will be a generalized Nash equilibrium point if and only if $(\overline{y}, \overline{y})$ solves the previous problem.

1.4.2 Applications from mathematical physics

One way of finding approximate solutions of control problems is by discretizing the domain and solving an approximate nonlinear problem. In this part we will present equilibrium problems appearing when solving control models by a discretization approach, *cf.* [43]. Let us first fix some notations. We assume that α is a function which is differentiable almost everywhere (a.e.). The set of feasible functions is:

$$U = \left\{ \alpha : [0, 1] \to [c_1, c_2] \;\middle|\; |\frac{\partial \alpha}{\partial x}(x)| \leq c_3,\ \text{a.e., } x \in [0, 1] \right\}.$$

For $\alpha \in U$, the set Ω_α is defined as $\{x \in \mathbb{R}^2 \mid x_1 \in [0, 1],\ 0 < x_2 < \alpha(x_1)\}$. Its area will be $J(\alpha)$.

Ω_0 denotes a fixed domain, that is assumed to be included in $(0,1) \times (0, c_1)$. For fixed $\alpha \in U$, u describes the deformation, by a force f, of a membrane in Ω_α. On the boundary of Ω_α, the membrane is not deformed, *i.e,*

$$u(x) = 0, \qquad x \in \partial\Omega_\alpha. \tag{1.4.5}$$

Packaging problem with rigid obstacle

In this problem, given $\chi, f : [0,1] \times [0, c_2] \to \mathbb{R}$, $f, \chi \in L_2([0,1] \times [0, c_2])$, we have to find α, $\alpha \in U$, that minimizes the area $J(\alpha)$ of Ω_α, under the conditions that there is a membrane in Ω_α, given by u, deformed by f, such that (1.4.5) is satisfied. The membrane lies over the rigid object described by the function χ and it has to be in contact with this object in the fixed set Ω_0. The model then is:

$$\begin{array}{lrcll} & \min\limits_{u,\alpha} J(\alpha) & & & \\ \text{s.t.} & -\Delta u(x) & \geq & f(x), & \text{a.e. in } \Omega_\alpha, \\ & u(x) & \geq & \chi(x), & x \in \Omega_\alpha, \\ & (\Delta u(x) + f(x))(u(x) - \chi(x)) & = & 0, & \text{a.e. in } \Omega_\alpha, \\ & u(x) & = & 0, & \forall x \in \partial\Omega_\alpha, \\ & \Omega_0 & \subset & Z(\alpha), & \\ & \alpha & \in & U, & \end{array}$$

where u denotes the deformation of the membrane, Δu the Laplacian of u and $Z(\alpha) = \{x \in \Omega_\alpha \mid u(x) = \chi(x)\}$.

In order to solve this problem, the condition $\Omega_0 \subset Z(\alpha)$ is eliminated via a penalty approach and the objective function becomes: $J(\alpha) + r\int_{\Omega_0}(u(\xi) - \chi(\xi))d\xi$.

Let $D_{\hat{\alpha}}(h)$ be a suitable discretization of the domain Ω_α with mesh size $h = \frac{1}{n}$, see [43] for details. If the involved functions are approximated by piecewise-linear interpolating functions, we obtain the following nonlinear approximate problem

$$\begin{array}{lrcl} & \min\limits_{\hat{\alpha},\hat{v}} J(\hat{\alpha}, h) + rh^2 \sum\limits_{i \in D_0(h)} \hat{v}_i & & \\ \text{s.t.} & A(\hat{\alpha}, h)\hat{v} + A(\hat{\alpha}, h)\hat{\chi}(\hat{\alpha}, h) - \hat{f}(\hat{\alpha}, h) & \geq & 0, \\ & \hat{v} & \geq & 0, \\ & \langle A(\hat{\alpha}, h)\hat{v} + A(\hat{\alpha}, h)\hat{\chi}(\hat{\alpha}, h) - \hat{f}(\hat{\alpha}, h), \hat{v}\rangle & = & 0, \\ & \hat{\alpha} & \in & \hat{U}. \end{array} \tag{1.4.6}$$

Here $\hat{\alpha}_i = \alpha(\frac{i}{n})$, $i = 0, 1, \ldots, n$, $D_0(h) = D_{\hat{\alpha}}(h) \cap \Omega_0$, and for $x_j \in D_{\hat{\alpha}}(h)$, $j = 1, \ldots, |D_{\hat{\alpha}}(h)|$, we have the following approximations: $\hat{u}_j = u(x_j)$, $\hat{f}(\hat{\alpha}, h)_j = f(x_j)$, $\hat{\chi}(\hat{\alpha}, h)_j = \chi(x_j)$. The vector $\hat{v}$ is equal to $\hat{u} - \hat{\chi}(\hat{\alpha}, h)$ and $A(\hat{\alpha}, h)$ denotes the matrix such that $[A(\hat{\alpha}, h)\hat{u}]_j \approx \Delta\overline{u}(x_j)$, where $\overline{u}(x)$ is the piecewise-linear function interpolating u in $D_{\hat{\alpha}}(h)$. Finally

$$\hat{U} = \left\{ \hat{\alpha} \in \mathbb{R}^{n+1} \;\middle|\; \begin{array}{c} \text{there is a piecewise linear function } \alpha \in U \\ \text{such that } \hat{\alpha}_i = \alpha(\frac{i}{n}),\ i = 0, 1, \ldots, n \end{array} \right\}.$$

It can be seen that $\hat{U}$ can be written as the set of vectors $\hat{\alpha} \in \mathbb{R}^{n+1}$ such that $\hat{\alpha}_i \in [c_1, c_2],\ i = 0, \ldots, n$, and $\left|\frac{\hat{\alpha}_i - \hat{\alpha}_{i-1}}{n}\right| \leq c_3,\ i = 1, 2, \ldots, n$.

The problem (1.4.6) is a MPCC problem. Note that it has also the structure of the P_{VI} problem (1.1.8), since the set of feasible solutions can be written as:

$$\hat{\alpha} \in \hat{U},\ \hat{v} \geq 0,\ \langle A(\hat{\alpha})\hat{v} + A(\hat{\alpha})\hat{\chi}(\hat{\alpha}) - \hat{f}(\hat{\alpha}), z - \hat{v} \rangle \geq 0,\ \forall z \in \mathbb{R}^m_+.$$

Packaging problem with compliant obstacle

In this example the object can be deformed by the membrane. The surface of the object is described by $G(u,x) = k(\Delta u(x) + f(x)) + \chi(x)$, where χ is the original shape of the object and $1/k$ is the compliance of the obstacle material, see [43] for details. The model is

$$\begin{array}{lrcll}
 & \min\limits_{u,\alpha} J(\alpha) & & & \\
\text{s.t.} & -\Delta u(x) & \geq & f(x), & \text{a.e. in } \Omega_\alpha, \\
 & u(x) & \geq & G(u,x), & \text{a.e. in } \Omega_\alpha, \\
 & (\Delta u(x) + f(x))(u(x) - G(u,x)) & = & 0, & \text{a.e. in } \Omega_\alpha, \\
 & u(x) & = & 0, & x \in \partial\Omega_\alpha, \\
 & \Omega_0 & \subset & Z(\alpha), & \\
 & \alpha & \in & U. &
\end{array}$$

Again the condition $\Omega_0 \subset Z(\alpha)$ is penalized and the objective function becomes $J(\alpha) + r\int_{\Omega_0}(u(\xi) - G(u,\xi))d\xi$. In this case, after applying the same discretization step of mesh size h, the resulting problem will be:

$$\begin{array}{lr}
 & \min\limits_{\hat{\alpha},\hat{u}} J(\hat{\alpha}, h) + rh^2 \sum\limits_{i \in D_0(h)} (\hat{u} - \hat{G}(\hat{\alpha}, h, \hat{u}))_i \\
\text{s.t.} & A(\hat{\alpha}, h)\hat{u} - \hat{f}(\hat{\alpha}, h) \geq 0, \\
 & \hat{u} - \hat{G}(\hat{\alpha}, h, \hat{u}) \geq 0, \\
 & \langle A(\hat{\alpha}, h)\hat{u} - \hat{f}(\hat{\alpha}, h), \hat{u} - \hat{G}(\hat{\alpha}, h, \hat{u}) \rangle = 0,
\end{array}$$

where $\hat{G}(\hat{\alpha}, h, \hat{u}) = k(\hat{f}(\hat{\alpha}, h) - A(\hat{\alpha}, h)\hat{u}) + \hat{\chi}(\hat{\alpha}, h)$. Here we have given a MPCC problem, which cannot be seen as a P_{VI} problem.

Chapter 2

Theoretical background

The aim of the present chapter is to introduce some notation and concepts in optimization and to sketch the deep genericity results of finite programming, see *e.g.* [29]. They form the basis of the structural and genericity analysis for equilibrium constrained programs presented later on.

2.1 Notations and basic results

As usual $\mathbb{R}^n$ denotes the n-dimensional Euclidean space. We often use the notation

$$\mathbb{R}^n_+ = \{x \in \mathbb{R}^n \mid x_i \geq 0,\ i = 1, 2, \dots, n\}$$

and

$$\mathbb{R}^n_{++} = \{x \in \mathbb{R}^n \mid x_i > 0,\ i = 1, 2, \dots, n\}.$$

Given $I \subset \{1, 2, \dots, n\}$, for a vector $x \in \mathbb{R}^n$, x_I denotes the $|I|$-dimensional vector with components $x_i,\ i \in I$. We define $x_{-I} = x_{I^c}$, with $I^c = \{1, 2, \dots, n\} \setminus I$. If $I = \{i\}$, obviously $x_I = x_i$ and we often write x_{-i} instead of x_{-I}. For a matrix $A \in \mathbb{R}^{m \times n}$, A_i denotes its i^{th}-column and A_I is the $m \times |I|$-matrix with columns $A_i,\ i \in I$. As usual, the matrix I_n represents the identity $n \times n$-matrix. If n is known, we simply write I.
For denoting a positive (semi)-definite matrix $A \in \mathbb{R}^{n \times n}$, we write $A \succ 0 (\succeq 0)$.
We give two classical results from matrix theory used later on.

Lemma 2.1.1 (Farkas Lemma) *If $M = \{x \in \mathbb{R}^n \mid Ax \leq 0\}$, $A \in \mathbb{R}^{m \times n}$, then $c^T x \leq 0$, $\forall x \in M$, if and only if $c = A^T y$, for some $y \geq 0,\ y \in \mathbb{R}^m$.*

Let B' be a linear subspace of $\mathbb{R}^n$, and V a matrix whose columns form a basis of B'. We will denote by $A|_{B'}$ the matrix $V^T A V$.

Proposition 2.1.1 *Let A be a symmetric matrix and B a $n \times p$ matrix, $n \geq p$. If $B' = \left\{x \mid B^T x = 0\right\}$, then the number of positive (negative, zero, respectively)* *eigenvalues of* $\begin{pmatrix} A & B \\ B^T & 0 \end{pmatrix}$ *is equal to the number of positive (negative, zero, respectively) eigenvalues of* $A|_{B'}$ *plus* $rank(B)$ *(plus* $rank(B)$*, plus* $(p - rank(B))$*, respectively).*

We further introduce $\|x\|_p = \left(\sum_{i=1}^{n} | x_i^p |\right)^{\frac{1}{p}}$. $\|x\|$ will always be the Euclidean norm $\|x\|_2$. The distance of a point $\hat{x} \in \mathbb{R}^n$ to a set $\mathcal{M} \subset \mathbb{R}^n$ is defined by $d(\hat{x}, \mathcal{M}) = \inf\{\|x - \hat{x}\| \mid x \in \mathcal{M}\}$.
We also use the notation $B_\varepsilon^n(\overline{x}) = \{x \in \mathbb{R}^n \mid \|x - \overline{x}\| < \varepsilon\}$ and denote the closure of $B_\varepsilon^n(\overline{x})$ by $\overline{B}_\varepsilon^n(\overline{x})$.
For $f : \mathbb{R}^n \to \mathbb{R}$, ∇f represents the gradient of f taken as a column vector.
Finally we consider $[C^k]_n^m$ as the space of $k-$times continuously differentiable functions with domain $\mathbb{R}^n$ and image in $\mathbb{R}^m$. $[C_S^k]_n^m$ is the space $[C^k]_n^m$ endowed with the strong topology, see Section 2.3.

2.2 Finite programming problems

In nonlinear programming a real valued function is minimized on the feasible set $M \subset \mathbb{R}^n$ described by finitely many equalities and inequalities. In most cases the involved functions are supposed to be C^k-functions. In this thesis a finite programming problem P is of the form

$$P: \quad \begin{array}{rc} & \min f(x) \\ \text{s.t.} & x \in M \end{array} \tag{2.2.1}$$

$$M = \left\{ x \in \mathbb{R}^n \;\middle|\; \begin{array}{rcll} h_i(x) & = & 0, & i = 1, \ldots, q_0, \\ g_j(x) & \geq & 0, & j = 1, \ldots, q \end{array} \right\} \tag{2.2.2}$$

with given functions $f, h_i, g_j \in [C^k]_n^1$, $i = 1, \ldots, q_0$, $j = 1, \ldots, q$, $k \geq 2$. We often use the abbreviation $h = (h_1, \ldots, h_{q_0})$, $g = (g_1, \ldots, g_q)$.

We want to find a (local) minimizer $\overline{x} \in M$. Assuming that the feasible set M is nonempty and compact, a minimizer always exists.

Definition 2.2.1 Given $f : \mathbb{R}^n \to \mathbb{R}$, $M \subset \mathbb{R}^n$, the point $\overline{x} \in M$ is *a local minimizer of f on M,* if there is a neighborhood $V(\overline{x})$ of $\overline{x}$ such that:

$$f(x) \geq f(\overline{x}), \; \forall x \in V(\overline{x}) \cap M.$$

It is called *a global minimizer* if this inequality holds $\forall x \in M$.

We say that $\overline{x}$ is *a local minimizer of order* $\omega, \omega > 0$, if there is a neighborhood $V(\overline{x})$ of $\overline{x}$, and a constant $\kappa > 0$, such that:

$$f(x) - f(\overline{x}) \geq \kappa \|x - \overline{x}\|^\omega, \; \forall x \in V(\overline{x}) \cap M.$$

We introduce some more definitions and notations.

Definition 2.2.2 For $\overline{x} \in M$ the set $J_0(\overline{x})$ of *active indices of* $\overline{x}$, is denoted as $J_0(\overline{x}) = \{j \mid g_j(\overline{x}) = 0\}$. *The condition LICQ holds at* $\overline{x}$ if the set of vectors

$$\{\nabla h_i(\overline{x}),\ i = 1, \ldots, q_0,\ \nabla g_j(\overline{x}),\ j \in J_0(\overline{x})\}$$

is linearly independent. *The constraint qualification MFCQ is satisfied at* $\overline{x}$ if

- $\nabla h_i(\overline{x}),\ i = 1, \ldots, q_0$, are linearly independent and
- there is a vector $\xi \in \mathbb{R}^n$ such that:

$$\begin{array}{rcl} \xi^T \nabla h_i(\overline{x}) & = & 0, \quad i = 1, \ldots, q_0, \\ \xi^T \nabla g_j(\overline{x}) & > & 0, \quad j \in J_0(\overline{x}). \end{array}$$

As usual the Lagrangean function of the finite problem P near $\overline{x}$ is defined by

$$L(x, \lambda, \mu) = f(x) - \sum_{i=1}^{q_0} \lambda_i h_i(x) - \sum_{j \in J_0(\overline{x})} \mu_j g_j(x),$$

where the numbers $\lambda_i,\ i = 1, \ldots, q_0,\ \mu_j,\ j \in J_0(\overline{x})$, are called Lagrange multipliers.

The following KKT optimality condition is standard, see *e.g.* Luenberger [41] or Bazara, Sherali and Shetty [4].

Theorem 2.2.1 (First order necessary conditions, *cf.* **[41], [4])** *Let* $\overline{x} \in M$ *be a local minimizer of* f *on* M *such that MFCQ holds at* $\overline{x}$. *Then there exists* λ_i, $i = 1, \ldots, q_0,\ \mu_j,\ j \in J_0(\overline{x})$, *such that*

$$\nabla f(\overline{x}) - \sum_{i=1}^{q_0} \lambda_i \nabla h_i(\overline{x}) - \sum_{j \in J_0(\overline{x})} \mu_j \nabla g_j(\overline{x}) = 0, \tag{2.2.3}$$

$$\mu_j \geq 0. \tag{2.2.4}$$

If LICQ holds at $\overline{x}$ *the multipliers* $\lambda_i,\ i = 1, \ldots, q_0,\ \mu_j,\ j \in J_0(\overline{x})$, *satisfying (2.2.3), are uniquely determined.*

Remark 2.2.1 *If* $\overline{x}$ *is a local minimizer and MFCQ fails, then the so-called Fritz John (FJ) condition is fulfilled, i.e., there exists* $(\lambda_0, \lambda, \mu) \neq 0,\ \lambda_0,\ \mu_j \geq 0$, $j \in J_0(\overline{x})$ *such that*

$$\lambda_0 \nabla f(\overline{x}) - \sum_{i=1}^{q_0} \lambda_i \nabla h_i(\overline{x}) - \sum_{j \in J_0(\overline{x})} \mu_j \nabla g_j(\overline{x}) = 0.$$

The points $\overline{x}$ satisfying these conditions are called *Fritz John points.*

Definition 2.2.3 The point $\overline{x} \in M$ is called *a critical point* if LICQ is satisfied at $\overline{x}$ and if there are multipliers $\lambda_i, \ i = 1, \ldots, q_0, \ \mu_j, \ j \in J_0(\overline{x})$, such that $(\overline{x}, \ \lambda, \ \mu)$ satisfies the system (2.2.3).
If for some (λ, μ), the point $\overline{x} \in M$ solves system (2.2.3) and for these multipliers also (2.2.4) holds, then we call $\overline{x}$ *a stationary point.*
A point $\overline{x}$ is *a generalized critical point (g.c. point),* if the set of vectors

$$\{\nabla f(\overline{x}), \ \nabla h_i(\overline{x}), \ i = 1, \ldots, q_0, \ \nabla g_j(\overline{x}), j \in J_0(\overline{x})\} \tag{2.2.5}$$

is linearly dependent. The set of all generalized critical points of P is denoted by Σ_{gc}.

Of course at a stationary point where LICQ fails to hold, the multipliers may not be unique. A necessary and sufficient condition for uniqueness is the so-called strong-MFCQ, obtained as a consequence of the Lemma of Farkas, see Lemma 2.1.1.

Definition 2.2.4 If $\overline{x}$ is a stationary point with multipliers $(\lambda, \mu), \ \mu \geq 0$, then *the strong-MFCQ condition is said to hold* if:

- The vectors $(\nabla h_1, \ldots, \nabla h_{q_0}, \nabla g_{J_0(\overline{x}) \cap J_+(\mu)})(\overline{x})$ are linearly independent, where $J_+(\mu) = \{j \mid \mu_j > 0\}$ and

- there is some $\xi \in \mathbb{R}^n$ such that

$$\begin{array}{rcl} [\nabla h_i(\overline{x})]^T \xi & = & 0, \quad i = 1, \ldots, q_0, \\ [\nabla g_j(\overline{x})]^T \xi & = & 0, \quad j \in J_0(\overline{x}) \cap J_+(\mu), \\ [\nabla g_j(\overline{x})]^T \xi & > & 0, \quad j \in J_0(\overline{x}) \setminus J_+(\mu). \end{array}$$

Let us assume that $\overline{x} \in M$ satisfies LICQ. We define the *tangent space* :

$$T_{\overline{x}}M = \left\{ \xi \in \mathbb{R}^n \ \middle| \ \begin{array}{rcl} [\nabla h_i(\overline{x})]^T \xi & = & 0, \quad i = 1, \ldots, q_0, \\ [\nabla g_j(\overline{x})]^T \xi & = & 0, \quad j \in J_0(\overline{x}). \end{array} \right\}$$

and denote by $A|_{T_{\overline{x}}M}$ the matrix $V^T A V$ where the columns of V form a basis of the space $T_{\overline{x}}M$.

Definition 2.2.5 A point $\overline{x} \in M$ is *a non-degenerate critical point of* P if it is a critical point, with unique multipliers (λ, μ), satisfying:

(i) $\mu_j \neq 0, \ j \in J_0(\overline{x})$.

(ii) $\nabla_x^2 L(\overline{x}, \lambda, \mu)|_{T_{\overline{x}}M}$ is non-singular.

A problem P is called *regular* if at all its feasible points the condition LICQ holds and if all its critical points are non-degenerate.

To formulate genericity results in finite programming, we will identify the set of problems P with the function space $\mathcal{P}_{q_0+q} := \{(f,h,g)\} = [C^k]_n^{1+q_0+q}$. The following theorem contains the main genericity result in finite nonlinear programming, see [22].

Theorem 2.2.2 (Genericity theorem, *cf.* **[22])** *Let $\mathcal{F} \subset \mathcal{P}_{q_0+q}$ denote the set of functions $(f,h,g) \in [C^2]_n^{1+q_0+q}$ such that for the associated optimization problem (2.2.1):*

(i) LICQ holds at all its feasible points.

(ii) All its critical points are non-degenerate.

Then the set $\mathcal{F}$ is a dense and open subset of $[C^2]_n^{1+q_0+q}$ with respect to the strong topology (see Definition 2.3.1).

2.3 Preliminaries from topology

In this section we will present some definitions on differential manifolds and topology for the space of smooth functions. For a more detailed discussion on the topic we refer to Hirsch [23] and [29].

Definition 2.3.1 For finite k, *the strong Whitney topology on* $[C^k]_n^1$ is obtained by considering the following local neighborhood system of the zero function:

$$V^k_{\varepsilon(x)} = \left\{ f(x) \;\middle|\; \left| \frac{\partial^r f}{\partial x_{i_1} \ldots . \partial x_{i_r}}(x) \right| < \varepsilon(x),\ \forall x \in \mathbb{R}^n, r \leq k \right\}$$

where ε is a continuous functions $\varepsilon : \mathbb{R}^n \to \mathbb{R}_{++}$. We will call this topology C^k_S topology and we denote the space $[C^k]_n^1$ endowed with this topology by $[C^k_S]_n^1$.
The C^∞_S topology in $[C^\infty]_n^1$ is the result of taking, as neighborhood system, the union of all sets $V^k_{\epsilon(x)}$ for all $k \in \mathbb{N} \cup \{0\}$.
In the case of the set $[C^k_S]_n^m$, $[C^\infty_S]_n^m$, *the strong topology* is obtained by the product topology.

As an important fact, it holds that the topological spaces $[C^k_S]_n^m$, $[C^\infty_S]_n^m$ are Baire spaces, see [29].

Definition 2.3.2 A set $\mathcal{B} \subset [C^k_S]_n^m$ is *generic in* $[C^k_S]_n^m$ if $\mathcal{B} = \cap_{i=1}^\infty \mathcal{B}_i$, with $\mathcal{B}_i$ open and dense sets in $[C^k_S]_n^m$.
We also say that *a property holds generically in* $[C^k_S]_n^m$ if it holds for a generic subset $\mathcal{B}$ in $[C^k_S]_n^m$.

Let us now present some definitions on differential manifolds in $\mathbb{R}^n$.

Definition 2.3.3 $M \subset \mathbb{R}^n$ is *an r-dimensional C^k-manifold* if and only if there are an open cover U_i, $i \in \Lambda$, $M \subset \bigcup_{i\in\Lambda} U_i$ and functions ϕ_i such that $\phi_i : U_i \to V_i$, $V_i \subset \mathbb{R}^r$ and $\phi_i(\phi_j^{-1})|_{V_i \bigcap V_j}$ is a C^k-function, $k \geq 1$.

We will mostly use manifolds of $\mathbb{R}^{n+r}$ that can be written as

$$M = \Phi^{-1}(0),$$

where $\Phi : \mathbb{R}^{n+r} \to \mathbb{R}^r$, $\Phi \in C^1$ and $\nabla\Phi(x)$ has full rank r for all x such that $\Phi(x) = 0$. In this case, M is an n-dimensional C^1 manifold in $\mathbb{R}^{n+r}$ and its co-dimension is r. Related with this fact there is the concept of regular values of a function.

Definition 2.3.4 Let the function $\phi : \mathbb{R}^n \to \mathbb{R}^r$ be in $[C^1]^r_n$. The value $y \in \mathbb{R}^r$ is said to be *a regular value of* ϕ if the matrix $\nabla\phi(x)$ has rank r for all $x \in \mathbb{R}^n$ such that $\phi(x) = y$.

The proof of the genericity results, later on, will mostly be based on the following important result, see [21].

Lemma 2.3.1 (Parameterized Sard Lemma, *cf.* **[21])** *Let $\phi(x, z)$ be in $[C^k]^r_{n+p}$, with $k > max\{0, n-r\}$ and $x \in \mathbb{R}^n$, $z \in \mathbb{R}^p$. Let us assume that 0 is a regular value of ϕ. Then for almost every $z \in \mathbb{R}^p$, 0 is a regular value of the function $\hat{\phi}_z : \mathbb{R}^n \to \mathbb{R}^r$, $\hat{\phi}_z(x) = \phi(x, z)$.*

We also give a main result in transversality theory.

Definition 2.3.5 Let M_1 and M_2 be C^1-manifolds in $\mathbb{R}^n$. We say that M_1, M_2 *intersect transversally, denoted by $M_1 \pitchfork M_2$,* if for every $\overline{x} \in M_1 \cap M_2$, it follows $T_{\overline{x}}M_1 + T_{\overline{x}}M_2 = \mathbb{R}^n$.

Definition 2.3.6 For $F \in [C^k]^m_n$, we define *the k-jet mapping or k-jet extension of F* by

$$j^kF(x) = (x, F(x), \nabla_x F(x), \nabla^2_x F(x), \ldots, \nabla^k_x F(x)),$$

where the elements of $\nabla^r F(x)$, $r \geq 2$ appear modulus symmetries.
The smallest Euclidean space containing the image of $j^kF(x)$ is called *jet space* and will be denoted as $J(n, m, k)$. $W_F = \left\{j^kF(x)|\ x \in \mathbb{R}^n\right\}$ is *the jet manifold.*

For a manifold V in $J(n, m, k)$, the set $\{F \in [C^\infty]^m_n \mid W_F \pitchfork V\}$ is written as $\pitchfork^k V$.

Theorem 2.3.1 (Jet Transversality theorem, *cf.* **[23])** *Let $k \in \mathbb{N}$ be fixed. Then for all $i \in \mathbb{N}$, the set $\pitchfork^k V$ is dense in $[C^\infty]^m_n$ with respect to the C^i_S topology. If V is a closed set of $J(n, m, k)$ then the set $\pitchfork^k V \subset [C^\infty]^m_n$ is open with respect to the C^i_S topology for $i \geq k + 1$.*

2.4 One-parametric optimization

The present section deals with one-parametric finite problems of the form

$$P(t): \qquad \min_x f(x,t) \quad \text{s.t.} \quad x \in M(t) \tag{2.4.1}$$

$$M(t) = \left\{ x \in \mathbb{R}^n \;\middle|\; \begin{array}{rcll} h_i(x,t) & = & 0, & i = 1, \ldots, q_0, \\ g_j(x,t) & \geq & 0, & j = 1, \ldots, q \end{array} \right\} \tag{2.4.2}$$

where $t \in \mathbb{R}$ is the parameter. For more details the reader is referred to Bank, Guddat, Klatte, Kummer and Tammer [1], [21] and [22]. The concepts in non-parametric optimization can be easily extended to the parametric case.

Definition 2.4.1 $(\overline{x}, \overline{t})$ is called *a local minimizer of* $P(t)$, also written as $(\overline{x}, \overline{t}) \in \Sigma_{loc}(P(t))$, if $\overline{x}$ is a local minimizer of $f(x, \overline{t})$ in $M(\overline{t})$.
The point $(\overline{x}, \overline{t})$ is *a generalized critical point of* $P(t)$, if $\overline{x} \in M(\overline{t})$ and the vectors $\left\{\nabla_x f(\overline{x}, \overline{t}),\ \nabla_x h_i(\overline{x}, \overline{t}),\ i = 1, \ldots, q_0,\ \nabla_x g_j(\overline{x}, \overline{t}),\ j \in J_0(\overline{x}, \overline{t})\right\}$ are linearly dependent, where $J_0(\overline{x}, \overline{t}) = \left\{j \mid g_j(\overline{x}, \overline{t}) = 0\right\}$. The set of g.c. points of $P(t)$ is denoted by $\Sigma_{gc}(P(t))$.
If the problem $P(t)$ is clearly identified, the sets of local minimizers and g.c. points are simply abbreviated as Σ_{loc} and Σ_{gc}, respectively.

The definitions of LICQ, critical points, Lagrange function $L(x, t, \lambda, \mu)$ near $(\overline{x}, \overline{t})$, Lagrange multipliers, etc., given in Section 2.2, are extended analogously.

For a vector $y \in \mathbb{R}^m$ we introduce the notation $J_{\neq}(y) = \{i \mid y_i \neq 0\}$.
Now we will present 5 types of g.c. points for one-parametric problems. They were defined and studied in [27], [28], see also [21] and [22].

In the following, Σ^i_{gc}, $i = 1, \ldots, 5$, denotes the set of g.c. points of type i. At a critical point $(\overline{x}, \overline{t})$, the vector $(\overline{\lambda}, \overline{\mu})$ represents the uniquely determined multipliers such that $\nabla_x L(\overline{x}, \overline{t}, \overline{\lambda}, \overline{\mu}) = 0$. In the g.c. points where LICQ fails, the multipliers (λ, μ) are such that

$$\sum_{i=1}^{q_0} \lambda_i \nabla_x h_i(\overline{x}, \overline{t}) - \sum_{j \in J_0(\overline{x}, \overline{t})} \mu_j \nabla_x g_j(\overline{x}, \overline{t}) = 0.$$

W.l.o.g. we assume that $J_0(\overline{x}, \overline{t}) = \{1, \ldots, p\}\,,\ p \leq q$.

Definition 2.4.2 For $(\overline{x}, \overline{t}) \in \Sigma_{gc}$, we write $(\overline{x}, \overline{t}) \in \Sigma^1_{gc}$, and say that $(\overline{x}, \overline{t})$ *is a generalized critical point of type 1,* if:

(1a) LICQ holds at $(\overline{x}, \overline{t})$.

(1b) $J_0(\overline{x}, \overline{t}) = J_{\neq}(\overline{\mu})$, *i.e.*, all multipliers associated to active inequality constraints are non-zero.

(1c) $\nabla_x^2 L(\overline{x},\overline{t},\overline{\lambda},\overline{\mu})|_{T_{\overline{x}}M(\overline{t})}$ is non-singular.

The previous definition means that $(\overline{x},\overline{t})$ is a non-degenerate critical point of $P(t)$. In view of Proposition 2.1.1, this can equivalently be expressed by the conditions

$$H := \nabla^2_{(x,\lambda,\mu)} L(\overline{x},\overline{t},\overline{\lambda},\overline{\mu}) \text{ is non-singular and } J_0(\overline{x},\overline{t}) = J_{\neq}(\overline{\mu}).$$

By using this result, we can apply the Implicit Function Theorem to the non-linear system that describes the critical point condition and show that, locally around $(\overline{x},\overline{t})$, the set of generalized critical points is a curve $(x(t),t)$ of non-degenerate critical points.

To define generalized critical points of type 2, we consider the problems

$$P^p(t): \qquad \begin{array}{rl} & \min\limits_x f(x,t) \\ \text{s.t.} & x \in M^p(t), \end{array}$$

where

$$M^p(t) = \left\{ x \in \mathbb{R}^n \;\middle|\; \begin{array}{ll} h_i(x,t) = 0, & i = 1,\ldots,q_0, \\ g_j(x,t) = 0, & j = 1,\ldots,p \end{array} \right\}$$

and

$$P^{p-1}(t): \qquad \begin{array}{rl} & \min\limits_x f(x,t) \\ \text{s.t.} & x \in M^{p-1}(t), \end{array}$$

with

$$M^{p-1}(t) = \left\{ x \in \mathbb{R}^n \;\middle|\; \begin{array}{ll} h_i(x,t) = 0, & i = 1,\ldots,q_0, \\ g_j(x,t) = 0, & j = 1,\ldots,p-1. \end{array} \right\}$$

Definition 2.4.3 $(\overline{x},\overline{t}) \in \Sigma_{gc}$ is *a generalized critical point of type 2,* $(\overline{x},\overline{t}) \in \sum^2_{gc}$, if:

(2a) LICQ holds at $(\overline{x},\overline{t})$.

(2b) $J_0(\overline{x},\overline{t})/J_{\neq}(\overline{\mu})$ consists of one index, w.l.o.g. $J_0(\overline{x},\overline{t})/J_{\neq}(\overline{\mu}) = \{p\}$.

(2c) $\overline{x}$ is a non-degenerate critical point of $P^p(\overline{t})$ and $P^{p-1}(\overline{t})$.

(2d) If $(x^{p-1}(t),t)$, denotes the curve of critical points of $P^{p-1}(t)$ near $\overline{t}$ then $D_t g_p(x^{p-1}(t),t)|_{t=\overline{t}} \neq 0$.

At a g.c. point of type 2, a bifurcation takes place in the set Σ_{gc}. There are two branches of critical points, one associated to problem $P^p(t)$ for $t \in [\overline{t}-\epsilon,\overline{t}+\epsilon]$ and the other corresponds to the feasible branch of critical points of problem $P^{p-1}(t)$, either for $t \in [\overline{t}-\epsilon,\overline{t}]$ or $t \in [\overline{t},\overline{t}+\epsilon]$.

Definition 2.4.4 We write $(\overline{x},\overline{t}) \in \Sigma^3_{gc}$ and say $(\overline{x},\overline{t})$ is *a g.c. point of type 3* if:

(3a) $(\overline{x},\overline{t})$ is a critical point of $P(t)$, with unique multipliers $(\overline{\lambda},\overline{\mu})$.

(3b) $J_0(\overline{x},\overline{t}) = J_{\neq}(\overline{\mu})$.

(3c) $(\overline{x},\overline{t},\overline{\lambda},\overline{\mu})$ is a non-degenerate critical point of

$$min\left\{t \mid \nabla_{(x,\lambda,\mu)}L(x,t,\lambda,\mu) = 0\right\}.$$

In such a point the matrix $\nabla^2_x L|_{T_{\overline{x}}M(\overline{t})}$ has exactly one eigenvalue equal to 0. Geometrically, condition (3c) implies that $(\overline{x},\overline{t})$ is a quadratic turning point in Σ_{gc}.

In the next two types of g.c. points the condition LICQ does not hold. For simplicity we introduce the notation

$$U(x,t) = (\nabla_x h_1(x,t),\ldots,\nabla_x h_{q_0}(x,t),\nabla_x g_1(x,t),\ldots,\nabla_x g_p(x,t)).$$

Definition 2.4.5 For a g.c. point $(\overline{x},\overline{t})$ of $P(t)$, we say $(\overline{x},\overline{t}) \in \Sigma^4_{gc}$, *i.e.,* $(\overline{x},\overline{t})$ is *a point of type 4* if:

(4a) $1 \le q_0 + p \le n$ and $rank(U(\overline{x},\overline{t})) = q_0 + p - 1$.

(4b) For all solutions $(\lambda,\mu) \in \mathbb{R}^{q_0+p}$ of $U(\overline{x},\overline{t})\begin{pmatrix}\lambda\\ \mu\end{pmatrix} = 0$, it follows $\mu_j \neq 0$, $j = 1,\ldots,p$.

(4c) Let us consider the function $H(x,t,\lambda,\mu,\lambda_0) = \nabla_{(x,\lambda,\mu)}\hat{L}(x,t,\lambda,\mu,\lambda_0)$, where $\hat{L}(x,t,\lambda,\mu,\lambda_0) = \lambda_0 f(x,t) - \sum_{i=1}^{q_0}\lambda_i h_i(x,t) - \sum_{j=1}^{p}\mu_j g_j(x,t)$ and let $(\overline{\lambda},\overline{\mu})$ be the unique solution of $U(\overline{x},\overline{t})\begin{pmatrix}\lambda\\ \mu\end{pmatrix} = 0$ with $\overline{\mu}_p = 1$. The point $(\overline{x},\overline{t},\overline{\lambda},\overline{\mu}_1,\ldots\overline{\mu}_{p-1},0)$ is a non-degenerate critical point of

$$\mathcal{G}: \quad \begin{array}{rl} & \min\limits_{x,t,\lambda,\mu_1,\ldots,\mu_{p-1},\lambda_0} t \\ \text{s.t.} & H(x,t,\lambda,\mu_1,\ldots,\mu_{p-1},1,\lambda_0) = 0. \end{array}$$

As in the case of a g.c. point of type 3, the points of type 4 are turning points of Σ_{gc}.

Definition 2.4.6 A g.c. point $(\overline{x},\overline{t})$ is *of type 5, i.e.* $(\overline{x},\overline{t}) \in \Sigma^5_{gc}$ if:

(5a) $q_0 + p = n + 1$ and $rank\begin{pmatrix} U(\overline{x},\overline{t}) \\ (\nabla_t h_1,\ldots,\nabla_t h_{q_0},\nabla_t g_1,\ldots,\nabla_t g_p)(\overline{x},\overline{t})\end{pmatrix} = n+1.$

(5b) For any solution $(\lambda, \mu) \in \mathbb{R}^{q_0} \times \mathbb{R}^p$ of $U(\overline{x}, \overline{t}) \begin{pmatrix} \lambda \\ \mu \end{pmatrix} = 0$ it holds $\mu_j \neq 0$ for $j = 1, \ldots, p$.

(5c) For any $(\overline{\lambda}, \overline{\mu}) \in \mathbb{R}^{q_0+p}$ such that $\nabla_x L(\overline{x}, \overline{t}, \overline{\lambda}, \overline{\mu}) = 0$, it follows $|J_0(\overline{x}, \overline{t})/J_{\neq}(\overline{\mu})| \leq 1$.

At a point of type 5 a bifurcation occurs in the following way: for $l = 1, 2, \ldots, p$ we define

$$P^l(t): \quad \begin{array}{rrcl} & \multicolumn{3}{c}{\min f(x,t)} \\ \text{s.t.} & h_i(x,t) & = & 0, \quad i = 1, \ldots, q_0, \\ & g_j(x,t) & = & 0, \quad j = 1, 2, \ldots, l-1, l+1, \ldots, p. \end{array}$$

Then there is some ε, $\varepsilon > 0$ such that exactly one of the sets

$$\Sigma_{gc}(P^l(t)) \cap \{(x,t) \in \mathbb{R}^n \times \mathbb{R} \mid x \in M(t), t \in (\overline{t}, \overline{t} + \varepsilon]\}$$

or

$$\Sigma_{gc}(P^l(t)) \cap \{(x,t) \in \mathbb{R}^n \times \mathbb{R} \mid x \in M(t), t \in [\overline{t} - \varepsilon, \overline{t})\}$$

is non-empty (and the other empty) around $(\overline{x}, \overline{t})$. Moreover, locally, $\Sigma_{gc}(P(t)) = \bigcup_{l=1}^{p} \left\{\Sigma_{gc}(P^l(t)) \cap \{(x,t) \in \mathbb{R}^n \times \mathbb{R} \mid x \in M(t)\}\right\}$.

Definition 2.4.7 We will say that a one-parametric problem, given by $(f, h, g) \in [C^3_S]^{1+q_0+q}_{n+1}$ is *JJT-regular on* $T \subset \mathbb{R}$ or that (f, h, g) *is in the class* $\mathcal{F}|_T$ if all its generalized critical points $(\overline{x}, \overline{t})$ with $\overline{t} \in T$, are of type 1, 2, 3, 4 or 5.

We end this chapter with the main genericity results for parametric optimization problems:

Theorem 2.4.1 (Genericity result, ***cf*** **[21])**
(a) Fix any parametric problem $P(t)$ (see (2.4.1)) and $T = [0,1]$ or $T = \mathbb{R}$. Consider the perturbed problems

$$P(A,b,c,d): \quad \begin{array}{rrcl} & \multicolumn{3}{l}{\min f(x,t) + x^T A x + b^T x} \\ s.t. & h_i(x,t) + c_i^T x + d_i & = & 0, \quad i = 1, \ldots, q_0, \\ & g_j(x,t) + c_{j+q_0}^T x + d_{j+q_0} & \geq & 0, \quad j = 1, \ldots, q \end{array}$$

where A is a symmetric $n \times n$-matrix, and $(b, c_1, \ldots c_{q_0+q}, d) \in \mathbb{R}^{n+n(q_0+q)+q_0+q}$. The set of perturbations (A, b, c, d) such that $P(A, b, c, d)$ is not JJT-regular on T has Lebesgue measure zero.
(b) The sets $\mathcal{F}|_{[0,1]}$ and $\mathcal{F}|_{\mathbb{R}}$ are open and dense with respect to the strong topology in $[C^3_S]^{1+q_0+q}_{n+1}$.

Remark 2.4.1 *The previous theorem means that the JJT-regularity is not a strong condition. It is stable under small perturbations and, for a problem defined by the functions* $P(t) = (f, h_1, \ldots h_{q_0}, g_1, \ldots, g_q)$*, there is a sequence of functions corresponding to JJT-regular problems* $P_k(t),\ k \in \mathbb{N}$*, converging to* $P(t)$ *with respect to the strong topology. Moreover we can find arbitrarily small quadratic and linear perturbations of the involved functions leading to regular problems.*

Chapter 3

Variational Inequality Problems

3.1 Introduction

This chapter is devoted to *Variational Inequalities (VI)*, *i.e.*, we consider the feasibility problem

$$VI: \qquad \begin{array}{c} \text{find } y \in Y \subset \mathbb{R}^m \\ \text{such that } \Phi(y)^T(z-y) \geq 0, \quad \forall z \in Y, \end{array}$$

and the one-parametric version

$$VI(t): \qquad \begin{array}{c} \text{for } t \in [0,1], \text{ find } y \in Y(t) \subset \mathbb{R}^m \\ \text{such that } \Phi(y,t)^T(z-y) \geq 0, \quad \forall z \in Y(t). \end{array}$$

These inequalities are particular cases of equilibrium constraints. A point solving VI, also called a feasible point of VI, can be obtained by means of a fixed point algorithm (Patrickson [45]), with the help of merit functions (Solodov [56]) or by applying regularization techniques (Ravindran and Gowda [48]).

Let us shortly discuss the problem of the existence of feasible solutions of VI. In general these problems may have no feasible solution. However convexity conditions lead to the following result.

Theorem 3.1.1 (Existence conditions, *cf.* **[45])** *If $Y \neq \varnothing$ is convex and $\Phi(y)$ is continuous on Y, then each of the following conditions is sufficient for the existence of a feasible solution of VI:*

1. *Y is compact.*
2. *Φ is coercive, i.e., $\exists y_0 \in Y$, such that $\lim\limits_{\|y\|\to\infty; y\in Y} \frac{\Phi(y)^T(y-y_0)}{\|y\|} = +\infty$ holds for any sequence of points $y \in Y$ with $\|y\| \to \infty$.*
3. *Φ is strongly monotone, i.e., $\exists \kappa,\ \kappa > 0$, such that for all $y_1, y_2 \in Y$ we have $(\Phi(y_1) - \Phi(y_2))^T(y_1 - y_2) \geq \kappa \|y_1 - y_2\|^2$.*

Proof. For a proof, which is based on the fixed point theory, we refer to [45]. See also [7].

$\square$

The chapter is organized as follows. We firstly apply the KKT approach to VI. Then in Section 3.3 we will outline the genericity results of [19] for the one parametric problem $VI(t)$. In the last section we analyze embedding procedures for solving VI and obtain genericity results for these approaches.

3.2 The KKT approach for variational inequalities

Let us consider the variational inequality problem

$$VI: \quad \begin{array}{c} \text{find } y \in Y \\ \text{such that } \Phi(y)^T(z-y) \geq 0, \quad \forall z \in Y, \end{array} \tag{3.2.1}$$

where the set Y is defined by:

$$Y = \left\{ y \in \mathbb{R}^m \;\middle|\; \begin{array}{rcl} h_i(y) & = & 0, \quad i = 1, \ldots, q_0, \\ g_j(y) & \geq & 0, \quad j = 1, \ldots, q. \end{array} \right\} \tag{3.2.2}$$

We denote this problem more explicitly by $VI(\Phi, Y)$ or by $VI(\Phi,\ h,\ g)$, where $h = (h_1, \ldots, h_{q_0})$ and $g = (g_1, \ldots, g_q)$.

Throughout the chapter we assume $Y \neq \varnothing$ and $(\Phi, h, g) \in [C^3]_m^{m+q_0+q}$. The points y satisfying condition (3.2.1) will be called feasible solutions of VI.

As it was already discussed, see Section 1.2, in view of the relation $\Phi(y)^T(y-y) = 0$, a point $\overline{y} \in Y$ is feasible for VI if and only if it solves the optimization problem

$$Q(\overline{y}): \quad \begin{array}{cc} & \min\limits_{z} \Phi(\overline{y})^T(z-\overline{y}) \\ \text{s.t.} & z \in Y. \end{array} \tag{3.2.3}$$

So a solution $\overline{y}$ of VI necessarily satisfies a Fritz John condition for the problem (3.2.3). With the active index set $J_0(y) = \{j \mid g_j(y) = 0\}$, and the function

$$\overline{L}(y, \lambda_0, \lambda, \mu) = \lambda_0 \Phi(y) - \sum_{i=1}^{q_0} \lambda_i \nabla h_i(y) - \sum_{j \in J_0(y)} \mu_j \nabla g_j(y),$$

corresponding to the derivative with respect to z of the Lagrangean of problem $Q(\overline{y})$, the necessary optimality conditions are summarized in the following proposition.

Theorem 3.2.1 (Necessary feasibility condition, *cf.* **[19])** *Let $\overline{y}$ be a feasible solution of $VI(\Phi, h, g)$. Then, there are multipliers λ and $\lambda_0,\ \mu \geq 0$ not all zero such that $\overline{L}(\overline{y}, \lambda_0, \lambda, \mu) = 0$. Moreover the second order optimality condition holds:*

$$\xi^T \left(-\sum_{i=1}^{q_0} \lambda_i \nabla_y^2 h_i(\overline{y}) - \sum_{j \in J_0(\overline{y})} \mu_j \nabla_y^2 g_j(\overline{y}) \right) \xi \geq 0 \ , \quad \forall \xi \in T_{\overline{y}} Y.$$

If LICQ or MFCQ holds at $\overline{y}$, then the KKT condition is satisfied, i.e., we can assume $\lambda_0 = 1$ in $\overline{L}(\overline{y}, \lambda_0, \lambda, \mu) = 0$.

Recall that if the problem (3.2.3) is convex (*i.e.*, the functions h_i are linear $i = 1, \ldots, q_0$ and $-g_j$ are convex, $j = 1, \ldots, q$) then the KKT condition $\overline{L}(\overline{y}, 1, \lambda, \mu) = 0$, $\mu \geq 0$, is sufficient for $\overline{y}$ to be a solution of (3.2.3). In view of Theorem 3.2.1, if in addition MFCQ holds, the KKT condition and the optimality condition for $Q(\overline{y})$ are equivalent.

Definition 3.2.1 For $\overline{y} \in Y$ we write $\overline{y} \in \Sigma_{gc}$, *i.e*, $\overline{y}$ is *a generalized critical point for VI*, if there exist $\lambda_0,\ \lambda_i,\ i = 1, \ldots, q_0,\ \mu_j,\ j \in J_0(\overline{y})$ not all zero such that $\overline{L}(\overline{y}, \lambda_0, \lambda, \mu) = 0$.
We write $\overline{y} \in \Sigma_{crit}$ if $\overline{y} \in \Sigma_{gc}$ and LICQ holds at $\overline{y} \in Y$. In this case we consider the unique multipliers (λ, μ) such that $\overline{L}(\overline{y}, 1, \lambda, \mu) = 0$. The notation $\overline{y} \in \Sigma_{stat}$ means that $\overline{y} \in \Sigma_{crit}$ and $\mu_j \geq 0$.

Definition 3.2.2 The point $\overline{y} \in \Sigma_{gc}$ is said to be *a non-degenerate critical point,* denoted as $\overline{y} \in \Sigma_{gc}^1$ if:

VI-1a: LICQ holds at $\overline{y}$.
So, there exist unique multipliers (λ, μ) such that $\overline{L}(\overline{y}, 1, \lambda, \mu) = 0$.

VI-1b: $\mu_j \neq 0$ for all $j \in J_0(\overline{y})$.

VI-1c: $\nabla_y \overline{L}(\overline{y}, 1, \lambda, \mu) \mid_{T_{\overline{y}} Y}$ is non-singular.

We say that $VI(\Phi, h, g)$ is *regular* if LICQ holds for all $y \in Y$ and all solutions of $VI(\Phi, h, g)$ satisfy VI-1a, VI-1b, VI-1c.

As in the case of nonlinear problems, it can be shown that, generically for $(\Phi, h, g) \in [C_s^2]_{n+1}^{n+q_0+q}$, the problem $VI(\Phi, h, g)$ is regular.

Remark 3.2.1 Under the conditions of Definition 3.2.2, as in standard finite optimization, the point $\overline{y}$ will be an isolated critical point.

We emphasize that in contrast to standard optimization, at a solution $\overline{y}$ of VI due to the term $\nabla_y \Phi(y)$, the second order matrix $\nabla_y \overline{L}|_{T_{\overline{y}}Y}$ does not need to be symmetric. Moreover, the condition $\nabla_y \overline{L}|_{T_{\overline{y}}Y} \succeq 0$ is not a second order necessary feasibility condition because at a solution $\overline{y}$, negative or even complex eigenvalues may appear as is shown in the following example.

Example 3.2.1 Consider $VI(\Phi, \mathbb{R}^3)$, see (3.2.1), with $\Phi(y) = \begin{pmatrix} -y_1 \\ y_2 - y_3 \\ 2y_2 + y_3 \end{pmatrix}$.

The point $\overline{y} = 0$ is the solution of the problem. However

$$\nabla_y \overline{L}(0)|_{\mathbb{R}^3} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 2 & 1 \end{pmatrix}$$

has negative and non-real eigenvalues.

3.2.1 Relations between Stampaggia and Minty variational inequalities

In this subsection, we consider the relations between two classical variational inequalities, the Stampaggia VI, in (3.2.1)

$$VI_S : \quad \begin{array}{c} \text{find } y \in Y \\ \text{such that} \quad \phi_S(y,z) := \Phi(y)^T(z-y) \geq 0, \quad \forall z \in Y, \end{array} \tag{3.2.4}$$

and the Minty VI

$$VI_M : \quad \begin{array}{c} \text{find } y \in Y \\ \text{such that} \quad \phi_M(y,z) := \Phi(z)^T(z-y) \geq 0, \quad \forall z \in Y. \end{array} \tag{3.2.5}$$

For details the reader is refereed to Kassay [35]. Assume that Y is defined by (3.2.2) and consider the associated problem

$$Q(y) : \quad \begin{array}{cc} & \min_z \phi(y,z) \\ \text{s.t.} & z \in Y. \end{array}$$

Let us apply the KKT approach. For both functions ϕ_S and ϕ_M we obtain $\nabla_z \phi_S(y,z)|_{z=y} = \Phi(y)$ and $\nabla_z \phi_M(y,z)|_{z=y} = \Phi(y) + \nabla_y \Phi(y)(z-y)|_{z=y} = \Phi(y)$. So the Stampaggia and the Minty variational inequalities lead to the same KKT system:

$$\begin{array}{rcl} \Phi(y) - \sum\limits_{i=1}^{q_0} \lambda_i \nabla h_i(y) - \sum\limits_{j=1}^{q} \mu_j \nabla g_j(y) & = & 0, \\ h_i(y) & = & 0, \quad i = 1, \ldots, q_0, \\ g_j(y) & \geq & 0, \quad j = 1, \ldots, q, \\ \mu_j & \geq & 0, \quad j = 1, \ldots, q, \\ g_j(y)\mu_j & = & 0, \quad j = 1, \ldots, q. \end{array} \tag{3.2.6}$$

Let us discuss the relations between these two variational inequalities. Assume that Y is convex and satisfies a constraint qualification. Since the function ϕ_S is linear in z, $(\overline{y}, \lambda, \mu)$ solves the KKT system (3.2.6) if and only if $\overline{y}$ is feasible for the Stampaggia variational inequality. In general the function ϕ_M is not convex in z. But in case it is, the fact that $\overline{y}$ solves the KKT system (3.2.6) with some (λ, μ) is equivalent to have $\overline{y}$ solving both, the Stampaggia and the Minty variational inequality.

It is easy to see that if Φ is monotone, *i.e.*, if

$$\Phi(z)^T(z-y) \geq \Phi(y)^T(z-y), \quad \forall y, z \in Y,$$

then each solution of VI_S is a solution of VI_M and vice versa. More precisely the following relations (partially proved in [35]) hold.

Lemma 3.2.1 *Let the set Y be convex. Then:*
(a) Any solution of VI_M is a solution of VI_S.
(b) Let $\overline{y}$ be a solution of VI_S and assume that the (partial) monotonicity condition:

$$\Phi(z)^T(z-\overline{y}) \geq \Phi(\overline{y})^T(z-\overline{y}), \quad \forall z \in Y, \tag{3.2.7}$$

holds. Then $\overline{y}$ is a solution of VI_M.
(c) Let $\overline{y}$ be a solution of VI_M (and thus of VI_S) and let the function $\phi_M(\overline{y}, z) = \Phi(z)^T(z-\overline{y})$ be convex in z. Then the condition (3.2.7) is satisfied.

Proof. (a) Let $\overline{y}$ be a solution of VI_M, *i.e.*,

$$\Phi(z)^T(z-\overline{y}) \geq 0, \ \forall z \in Y.$$

Take any point $v \in Y$ and consider $z(\alpha) = \alpha\overline{y} + (1-\alpha)v$. As Y is convex, if $\alpha \in (0,1)$, it follows that $z(\alpha) \in Y$ and

$$\Phi(\alpha\overline{y}+(1-\alpha)v)^T(\alpha\overline{y}+(1-\alpha)v-\overline{y}) = (1-\alpha)\Phi(\alpha\overline{y}+(1-\alpha)v)^T(v-\overline{y}) \geq 0, \qquad \forall v \in Y.$$

Dividing by $1-\alpha$ and letting $\alpha \to 1^-$ it follows that $\Phi(\overline{y})^T(v-\overline{y}) \geq 0, \ \forall v \in Y$.

(b) For a solution $\overline{y}$ of VI_S under (3.2.7) we directly obtain:

$$\Phi(z)^T(z-\overline{y}) \geq \Phi(\overline{y})^T(z-\overline{y}) \geq 0, \quad \forall z \in Y.$$

(c) If the function $\Phi(z)^T(z-\overline{y})$ is convex in z, then for all $\alpha \in (0,1)$ and $z \in Y$,

$$\Phi(\alpha\overline{y}+(1-\alpha)z)^T(\alpha\overline{y}+(1-\alpha)z-\overline{y}) \leq \alpha\Phi(\overline{y})^T(\overline{y}-\overline{y}) + (1-\alpha)\Phi(z)^T(z-\overline{y}).$$

So for all $\alpha \in (0,1)$ we find

$$(1-\alpha)\Phi(\alpha\overline{y}+(1-\alpha)z)^T(z-\overline{y}) \leq (1-\alpha)\Phi(z)^T(z-\overline{y}).$$

Dividing by $(1-\alpha)$ and letting $\alpha \uparrow 1$ yields the monotonicity property,

$$\Phi(\overline{y})^T(z-\overline{y}) \le \Phi(z)^T(z-\overline{y}), \text{ for all } z \in Y.$$

□

The next example shows that the converse of Lemma 3.2.1(c) is not necessarily true, *i.e.*, the monotonicity condition (3.2.7) does not necessarily imply the convexity of the function $\phi_M(y,z)$ w.r.t. z.

Example 3.2.2 Consider the Minty inequality with the function $\Phi(y) = \sin y$ and $Y = [-\frac{\pi}{2}, \frac{\pi}{2}]$. The unique solution $y \in Y$ of:

$$\sin(z) \cdot (z-y) \ge 0, \quad \forall z \in Y,$$

is given by $\overline{y} = 0$.

As $\Phi'(y) = \cos(y) \ge 0$ on Y, the function $\Phi(y)$ is monotonically increasing on the interval $[-\frac{\pi}{2}, \frac{\pi}{2}]$. So the monotonicity relation $(\sin(z) - \sin(y))(z-y) \ge 0$ holds. However, the function $\phi(\overline{y}, z)_M = \sin(z)(z-\overline{y})$ is not convex in $z \in Y$. To see this, note that a C^2-function ϕ is convex on Y if and only if $\phi''(z) \ge 0$, for $z \in Y$. Differentiating w.r.t. z yields

$$\nabla_z^2 \phi_M(\overline{y}, z) = 2\cos(z) - \sin(z)(z-\overline{y}),$$

and we see that the second derivative is negative for $z = \frac{\pi}{2}$.

3.3 One-parametric variational inequalities

In this section we shortly describe the genericity results of [19] for one parametric variational inequalities. So we consider the parametric VI

$$VI(t): \quad \begin{array}{rl} & \text{find } y \in Y(t) \\ \text{such that} & \Phi(y,t)^T(z-y) \ge 0, \quad \forall z \in Y(t), \end{array} \tag{3.3.1}$$

depending on the variable $t \in T \subset \mathbb{R}$, where the set $Y(t)$ is defined by

$$Y(t) = \left\{ y \in \mathbb{R}^m \;\middle|\; \begin{array}{rcll} h_i(y,t) & = & 0, & i = 1, \dots, q_0, \\ g_j(y,t) & \ge & 0, & j = 1, \dots, q \end{array} \right\} \tag{3.3.2}$$

and $(\Phi, h, g) \in [C^3]_{m+1}^{m+q_0+q}$, with $h = (h_1, \dots, h_{q_0})$ and $g = (g_1, \dots, g_q)$. $VI(t)$ is also denoted as $VI(t; \Phi, Y(t))$ or as $VI(t; \Phi, h, g)$.

We will assume that T is a compact connected set, w.l.o.g., $T = [0,1]$. We say that a point $(y,t) \in \mathbb{R}^m \times \mathbb{R}$ is feasible for $VI(t)$ if y is feasible for the problem $VI(t)$. In the same way, we can extend the other definitions of Section 3.2 and speak about stationary points, non-degenerate critical points and generalized critical points $(y,t) \in \mathbb{R}^{m+1}$ of $VI(t)$ (for details we refer to [19]).

The KKT approach for solving the parametric VI (*cf.* Sectioon 3.2) leads us to a one-parametric KKT system:

$$\begin{array}{rcll} \overline{L}(y,t,1,\lambda,\mu) & = & 0, & \\ h_i(y,t) & = & 0, & i=1,\dots,q_0, \\ g_j(y,t) & \geq & 0, & j=1,\dots,q, \\ \mu_j & \geq & 0, & j=1,\dots,q, \\ g_j(y,t)\mu_j & = & 0, & j=1,\dots,q. \end{array} \tag{3.3.3}$$

where $\overline{L}(y,t,1,\lambda,\mu) = \Phi(y,t) - \sum\limits_{i=1}^{q_0} \lambda_i \nabla_y h_i(y,t) - \sum\limits_{j=1}^{q} \mu_j \nabla_y g_j(y,t)$.
Near a non-degenerate critical point $(\overline{y},\overline{t})$ of $VI(t)$, also called g.c. point of type 1, there exists a unique curve $(y(t),t)$ of non-degenerate critical points, with $(y(\overline{t}),\overline{t}) = (\overline{y},\overline{t})$.

We are interested in the types of degeneracies which generically may occur in the set of generalized critical points Σ_{gc} of $VI(t)$. Extending the singularities appearing in one-parametric finite programming (see [27], [28]) 4 types of degenerate generalized critical points $(\overline{y},\overline{t})$ were defined for $VI(t)$ in [19]. Roughly speaking at the singular points the following takes place:

- G.C. points of type 2, VI-2: Condition VI-1b does not hold.

- G.C. points of type 3, VI-3: Condition VI-1c fails.

- G.C. points of type 4, VI-4: LICQ does not hold at $(\overline{y},\overline{t})$ w.r.t. $Y(\overline{t})$ and $q_0 + |J_0(\overline{y},\overline{t})| \leq m$.

- G.C. points of type 5, VI-5: LICQ does not hold at $(\overline{y},\overline{t})$ w.r.t. $Y(\overline{t})$ and $q_0 + |J_0(\overline{y},\overline{t})| = m+1$.

Here we have only listed the condition which is violated in each kind of singularity. In all cases, the defined types have to fulfill additional properties. For a complete definition of the types and their properties we refer to [19].
For $VI(t)$ we denote the set of generalized critical points $(\overline{y},\overline{t})$ of type i by Σ_{gc}^i. Let T be a subset of $\mathbb{R}$. A VI problem where all its g.c. points $(y,t),\ t \in T$ are of type 1, 2, 3, 4 or 5, is called regular on T. In terms of the defining functions $(\Phi(y,t), h(y,t), g(y,t))$ the set of regular one-parametric variational inequalities is:

$$\mathcal{F}_{VI(t)}|_T = \left\{ (\Phi,h,g) \in [C_S^3]_{m+1}^{m+q_0+q} \;\middle|\; \Sigma_{gc}(VI(t;\Phi,h,g)) \bigcap [\mathbb{R}^m \times T] \subset \cup_{i=1}^5 \Sigma_{gc}^i \right\}.$$

The following result has been shown in [19].

Theorem 3.3.1 (*cf.* **[19]**) *Given* $(\Phi(y,t), h(y,t), g(y,t)) \in [C_S^3]_{m+1}^{m+q_0+q}$*, for almost all* $(A, b, C_h, d_h, C_g, d_g) \in \mathbb{R}^{m^2+m+q_0m+q_0+qm+q}$ *it holds that*

$$(\Phi(y,t) + Ay + b, h(y,t) + [C_h y + d_h]^T, g(y,t) + [C_g y + d_g]^T) \in \mathcal{F}_{VI(t)}|_{[0,1]}.$$

Furthermore, the set $\mathcal{F}_{VI(t)}|_{[0,1]}$ *is open and dense with respect to the topology in* $[C_S^3]_{m+1}^{m+q_0+q}$.

In Figure 3.1 (see [19]) the local structure of Σ_{stat} and Σ_{gc} is sketched around the 5 types of g.c. points appearing in the generic case.

In particular at points of type 5 where MFCQ fails, there exists a neighborhood U of $\overline{y}$ and $\delta > 0$ such that for all $\epsilon \in (0, \delta)$ either $Y(\bar{t} + \epsilon) \cap U = \varnothing$ or $Y(\bar{t} - \epsilon) \cap U = \varnothing$
Under additional convexity assumptions the points of type 3 are excluded in the set Σ_{stat}.

Proposition 3.3.1 *Let* $h_i(y,t) = c_i^T(t)y + d_i(t),\ i = 1, \ldots, q_0$*, and let, for any* t*,* $-g_j(y,t),\ j = 1, \ldots, q$*, be convex in* y*. If* $\Phi_y(y,t) \succ 0$ *for all* (y,t)*, then for the corresponding problem* $VI(t)$ *it follows that* $\Sigma_{stat} \cap \Sigma_{gc}^3 = \varnothing$*.*

Proof. By assumption

$$\nabla_y \overline{L} = \nabla_y \left(\Phi - \sum_{i=1}^{q_0} \lambda_i \nabla_y h_i - \sum_{j \in J_0(y,t)} \mu_j \nabla_y g_j \right) \succ 0$$

at all (y, t, λ, μ) with $\mu \geq 0$. So, the matrix $\nabla_y \overline{L}|_{T_y Y(t)}$ is regular and singular points of type 3 are excluded.

□

Remark 3.3.1 *In particular if for all* $t \in [0,1]$ *the set* $Y(t)$ *is convex and* $\Phi(y,t)$ *is strongly monotone for* $y \in \mathbb{R}^m$ *(i.e., there is some* $\kappa,\ \kappa > 0$*, such that* $\Phi(y_1,t)^T(y_1 - y_2) - \Phi(y_2,t)^T(y_1 - y_2) \geq \kappa \|y_1 - y_2\|^2,\ \forall y_1, y_2 \in \mathbb{R}^m$*) it is not difficult to show that* $\nabla_y \Phi(y,t) \succ 0,\ \forall (y,t)$*. So, we have* $\Sigma_{stat} \cap \Sigma_{gc}^3 = \varnothing$*.*

3.4 Embeddings for variational inequalities

The idea of an embedding approach to solve a non-parametric optimization problem P, is to construct a one-parametric problem $P(t)$, $t \in [0,1]$, with end problem $P(1) = P$ and an easy starting problem $P(0)$. Then, by using continuation methods, we try to follow the solutions of $P(t)$ from $t = 0$ to $t = 1$. We adapt this approach to solve non-parametric variational inequalities.

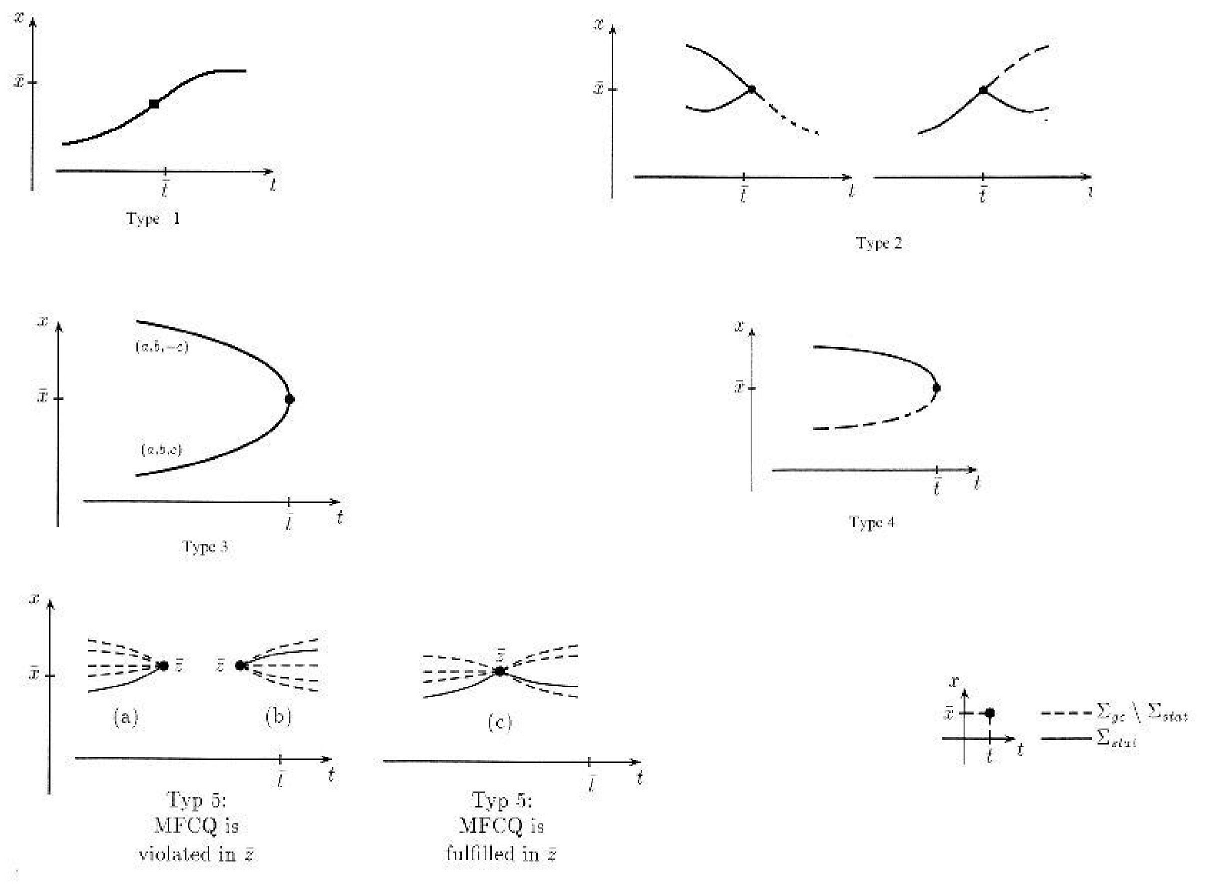


Figure 3.1: The behavior of Σ_{stat} around the singularities

Let be given a non-parametric VI problem $VI(\Phi(y), h(y), g(y))$ in (3.2.1) defined by C^3-functions $(\Phi, h, g) : \mathbb{R}^m \to \mathbb{R}^{m+q_0+q}$. We try to construct one-parameter depending functions $(\hat{\Phi}, \hat{h}_1, \ldots, \hat{h}_{q_0}, \hat{g}_1, \ldots, \hat{g}_q) : \mathbb{R}^{m_1} \times \mathbb{R} \to \mathbb{R}^{m_1+q_{01}+q_1}$ such that the corresponding parametric problem $VI(t) = VI(t; \hat{\Phi}(y,t), \hat{h}(y,t), \hat{g}(y,t))$ satisfies:

- $VI(0)$ has a trivial solution.

- For all $t \in [0, 1]$, the problem $VI(t)$ has a solution.

- $VI(1)$ is equivalent to $VI(\Phi, h, g)$.

Under the assumption that all generalized critical points of $VI(t)$ are of type 1, there exists a solution curve $(y(t), t)$, $t \in [0, 1]$, which can be followed by continuation methods. However this assumption is not generically satisfied. So, we will consider this approach under the weaker regularity assumption that the functions $(\hat{\Phi}, \hat{h}, \hat{g})$, defining the parametric embedding $VI(t)$, are contained in the generic subset $\mathcal{F}_{VI(t)}|_{(0,1)}$ introduced in the previous section.

We will discuss two different approaches, the standard embedding and the penalty embedding. For both methods we will prove genericity results similar to the general results in Theorem 3.3.1.

3.4.1 Standard embedding

Consider the functions (Φ, h, g) and the associated non-parametric problem $VI(\Phi, h, g)$, see (3.2.1). The standard embedding is defined by functions of the form

$$\Psi_S(t; \Phi, h, g) = \begin{pmatrix} t\Phi(y) + (1-t)(y - y_0) \\ th_i(y) + (1-t), \ i = 1, \ldots, q_0 \\ -t\sum_{i=1}^{q_0} h_i(y) + (1-t) \\ tg_j(y) + (1-t), \ j = 1, \ldots, q \end{pmatrix}$$

and leads to the parametric variational inequality problem

$$VI_S(t; \Phi, h, g) : \quad \begin{array}{c} \text{for } t \in [0,1], \text{ find } y \in Y_S(t) \\ \text{such that } (t\Phi(y) + (1-t)(y - y_0))^T(z - y) \geq 0, \ \forall z \in Y_S(t), \end{array} \tag{3.4.1}$$

where the sets $Y_S(t)$ are given by, (recall $h = (h_1, \ldots, h_{q_0})$ and $g = (g_1, \ldots, g_q)$):

$$Y_S(t) = \left\{ y \in \mathbb{R}^m \;\middle|\; \begin{array}{rcl} th_i(y) + (1-t) & \geq & 0, \quad i = 1, \ldots, q_0, \\ -t\sum_{i=1}^{q_0} h_i(y) + (1-t) & \geq & 0, \\ tg_j(y) + (1-t) & \geq & 0, \quad j = 1, \ldots, q. \end{array} \right\}$$

Clearly, for $t = 0$ the point $y_0 \in \mathbb{R}^m$ is a feasible starting point. Note that $Y_S(1)$ coincides with the set of feasible solutions of $VI(\Phi, h, g)$ because the original constraints $h_i(y) = 0$ can be written equivalently as

$$h_i(y) \geq 0,\ i = 1, \ldots, q_0, \text{ and } -\sum_{i=1}^{q_0} h_i(y) \geq 0.$$

So $VI_S(1; \Phi, h, g)$ coincides with $VI(\Phi, h, g)$. We refer to Schmidt [50] for a study of similar embeddings for solving standard mathematical programs. For our embedding we can prove the following genericity result.

Proposition 3.4.1 *The set* $I = \left\{(\Phi, h, g) \mid \Psi_S(t; \Phi, h, g) \in \mathcal{F}_{VI(t)}|_{t\in(0,1)}\right\}$ *is a generic set in* $[C^3_S]^{m+q_0+q}_m$.

Proof. Firstly we prove that for any fixed $k > 2$, the sets

$$I_k = \left\{(\Phi, h, g) \mid \Psi_S(t; \Phi, h, g) \in \mathcal{F}_{VI(t)}|_{t\in[\frac{1}{k}, 1-\frac{1}{k}]}\right\}$$

are open and dense in $[C^3_S]^{m+q_0+q}_m$.

<u>I_k is open in $[C^3_S]^{m+q_0+q}_m$</u>: Let $(\overline{\Phi}, \overline{h}, \overline{g}) \in I_k$. By Theorem 3.3.1, $\mathcal{F}_{VI(t)}|_{[0,1]}$ is open, and it can be proven that $\mathcal{F}_{VI(t)}|_{[a,b]}$ is also open for all $a, b,\ 0 < a < b < 1$. So, there is a strong neighborhood $U \subset [C^3_S]^{m+q_0+1+q}_{m+1}$ of $\Psi_S(t; \overline{\Phi}, \overline{h}, \overline{g})$ (defined by a continuous function $\varepsilon(x, t) : \mathbb{R}^{m+1} \to \mathbb{R}_{++}$) such that $U \subset \mathcal{F}_{VI(t)}|_{t\in[\frac{1}{k}, 1-\frac{1}{k}]}$. Clearly, $(\Phi, h_1, \ldots h_{q_0}, h_0, g_1, \ldots, g_q)(y, t) \in [C^3_S]^{m+q_0+1+q}_{m+1}$ is in U if and only if for all $(y, t) \in \mathbb{R}^m \times [\frac{1}{k}, 1 - \frac{1}{k}]$

$$\begin{array}{rcll}
\|\Phi(y,t) - \left[t\overline{\Phi}(y) + (1-t)(y - y_0)\right]\| & < & \varepsilon(y,t), & \\
\|h_i(y,t) - \left[t\overline{h}_i(y) + (1-t)\right]\| & < & \varepsilon(y,t), & i = 1, \ldots, q_0, \\
\|h_0(y,t) - \left[-t\displaystyle\sum_{i=1}^{q_0}\overline{h}_i(y) + (1-t)\right]\| & < & \varepsilon(y,t), & \\
\|g_j(y,t) - \left[t\overline{g}_j(y) + (1-t)\right]\| & < & \varepsilon(y,t), & j = 1, \ldots, q,
\end{array}$$

and analogous relations hold for the first and second order partial derivatives.

Now we consider an open neighborhood of $(\overline{\Phi}, \overline{h}, \overline{g}) \in [C^3]^m_m \times [C^3]^{q_0}_m \times [C^3]^q_m$ defined by

$$\hat{\varepsilon}(y) = \begin{cases} \displaystyle\min_{t\in[\frac{1}{k}, 1-\frac{1}{k}]} \frac{\varepsilon(y,t)}{q_0} & \text{if } q_0 \neq 0, \\ \displaystyle\min_{t\in[\frac{1}{k}, 1-\frac{1}{k}]} \varepsilon(y,t) & \text{if } q_0 = 0. \end{cases}$$

As the minimum is taken over a compact set and $\varepsilon(y, t)$ is a continuous and positive function, also $\hat{\varepsilon}(y)$ will be a positive and continuous function of y. Let

(Φ, h, g) be an element in the neighborhood of $(\overline{\Phi}, \overline{h}, \overline{g})$ defined by $\hat{\varepsilon}(y)$. We claim that $\Psi_S(t; \Phi, h, g) \in U$.
To show this, for $q_0 \geq 1$ we obtain, for $t \in [\frac{1}{k}, 1 - \frac{1}{k}]$:

$$\begin{aligned} \|t\Phi(y) + (1-t)(y - y_0) - \left[t\overline{\Phi}(y) + (1-t)(y - y_0)\right]\| &= t\|\Phi(y) - \overline{\Phi}(y)\| \\ &< t\hat{\varepsilon}(y) \leq \tfrac{\varepsilon(y,t)}{q_0} \leq \varepsilon(y,t) \end{aligned}$$

and for $q_0 = 0$,

$$\|t\Phi(y) + (1-t)(y - y_0) - \left[t\overline{\Phi}(y) + (1-t)(y - y_0)\right]\| < t\hat{\varepsilon}(y) \leq \varepsilon(y,t).$$

Similarly, it is easy to see that $\|th_i(y) + (1-t) - \left[t\overline{h}_i(y) + (1-t)\right]\| < \varepsilon(y,t)$ and $\|tg_j(y) + (1-t) - \left[t\overline{g}_j(y) + (1-t)\right]\| < \varepsilon(y,t)$. The partial derivatives of first and second order of (Φ, h, g) satisfy an analogous inequality.
For $q_0 = 0$, the proof is completed. In the other cases we also have to consider the bound ($t \in (\frac{1}{k}, 1 - \frac{1}{k})$)

$$\begin{aligned} \| - t\sum_{i=1}^{q_0} h_i(y) + (1-t) - \left[-t\sum_{i=1}^{q_0} \overline{h}_i(y) + (1-t)\right] \| &= t\|\sum_{i=1}^{q_0} h_i(y) - \sum_{i=1}^{q_0} \overline{h}_i(y)\| \\ &\leq t\sum_{i=1}^{q_0} \|h_i(y) - \overline{h}_i(y)\| \\ &< q_0\hat{\varepsilon}(y) \leq \varepsilon(y,t). \end{aligned}$$

That means, we have found a strong neighborhood $\hat{U}$ of $(\overline{\Phi}, \overline{h}, \overline{g})$ given by $\hat{\varepsilon}(y)$ such that $\hat{U} \subset I_k$, hence, I_k is open.

$\underline{I_k \text{ is dense in } [C^3_S]^{m+q_0+q}_m}$: To show this we will first fix the functions (Φ, h, g) and prove that for almost all $(A, b, C_h, d_h, C_g, d_g)$ it holds that $(\Phi + Ay + b, h + [C_h y + d_h]^T, g + [C_g y + d_g]^T) \in I_k$.
We begin by considering the g.c. points (y, t) where $LICQ$ fails, *i.e.*, there exists $\mu \neq 0$ such that:

$$\sum_{j \in J_0(y,t)} \mu_j \nabla_y \hat{g}_j(y,t) = 0 \tag{3.4.2}$$

where $\hat{g}_j(y,t) = th_j(y) + (1-t)$ if $j = 1, \ldots, q_0$, $\hat{g}_{q_0+1}(y,t) = -t\sum_{i=1}^{q_0} h_i(y) + (1-t)$ and $\hat{g}_{j+q_0+1}(y,t) = tg_j(y) + (1-t)$ if $j = 1, \ldots, q$. We will show that for almost all (C_h, d_h, C_g, d_g) for the corresponding perturbed problem it holds:

(a) LICQ fails only in a discrete set of feasible points (y, t) with $t \in [\frac{1}{k}, 1 - \frac{1}{k}]$ (we will denote this set by Y_0).

(b) For all $(y, t) \in Y_0$ and (y, t, μ) solving (3.4.2), it holds that $\mu_j \neq 0$ for all $j \in J_0(y, t)$.

(c) For all $(y,t) \in Y_0$ and (y,t,μ) solving (3.4.2) and $j^* \in J_0(y,t)$ the matrix $\begin{pmatrix} \sum_{j\in J_0(y,t)} \mu_j \nabla_{(y,t)}[\nabla_y \hat{g}_j(y,t)]^T & \nabla_{(y,t)}\hat{g}_{J_0(y,t)}(y,t) \\ [\nabla_y \hat{g}_{J_0(y,t)\setminus\{j^*\}}]^T(y,t) & 0 \end{pmatrix}$ is non-singular.

These conditions will guarantee the fulfillment of conditions $(4a)$-$(4b)$, $(5a)$-$(5b)$ in Definition 2.4.5 and Definition 2.4.6. Note that the feasible set $Y_S(t)$ has a special structure because the same functions h_i appear also in the $(q_0+1)^{th}$-constraint. For $y \in Y_S(t),\ 0 < t < 1$, the first q_0+1 inequality constraints cannot be active simultaneously. Indeed if the first q_0 constraints are active,

$$th_i(y) + (1-t) = 0,\ i = 1,\dots,q_0,$$

then

$$-t\sum_{i=1}^{q_0} h_i(y) + (1-t) = (q_0+1)(1-t) > 0.$$

For $i = 1,\dots,q_0+1$, we consider the sets $Y_{i,S}(t)$ which are obtained from $Y_S(t)$ by skipping the i^{th}-inequality. Then, in particular, the following holds: for all $t \in \left[\frac{1}{k}, 1-\frac{1}{k}\right]$ if $y \in Y_S(t)$ then $y \in Y_{i,S}(t)$ for some $i \in \{1,\dots,q_0+1\}$, and the i^{th}-inequality is strictly positive at y.
Fixing $i \in \{1,\dots,q_0+1\}$, and following the ideas of the proof of Lemma 6.17 pp. 119, in [21], we have that for almost all perturbations (C_h, d_h, C_g, d_g), the set $Y_0 \cap \{(y,t) \mid y \in Y_{iS}(t)\}$ is a discrete set, see (a), and for its elements, conditions (b)-(c) hold, *i.e.*, for all (y,t), with $y \in Y(t),\ t \in \left[\frac{1}{k}, 1-\frac{1}{k}\right]$ and $J_0(y,t) \subset \{1,2,\dots,i-1,i+1,q_0+1+q\}$, if (y,t,μ) solves (3.4.2), then $\mu_j \neq 0,\ \forall j \in J_0(y,t)$. Now if we consider all possible indices $i = 1,\dots,q_0+1$ and intersect the resulting sets of perturbations, we find that, for almost all parameters (C_h, d_h, C_g, d_g), conditions (a)-(b)-(c) are fulfilled for $Y_S(t)$, and this leads to the desired result.

Now we fix the parameters (C_h, d_h, C_g, d_g), and thus the feasible set, such that the resulting perturbed problem satisfies conditions (a)-(b)-(c). Following the same lines of the proof of Theorem 2.4.1 (see Theorem 6.18 pp. 121 in [21]) we can prove that for almost all (A,b) for the associated perturbed problem the feasible points where LICQ fails are g.c. points of type 4 or 5 and the g.c. points where *LICQ* holds are of type 1, 2 or 3. The perturbation result is now a consequence of the Fubini theorem applied to the set of perturbations $(C_h, d_h, C_g, d_g) \times (A,b)$.

Based on this result and with the help of partitions of the unity, the density of I_k now follows as usual, see [21]. As can easily be seen, the relation $I = \cap_{k=3}^{\infty} I_k$ holds. So the set I is generic. □

Remark 3.4.1 *If instead of* $(t\Phi(y) + (1-t)(y-y_0))$ *we choose the simpler parametric function* $(t\Phi(y) + (1-t)c)$ *with some* $c \in \mathbb{R}^m,\ c \neq 0$*, then at* $t = 0$

no initial solution exists for $Y_S(0) = \mathbb{R}^m$. Indeed the initial problem is to find a point $y \in \mathbb{R}^m$ such that $c^T(z - y) \geq 0$, $\forall z \in \mathbb{R}^m$, which is impossible.

Remark 3.4.2 *Assume that $VI(\Phi, h, g)$ is a regular problem, see Definition 3.2.2, and that for all critical points with associated multipliers (λ, μ) it holds that $\lambda_i \neq \lambda_j$, $\forall i \neq j$. Let us consider a sequence $(y_k, t_k) \in \Sigma_{gc}$, $(y_k, t_k) \to (\overline{y}, 1)$, such that $\overline{y}$ is a feasible point of $VI(\Phi, h, g)$. By the regularity assumptions $\overline{y}$ is a non-degenerate g.c. point of the original problem $VI(\Phi, h, g)$. Using this fact under our assumptions, we can prove that $(y_k, t_k) \in \Sigma^1_{gc}$ for $k \gg 1$.*

We shortly discuss the particularities of the standard embedding when applied to some special instances of VI problems. The first case is

$$VI(\Phi, \mathbb{R}^m_+): \quad \begin{array}{rc} & \text{find } y \in \mathbb{R}^m_+ \\ \text{such that} & \Phi(y)^T(z - y) \geq 0, \forall z \in \mathbb{R}^m_+. \end{array} \tag{3.4.3}$$

It is not difficult to see that it is equivalent with the *Non-Linear Complementarity Problem (NLCP)*,

$$NLCP: \quad \begin{array}{rrcl} & \multicolumn{3}{l}{\text{find } y \in \mathbb{R}^m} \\ \text{such that} & y & \geq & 0, \\ & \Phi(y) & \geq & 0, \\ & \Phi(y)^T y & = & 0. \end{array} \tag{3.4.4}$$

The second case is the box constrained problem $VI(\Phi, [0,1]^m)$. It can also be written as follows:

$$\begin{array}{rlllllll} & \multicolumn{7}{l}{\text{find } y \in \mathbb{R}^m} \\ \text{such that} & y_i & = & 0 & \Rightarrow & \Phi_i(y) & \geq & 0, \\ & y_i & = & 1 & \Rightarrow & \Phi_i(y) & \leq & 0, \\ & y_i & \in & (0,1) & \Rightarrow & \Phi_i(y) & = & 0, \\ & \multicolumn{7}{l}{\forall i = 1, 2, \ldots, m.} \end{array} \tag{3.4.5}$$

For numerical reasons we prefer to deal with compact feasible sets. So in the case of the VI problem defined in (3.4.3), the feasible set $\mathbb{R}^m_+$ is replaced by $\mathbb{R}^m_+ \cap \overline{B}^m_{\sqrt{p}}(0)$ for some p, $p > 0$ large enough, and we consider the problems,

$$\begin{array}{rc} & \text{find } y \in Y \\ \text{such that} & \Phi(y)^T(z - y) \geq 0, \ \forall z \in Y, \end{array}$$

$$\begin{array}{rrcll} \text{where} & Y & = & \left\{ y \in \mathbb{R}^m_+ \mid \|y\|^2 \leq p \right\} & \text{in case of NLCP} \\ & Y & = & [0,1]^m & \text{for box-constrained } VI. \end{array}$$

In both cases LICQ holds for Y and it is easy to find a point $y_0 \in Y$. So, in the embedding approach we can leave the set Y unchanged. This leads to the embedding

$$VI(t): \quad \begin{array}{rc} & \text{for } t \in [0,1], \text{ find } y \in Y \\ \text{such that} & \left(t\Phi(y) + (1-t)(y - y_0)\right)^T (z - y) \geq 0, \ \forall z \in Y. \end{array}$$

As LICQ holds at all feasible points, by using a proof similar to the proof of Theorem 3, pp. 21 in [19], we can show that generically with respect to the function $\Phi \in [C^3_S]^m_m$, the problems $VI(\Phi, \mathbb{R}^m_+ \cap \overline{B}^m_{\sqrt{p}}(0))$ and $VI(\Phi, [0,1]^m)$ are regular for $t \in (0,1)$. Moreover the only possible types of g.c. points, in the generic case, are g.c. points of type 1, 2 and 3.

Here in both cases we can choose the embedding $t\Phi(y) + (1-t)c$. A starting solution $y(0) = y_0$ at $t = 0$ must necessarily be a boundary point of Y. So we will start with an active index set $J_0(y_0) \neq \emptyset$. In particular if c is a strictly positive vector, $y_0 = 0$ will be an starting solution with $J_0(y_0) = \{1, \ldots, m\}$ and a non-degenerate point of $VI_S(0, \Phi, h, g)$.

3.4.2 Penalty embedding

Penalty embeddings for common optimization problems have been already developed in Dentcheva, Gollmer, Guddat and Rückmann [12] and in Gómez [20]. The main advantage of these embeddings is that under regularity assumptions, no points of type 5 can occur. We extend this approach to variational inequalities, and define, for the non-parametric problem $VI(\Phi, h, g)$ the parametric VI

$$VI_P(t): \qquad \text{for } t \in [0,1], \text{ find } (y, v, w)$$

$$\text{such that } \Phi_P(y, v, w, t)^T \begin{pmatrix} z_y - y \\ z_v - v \\ z_w - w \end{pmatrix} \geq 0, \ \forall z = (z_y, z_v, z_w) \in Y_P(t),$$

$$\text{where } \Phi_P(y, v, w, t) = \begin{pmatrix} t\Phi(y) + (1-t)y \\ v \\ w - e_q \end{pmatrix} \text{ and}$$

$$Y_P(t) = \left\{ (y, v, w) \in \mathbb{R}^{m+q_0+q} \;\middle|\; \begin{array}{rcll} th_i(y) + (1-t)v_i & = & 0, & i = 1, \ldots, q_0, \\ tg_j(y) + (1-t)w_j & \geq & 0, & j = 1, \ldots, q, \\ \|y, v, w\|^2 & \leq & p. & \end{array} \right\}$$

Here $(v, w) \in \mathbb{R}^{q_0} \times \mathbb{R}^q$ are additional variables, e_q is the vector in $\mathbb{R}^q$ with all components equal to 1, and p, $p \gg 1$, is a fixed parameter. A starting solution for $t = 0$ is given by $(y, v, w) = (0, 0, e_q)$. This is a point of type 1 for $VI_P(0)$. More precisely, it is the unique generalized critical point for $VI_P(0)$, see Definition 3.2.1, with multipliers μ, $\mu \geq 0$.

As in the standard case, we will refer to the functions that define the parametric embedding $VI_P(t)$ as

$$\Psi_P(t; \Phi, h, g) = \begin{pmatrix} \Phi_P(y, v, w, t) \\ th_i(y) + (1-t)v_i, \ i = 1, \ldots, q_0 \\ tg_j(y) + (1-t)w_j, \ j = 1, \ldots, q \\ p - \|y, v, w\|^2 \end{pmatrix}$$

Proposition 3.4.2 *The set* $\left\{(\Phi,h,g)\in[C_S^3]_m^{m+q_0+q}\,\middle|\,\Psi_P(t;\Phi,h,g)\in\mathcal{F}_{VI(t)}|_{t\in(0,1]}\right\}$ *is generic in* $[C_S^3]_m^{m+q_0+q}$.

Proof. We follow the lines of the proof of Proposition 3.4.1. We consider the sets $I_k=\left\{(\Phi,h,g)\mid\Psi_P(t;\Phi,h,g)\in\mathcal{F}_{VI(t)}|_{t\in[\frac{1}{k},1]}\right\},\ k=2,3,\ldots,$ and prove that they are open and dense in $[C_S^3]_m^{m+q_0+q}$. To show that the sets I_k are dense, we again have to apply a perturbation argument and a partition of unity. For the open part, we use a local stability theorem and a locally finite cover, see the proof of Theorem 6.25 in [21].
The result follows then by noticing that $\left\{(\Phi,h,g)\mid\ \Psi_P(t;\Phi,h,g)\in\mathcal{F}_{VI(t)}|_{t\in(0,1]}\right\}$ is equal to $\cap_{k=2}^{\infty}I_k$, and $I_k,\ k\geq 2$, are open and dense sets. □

Remark 3.4.3 *In contrast to the standard embedding, under regularity, there will not appear points of type 5 for the penalty embedding, because the number of variables is always greater than or equal to the number of possible active constraints. However the penalty embedding has the disadvantage of a larger number of variables.*

Consider again the special case of VI given by the NLCP, see (3.4.4), with $Y=\mathbb{R}_+^m$. The penalty embedding leads to the problem,

$$\begin{array}{c}\text{for } t\in[0,1],\ \text{find } y\\ \text{such that} \begin{pmatrix}t\Phi(y)+(1-t)y\\ w-e_n\end{pmatrix}^T\begin{pmatrix}z_y-y\\ z_w-w\end{pmatrix}\geq 0,\ \forall(z_y,z_w)\in Y_P(t)\end{array}\tag{3.4.6}$$

with sets

$$Y_P(t)=\left\{(y,w)\in\mathbb{R}^{m+m}\ \middle|\ \begin{array}{rcl}ty+(1-t)w&\geq&0,\\ \|y,w\|^2&\leq&p\end{array}\right\}$$

which also satisfy the LICQ condition at all their feasible points.

Proposition 3.4.3 *Let* $p>0$ *be fixed. Then for all* $t\in[0,1]$ *the condition LICQ is satisfied for problem (3.4.6).*

Proof. As the first n constraints are linearly independent, the LICQ condition could only fail if the compactification restriction is active and its gradient is a linear combination of the gradients of the other $|J_0(y,w,t)|-1$ active constraints. This linear combination reads:

$$\begin{pmatrix}tI_{J_0(y,w,t)\setminus\{m+1\}}\\ (1-t)I_{J_0(y,w,t)\setminus\{m+1\}}\end{pmatrix}\lambda=2\begin{pmatrix}y\\ w\end{pmatrix}$$

for some $\lambda\in\mathbb{R}^m$. This means that

$$\begin{array}{rcl}t\lambda_{J_0(y,w,t)\setminus\{m+1\}}&=&2y_{J_0(y,w,t)\setminus\{m+1\}}\\ (1-t)\lambda_{J_0(y,w,t)\setminus\{m+1\}}&=&2w_{J_0(y,w,t)\setminus\{m+1\}}\\ 0&=&y_{J_0^c(y,w,t)}\\ 0&=&w_{J_0^c(y,w,t)}.\end{array}$$

So, $J_0^c(y,w,t)=\emptyset$ because, for $i \in J_0^c(y,w,t)$, the inequality $ty_i+(1-t)w_i>0$ should hold, in contradiction to $w_i=y_i=0$. Consequently $ty_i+(1-t)w_i=0$, $i=1,\dots,m$. Using the first two equations yields $\frac{t^2\lambda_i+(1-t)^2\lambda_i}{2}=0,\ i=1,\dots,m$. So, $\lambda=0=y=w$. But as the compactification constraint is active it would follow $p=\|y\|^2+\|w\|^2=0$, contradicting $p>0$. □

By Proposition 3.4.3, g.c. points of type 4 and 5 are excluded for problem (3.4.6) and it can be proven that generically w.r.t. Φ, all generalized critical points are of type 1, 2 or 3.

The box constrained VI, *i.e.*, $Y=[0,1]^m$, can be analogously embedded, leading to the parametric problem

$$\begin{aligned}&\text{for } t\in[0,1] \text{ find } y\\ \text{such that}\quad &\begin{pmatrix} t\Phi(y)+(1-t)y\\ w_1-e_n\\ w_2+e_n\end{pmatrix}^T\begin{pmatrix} z_y-y\\ z_{w_1}-w_1\\ z_{w_2}-w_2\end{pmatrix}\geq 0,\ \forall(z_y,z_{w_1},z_{w_2})\in Y_P(t),\end{aligned} \tag{3.4.7}$$

where

$$Y_P(t)=\left\{(y,w_1,w_2)\in\mathbb{R}^{m+m+m}\ \middle|\ \begin{array}{rcl} ty+(1-t)w_1 & \geq & 0,\\ ty+(1-t)w_2 & \leq & t,\\ \|y,w_1,w_2\|^2 & \leq & p.\end{array}\right\}$$

In contrast to the standard embedding the first $2m$ constraints do not guarantee that $Y_P(t)$ is compact. Moreover, LICQ may be violated as is shown by the following result.

Proposition 3.4.4 *Let $m=1$. For any fixed $p,\ p\gg 0$, there is some $t,\ t\in[0,1]$ and $(y,w_1,w_2)\in Y_P(t)$, described in (3.4.7), such that LICQ fails.*

Proof. We will construct a point $(y,w_1,w_2,t)\in\mathbb{R}\times\mathbb{R}\times\mathbb{R}\times(0,1)$ such that all constraints are active and the gradient of the compactification constraint is a linear combination of the others. This means that there exists (λ_1,λ_2) such that the following system has a solution.

$$\begin{array}{rcl} t\lambda_1+t\lambda_2 & = & 2y,\\ (1-t)\lambda_1 & = & 2w_1\\ (1-t)\lambda_2 & = & 2w_2\\ ty+(1-t)w_1 & = & 0\\ ty+(1-t)w_2 & = & t,\\ y^2+w_1^2+w_2^2 & = & p.\end{array}$$

Here the first three equations represent the linear combination of the gradients and the last three, the activity condition $J_0(y, w_1, w_2, t) = \{1, 2, 3\}$.

Substituting $2(\frac{w_1}{1-t}, \frac{w_2}{1-t})$ for (λ_1, λ_2) in the first equation, we obtain:

$$y = t\frac{w_1 + w_2}{1 - t}.$$

Using this expression in the equations $ty + (1-t)w_1 = 0, \quad ty + (1-t)w_2 = t$, it follows

$$\begin{aligned} \tau w_1 + t^2 w_2 &= 0 \\ t^2 w_1 + \tau w_2 &= (1-t)t \end{aligned}$$

where $\tau = t^2 + (1-t)^2$. Therefore, we find $(w_1, w_2) = (\frac{-t^3(1-t)}{(\tau - t^2)(\tau + t^2)}, \frac{t\tau(1-t)}{(\tau - t^2)(\tau + t^2)})$ and $y = \frac{t^2}{\tau + t^2}$.

Finally, as $y^2 + w_1^2 + w_2^2 = p$, we need to prove that there is some $\bar{t},\ \bar{t} \in [0, 1]$ such that the function $p(t) := \frac{t^4}{(\tau+t^2)^2} + t^2 \frac{(t^4+\tau^2)(1-t)^2}{(\tau-t^2)^2(\tau+t^2)^2}$ satisfies $p(\bar{t}) = p$.

As $\tau - t^2 = (1-t)^2$ the previous expression simplifies to:

$$p(t) = \frac{t^4}{(\tau + t^2)^2} + t^2 \frac{(t^4 + \tau^2)}{(1-t)^2(\tau + t^2)^2}$$

When $t \to 1$, it follows that $p(t) \to +\infty$ while $p(0) = 0 < p$. So there is a point $\bar{t},\ \bar{t} \in (0, 1)$, such that $p(\bar{t}) = p$ and at the feasible point

$$(y, w_1, w_2, t) = \left(\frac{\bar{t}^2}{1 - 2\bar{t} + 3\bar{t}^2}, \frac{-\bar{t}^3}{(1-t)(1 - 2\bar{t} + 3\bar{t}^2)}, \frac{\bar{t}(1 - 2\bar{t} + 2\bar{t}^2)}{(1-\bar{t})(1 - 2\bar{t} + 3\bar{t}^2)}, \bar{t} \right)$$

the LICQ condition fails. $\square$

With regard to the preceding proposition, for the box constrained VI, the standard embedding behaves better than the penalty embedding.

Chapter 4

Problems with complementarity constraints

4.1 Introduction

In this chapter, we will study optimization problems of the form:

$$
\begin{array}{rl}
\min\limits_{x} & f(x) \\
\text{s.t.} & x \in \mathcal{M}_{CC}
\end{array}
$$
$$
\mathcal{M}_{CC} = \left\{ x \in \mathbb{R}^n \;\middle|\; \begin{array}{rcll} h_k(x) & = & 0, & k = 1, \dots, q_0, \\ g_j(x) & \geq & 0, & j = 1, \dots, q, \\ r_i(x)s_i(x) & = & 0, & i = 1, \dots, l, \\ r_i(x), s_i(x) & \geq & 0, & i = 1, \dots, l \end{array} \right\} \tag{4.1.1}
$$

where $f, h_k, g_j, r_i, s_i : \mathbb{R}^n \to \mathbb{R}$, $k = 1, \dots, q_0$, $j = 1, \dots, q$, $i = 1, \dots, l$ are assumed to be C^2-functions.
The constraints $r_i(x)s_i(x) = 0$, $r_i(x), s_i(x) \geq 0$ are called *Complementarity Constraints* and such problems will be termed problems with complementarity constraints, denoted as P_{CC}. To keep the presentation as clear as possible, we omit the equality constraints $h_k(x) = 0$. So, in the Sections 4.2-4.6, the P_{CC} problem will be

$$
\begin{array}{rrl}
P_{CC}: & \min\limits_{x} & f(x) \\
& \text{s.t.} & x \in \mathcal{M}_{CC}
\end{array}
$$
$$
\mathcal{M}_{CC} = \left\{ x \in \mathbb{R}^n \;\middle|\; \begin{array}{rcll} g_j(x) & \geq & 0, & j = 1, \dots, q, \\ r_i(x)s_i(x) & = & 0, & i = 1, \dots, l, \\ r_i(x), s_i(x) & \geq & 0, & i = 1, \dots, l. \end{array} \right\} \tag{4.1.2}
$$

We, however, emphasize that all results of these sections remain true for problem (4.1.1) if additional LICQ conditions are assumed for the equality constraints $h_k(x) = 0$. We will often use the abbreviations $\hat{h}_i(x) := r_i(x)s_i(x)$ $s(x) = (s_1(x), \dots, s_l(x))$, $r(x) = (r_1(x), \dots, r_l(x))$ and $g(x) = (g_1(x), \dots, g_q(x))$.

Complementarity constraints arise in problems with equilibrium conditions, see [43], or as special cases in a so-called Kuhn Tucker approach for solving problems with a bilevel structure, see *e.g.* [60] and Section 1.2. Other practical examples appear when solving numerically problems from mathematical physics, as presented in Section 1.4, see also [43].

It is well-known that problems with complementarity constraints cannot be solved by using standard nonlinear programming approaches because the MFCQ will never hold. Indeed, if at a feasible point $\overline{x}$ the relation $r_{i_0}(\overline{x}) = s_{i_0}(\overline{x}) = 0$ holds for some $i_0 \in \{1, \ldots, l\}$, it follows

$$\nabla \hat{h}_{i_0}(\overline{x}) = s_i(\overline{x}) \nabla r_{i_0}(\overline{x}) + r_{i_0}(\overline{x}) \nabla s_{i_0}(\overline{x}) = 0.$$

If $r_i(\overline{x}) = 0$ and $s_i(\overline{x}) > 0$, then $\nabla \hat{h}_i(\overline{x}) = s_i(\overline{x}) \nabla r_i(\overline{x})$ and there is no vector ξ such that $\nabla \hat{h}_i(\overline{x})\xi = 0$, $\nabla r_i(\overline{x})\xi > 0$.

To circumvent this problem and to obtain necessary optimality conditions, generalized derivatives have been used in Ye [63], [64]. Pang and Fukushima [44] derived optimality conditions by assuming an Abadie-type constraint qualification. Fritz-John type necessary conditions were established by [14] based on the relation between P_{CC} and certain nonlinear optimization problems without complementarity constraints via a disjunctive analysis. In this way a KKT-type optimality condition can be obtained if a natural constraint qualification holds.

The sensitivity of P_{CC} with respect to parameters has also been studied. Sufficient conditions for stability of the value function and the stationary points were obtained in [24] and [49], respectively.
Some specific algorithms for solving complementarity constrained problems have been developed. In Schramm and Zowe [54], for example, a bundle method is constructed. The algorithm PIPA, presented in [42], applies a penalty interior approach and solves the associated optimization problems using Sequential Quadratic Programming (SQP) methods. The convergence of this algorithm is proven under a strong hypothesis. SQP-type methods have been adapted to particular cases such as $s(x) = (x_1, \ldots, x_l)$ and $r(x)$ satisfies a kind of convexity condition (see Jiang and Ralph [26]) or $s(x) = (x_1, \ldots, x_l)$, $r(x) = Qx + q$, see Fukushima, Luo and Pang [15], Liu, Perakis and Sun [40] and Zhang and Liu [65].

A promising algorithmic idea is to substitute the complementarity constraints $r_i(x)s_i(x) = 0$ by the (regularizing) inequality $r_i(x)s_i(x) \leq \tau$ or by the smoothing equality $r_i(x)s_i(x) = \tau$ and to let $\tau \to 0^+$. This kind of techniques have been discussed in [51] and Ralph and Wright [47]. In Chen and Fukushima [10], this approach is applied when $s(x) = x$, $r(x) = Qx + q$, and in Facchinei, Houyuan and Liqun [13], for mathematical problems with VI constraints (1.1.8) using NCP-functions.

Exact and L^1-penalty strategies have also been used, see for example Lin and Fukushima [38] and Scholtes and Stöhr [52], respectively. Besides, in Hu and Ralph [25], the complementarity constraints are smoothed and then solved via a penalty function. They also presented a convergence analysis. Interior point methods have also been implemented, see Benson, Shanno and Vanderbei [5], Benson, Sen, Shanno and Vanderbei [6] and Raghunathan and Biegler [46]. Other algorithmic ideas are a branch and bound method in Liu and Zhang [39], relaxation algorithms via a simplex representation in [37], an active-set algorithm in Fukushima and Tseng [17], and an heuristic in Braun and Mitchell [8].

In this thesis, we consider the parametric smoothing approach P_τ, based on the perturbation $r_i(x)s_i(x) = \tau,\ i = 1, \ldots, l$. We study the convergence of the solutions of P_τ when $\tau \to 0^+$ and discuss the regularity of P_τ as a parametric problem. We also present the generic singularities appearing in one-parametric mathematical programs with complementarity constraints.
The chapter is organized as follows. In the next section the structure of the feasible set, the active index set, the cone of feasible directions, etc., are presented. The third section includes necessary optimality conditions, constraint qualifications and different types of stationarity concepts for P_{CC}. Section 4.4 is divided into two parts. In the first, we review the generic properties of P_{CC} and, in the second, necessary and sufficient primal-dual optimality conditions for local minimizers of order 1 and 2 are presented. In Section 4.5, for the parametric approach P_τ, we prove the existence of a sequence of local minimizers (stationary points) of P_τ converging to a local minimizer (stationary point) of P_{CC} with rate $\mathcal{O}(\sqrt{\tau})$ when $\tau \to 0$. In Section 4.6, we show that, generically the problem P_τ is in $\mathcal{F}_{(0,1]}$ (*cf.* Definition 2.4.7). The chapter ends with a genericity analysis of one-parametric mathematical programs with complementarity constraints. We study the types of singularities that may appear at a generalized critical point in the generic case and present the local behavior of the set of generalized critical points around such a point.

4.2 Structure of the feasible set

We now analyze the structure of the feasible set $\mathcal{M}_{\mathcal{CC}}$ of P_{CC} in (4.1.2).

To do so we make use of the disjunctive or *piecewise* structure of the problem. Note that in each feasible point x either $r_i(x)$ or $s_i(x)$ should be zero. So all feasible points of P_{CC} are given as feasible points of a problem

$$
\begin{array}{llcccccl}
P_{\mathcal{I}}: & & & & \min f(x) & & & \\
 & \text{s.t.} & & & g_j(x) \geq 0, \ \ j = 1, \ldots, q, & & & \\
 & & r_i(x) & = & 0, \ \ s_i(x) & \geq & 0, & i \in \mathcal{I}, \\
 & & s_i(x) & = & 0, \ \ r_i(x) & \geq & 0, & i \in \{1, \ldots, l\} \setminus \mathcal{I}.
\end{array}
\tag{4.2.1}
$$

for some $\mathcal{I} \subset \{1, \ \ldots, l\}$.

To be more precise, for a feasible point $\overline{x} \in \mathcal{M}_{CC}$ we introduce the active index sets:

$$\begin{array}{rcl} J_0(\overline{x}) & = & \{j \mid g_j(\overline{x}) = 0\,\}\,, \\ I_r(\overline{x}) & = & \{i \mid r_i(\overline{x}) = 0,\ s_i(\overline{x}) > 0\}\,, \\ I_s(\overline{x}) & = & \{i \mid s_i(\overline{x}) = 0,\ r_i(\overline{x}) > 0\}\,, \\ I_{rs}(\overline{x}) & = & \{i \mid r_i(\overline{x}) = s_i(\overline{x}) = 0\}\,. \end{array}$$

For any $I \subset I_{rs}(\overline{x})$ the problem $P_{\mathcal{I}}$ with $\mathcal{I} = I_r(\overline{x}) \cup I$ will be denoted by $P_I(\overline{x})$, *i.e.*,

$$P_I(\overline{x}): \quad \begin{array}{ll} & \min f(x) \\ \text{s.t.} & g_j(x) \geq 0,\ j = 1, \dots, q, \\ & \begin{array}{rclrcll} r_i(x) & = & 0, & s_i(x) & \geq & 0, & i \in I_r(\overline{x}), \\ s_i(x) & = & 0, & r_i(x) & \geq & 0, & i \in I_s(\overline{x}), \\ r_i(x) & = & 0, & s_i(x) & \geq & 0, & i \in I, \\ s_i(x) & = & 0, & r_i(x) & \geq & 0, & i \in I_{rs}(\overline{x}) \setminus I. \end{array} \end{array} \tag{4.2.2}$$

Let $\mathcal{M}_I(\overline{x})$ denote the feasible set of this program. Then obviously, the following *disjunctive or piecewise* description holds.

Lemma 4.2.1 (*cf.* **[42])**
(a) There exists a neighborhood $B_\varepsilon(\overline{x})$ ($\varepsilon > 0$) of $\overline{x}$ such that

$$\mathcal{M}_{CC} \cap B_\varepsilon(\overline{x}) = \bigcup_{I \subset I_{rs}(\overline{x})} \left(\mathcal{M}_I(\overline{x}) \cap B_\varepsilon(\overline{x})\right).$$

(b) The point $\overline{x} \in \mathcal{M}_{CC}$ is a local minimizer of order ω of P_{CC} if and only if $\overline{x}$ is a local minimizer of order ω of $P_I(\overline{x})$ for all $I \subset I_{rs}(\overline{x})$.

Based on this lemma, optimality conditions and genericity results for $P_I(x)$ lead directly to corresponding results for the complementarity constrained problem P_{CC}. This will be done in the next sections.

4.3 Necessary optimality conditions

This section introduces some notations and deals with different types of necessary optimality conditions for P_{CC}. We begin with some definitions. As usual

$$L(x, \lambda_0, \mu, \rho, \sigma) = \lambda_0 f(x) - \sum_{j \in J_0(x)} g_j(x)\mu_j - \sum_{i \in I_r(x) \cup I_{rs}(x)} r_i(x)\rho_i - \sum_{i \in I_s(x) \cup I_{rs}(x)} s_i(x)\sigma_i$$

denotes the Lagrangean function. Given $\overline{x} \in \mathcal{M}_{CC}$ and $I \subset I_{rs}(\overline{x})$ we introduce the set of *critical directions* for $P_I(\overline{x})$:

$$C_I(\overline{x}) = \left\{ d \in \mathbb{R}^n \;\middle|\; \begin{array}{rclrcll} \nabla f(\overline{x})d & \leq & 0, & \nabla g_j(\overline{x})d & \geq & 0, & j \in J_0(\overline{x}), \\ \nabla r_i(\overline{x})d & = & 0, & \nabla s_i(\overline{x})d & \geq & 0, & i \in I, \\ \nabla r_i(\overline{x})d & \geq & 0, & \nabla s_i(\overline{x})d & = & 0, & i \in I_{rs}(\overline{x}) \setminus I, \\ \nabla r_i(\overline{x})d & = & 0, & i \in I_r(\overline{x}), & & & \\ \nabla s_i(\overline{x})d & = & 0, & i \in I_s(\overline{x}). & & & \end{array} \right\} \tag{4.3.1}$$

We need some constraint qualifications.

Definition 4.3.1 Let $\overline{x} \in \mathcal{M}_{CC}$:

- *MPCC-LICQ holds at* $\overline{x}$ if $\nabla r_i(\overline{x}), \nabla s_k(\overline{x}), \nabla g_j(\overline{x}),\ i \in I_r(\overline{x}) \cup I_{rs}(\overline{x})$, $k \in I_s(\overline{x}) \cup I_{rs}(\overline{x})\ j \in J_0(\overline{x})$, are linearly independent.

- *MPCC-MFCQ holds at* $\overline{x}$ if the gradients $\nabla r_i(\overline{x}), \nabla s_k(\overline{x}),\ \ i \in I_r(\overline{x}) \cup I_{rs}(\overline{x}),\ k \in I_s(\overline{x}) \cup I_{rs}(\overline{x})$ are linearly independent and there is a vector ξ such that $\nabla g_j(\overline{x})^T\xi > 0,\ j \in J_0(\overline{x})$ and $\nabla r_i(\overline{x})^T\xi = 0, \nabla s_k(\overline{x})^T\xi = 0$, $i \in I_r(\overline{x}) \cup I_{rs}(\overline{x}),\ k \in I_s(\overline{x}) \cup I_{rs}(\overline{x})$.

We now introduce different types of stationarity.

Definition 4.3.2 Let $\overline{x} \in \mathcal{M}_{CC}$

- $\overline{x}$ is *a Fritz John point of* P_{CC} if it is a Fritz John point for all problems $P_I(\overline{x}),\ I \subset I_{rs}(\overline{x})$ (see (4.2.2)).

- $\overline{x}$ will be called *weakly stationary* if there are multipliers $(\lambda_0, \mu, \rho, \sigma)$ not all zero satisfying $\nabla L(\overline{x}, \lambda_0, \mu, \rho, \sigma) = 0$ and $\lambda_0,\ \mu \geq 0$.

- If $\overline{x}$ is a weakly stationary point with $\sigma_i \cdot \rho_i \geq 0, \forall i \in I_{rs}(\overline{x})$ for some associated multiplier vector $(\lambda_0, \mu, \rho, \sigma)$, then $\overline{x}$ is called *a C-stationary point.*

- $\overline{x}$ is *M-stationary* if $\overline{x}$ is a weakly stationary point with an associated multiplier $(\lambda_0, \mu, \rho, \sigma)$ such that $\lambda_0 = 1$ and for each $i \in I_{rs}(\overline{x})$ either $\sigma_i > 0,\ \rho_i > 0$ or $\sigma_i \cdot \rho_i = 0$ holds.

- If for all $I \subset I_{rs}(\overline{x})$, $\overline{x}$ is a stationary point of the nonlinear program $P_I(\overline{x})$ (see Definition 2.2.3) then $\overline{x}$ is called a *B-stationary point.*

- *A-stationary points* are weakly stationary points with associated multiplier $(\lambda_0, \mu, \rho, \sigma)$ satisfying $\lambda_0 = 1$ and for all $i \in I_{rs}(\overline{x})$, σ_i or ρ_i is non-negative, *i.e,* there is some $I,\ I \subset I_{rs}(\overline{x})$ such that $\overline{x}$ is a stationary point of $P_I(\overline{x})$.

- $\overline{x}$ is called *a strongly stationary point* if the weak stationarity conditions are satisfied with multiplier $(\lambda_0, \mu, \rho, \sigma)$, fulfilling $\lambda_0 = 1$ and $\sigma_i, \rho_i \geq 0$, $i \in I_{rs}(\overline{x})$.

Let $\overline{x}$ be feasible for the complementarity constrained problem P_{CC} in (4.1.2). Consider two related problems, the *tightened problem*

$$P_T(\overline{x}) : \qquad \begin{array}{rl} & \min f(x) \\ \text{s.t.} & x \in \mathcal{M}_T \end{array} \tag{4.3.2}$$

$$\mathcal{M}_T = \left\{ x \in \mathbb{R}^n \;\middle|\; \begin{array}{ccccccc} & & g_j(x) \geq 0, & j = 1, \ldots, q, & & & \\ r_i(x) & = & 0, & s_i(x) & \geq & 0, & i \in I_r(\overline{x}), \\ s_i(x) & = & 0, & r_i(x) & \geq & 0, & i \in I_s(\overline{x}), \\ r_i(x) & = & 0, & s_i(x) & = & 0, & i \in I_{rs}(\overline{x}) \end{array} \right\}$$

and the *relaxed problem*

$$P_R(\overline{x}): \qquad \begin{array}{rl} & \min f(x) \\ \text{s.t.} & x \in \mathcal{M}_R \end{array} \tag{4.3.3}$$

$$\mathcal{M}_R = \left\{ x \in \mathbb{R}^n \;\middle|\; \begin{array}{ccccccc} & & g_j(x) \geq 0, & j = 1, \ldots, q, & & & \\ r_i(x) & = & 0, & s_i(x) & \geq & 0, & i \in I_r(\overline{x}), \\ s_i(x) & = & 0, & r_i(x) & \geq & 0, & i \in I_s(\overline{x}), \\ r_i(x) & \geq & 0, & s_i(x) & \geq & 0, & i \in I_{rs}(\overline{x}). \end{array} \right\}$$

Note that, if $\overline{x}$ is a strongly stationary point of P_{CC}, then $\overline{x}$ is also a stationary point (in the classical sense) of $P_R(\overline{x})$. Note also that MPCC-LICQ holds at $\overline{x} \in \mathcal{M}_{CC}$ if and only if *LICQ* is satisfied at $\overline{x}$ in $\mathcal{M}_R$. The condition MPCC-MFCQ is stronger than the fulfillment of MFCQ for P_R. From the definitions, the following proposition is immediate.

Proposition 4.3.1 *For a point $\overline{x} \in \mathcal{M}_{CC}$, the following holds:*
MPCC-LICQ for P_{CC} $\Leftrightarrow$ LICQ for $P_T(\overline{x})$ $\Leftrightarrow$ LICQ for $P_R(\overline{x})$.
MPCC-MFCQ for P_{CC} $\Leftrightarrow$ MFCQ for $P_T(\overline{x})$ $\Rightarrow$ MFCQ for $P_R(\overline{x})$.

We mention the following necessary optimality conditions without proof.

Proposition 4.3.2 (*cf.* **[14])** *For P_{CC}, the following relations hold:*

(a) Strong $\Rightarrow$ M- $\Rightarrow$ C- $\Rightarrow$ weak stationarity.
Strong $\Rightarrow$ B- $\Rightarrow$ A- $\Rightarrow$ weak stationarity.

(b) If $\overline{x}$ is a local minimizer then it is a C-stationary point.

(c) If MPCC-MFCQ holds at a local minimizer $\overline{x}$ of P_{CC} then $\overline{x}$ is A-stationary.

(d) If $\overline{x}$ is a local minimizer and MPCC-LICQ is satisfied, then $\overline{x}$ is a strongly stationary point. In this case B- and strong stationarity are equivalent.

We will say that the strict complementarity property holds for P_{CC} at a feasible point $\overline{x}$ if

$$r_i(\overline{x}) + s_i(\overline{x}) > 0, \qquad \forall i = 1, \ldots, l. \qquad (SC)$$

In this case all the problems $P_T(\overline{x})$, $P_I(\overline{x})$, $P_R(\overline{x})$ coincide and the stationarity definitions are equivalent. Furthermore, the problem P_{CC} can then be treated as a standard finite program. In [42] the convergence of an algorithm for MPCC has been obtained under the condition *SC*. However, for complementarity constrained problems this condition is not usually fulfilled, as can be seen in the following standard example.

Example 4.3.1

$$\begin{array}{rcrl} & \min x_1 + x_2 & & \\ s.t. & x_1 x_2 & = & 0, \\ & x_1,\ x_2 & \geq & 0. \end{array}$$

Here SC fails at the minimizer $\overline{x} = (0,0)$ even if small perturbations of the involved functions are considered.

This example shows that, even in the generic case, around local minimizers the SC condition may fail. We therefore analyze the P_{CC} problem under weaker assumptions. To do so we make use of the disjunctive structure of $\mathcal{M}_{CC}$ (see Section 4.2) and define the cone of critical directions at $\overline{x} \in \mathcal{M}_{CC}$ by

$$C_{\overline{x}} = \bigcup_{I \subset I_{rs}(\overline{x})} C_I(\overline{x}), \tag{4.3.4}$$

where $C_I(\overline{x})$ is the cone of critical directions of the problem $P_I(\overline{x})$.

Definition 4.3.3 We say that $\overline{x} \in \mathcal{M}_{CC}$ is *a g.c. point of* P_{CC} if there is a multiplier vector $(\lambda_0, \mu, \rho, \sigma) \neq 0$ such that $\nabla_x L(x, \lambda_0, \mu, \rho, \sigma) = 0$.
Let $\overline{x}$ be a g.c. point with associated multiplier vector $(1, \mu, \rho, \sigma)$. We say that *the MPCC-strict complementarity condition (MPCC-SC) holds* if

$$\mu_j \neq 0,\ \forall j \in J_0(\overline{x}), \quad \rho_i \neq 0, \sigma_i \neq 0,\ \forall i \in I_{rs}(\overline{x}). \tag{4.3.5}$$

The MPCC second order condition (MPCC-SOC) is satisfied if

$$d^T \nabla_x^2 L(\overline{x}, 1, \mu, \rho, \sigma) d \neq 0, \qquad \forall d \in T_{\overline{x}} \mathcal{M}_R \backslash \{0\}. \tag{4.3.6}$$

A point $\overline{x} \in \mathcal{M}_{CC}$ such that MPCC-LICQ, MPCC-SC, and MPCC-SOC holds is called *a non-degenerate critical point in the MPCC-sense.*

If at a g.c. point $\overline{x}$ the MPCC-LICQ condition holds, then there is a unique multiplier vector $(1, \mu, \rho, \sigma)$ such that $\nabla_x L(\overline{x}, 1, \mu, \rho, \sigma) = 0$. If $\overline{x}$ is also a B-stationary point, then the same unique multiplier vector $(1, \mu, \rho, \sigma)$ solves the KKT system corresponding to $P_I(\overline{x})$ for all $I \subset I_{rs}(\overline{x})$. Moreover it is not difficult to see that in this case the set $C_{\overline{x}}$ simplifies:

$$C_{\overline{x}} = \left\{ d \in \mathbb{R}^n \;\middle|\; \begin{array}{lcllccl} \nabla g_j(\overline{x}) d & \begin{array}{c} = \\ \geq \end{array} & 0 & \text{if} & \mu_j & \begin{array}{c} > \\ = \end{array} & 0, \quad j \in J_0(\overline{x}), \\ \nabla r_i(\overline{x}) d & \begin{array}{c} = \\ \geq \end{array} & 0 & \text{if} & \rho_i & \begin{array}{c} > \\ = \end{array} & 0, \quad i \in I_{rs}(\overline{x}), \\ \nabla s_i(\overline{x}) d & \begin{array}{c} = \\ \geq \end{array} & 0 & \text{if} & \sigma_i & \begin{array}{c} > \\ = \end{array} & 0, \quad i \in I_{rs}(\overline{x}), \\ \nabla r_i(\overline{x}) d & = & 0, & i \in I_r(\overline{x}), & & & \\ \nabla s_i(\overline{x}) d & = & 0, & i \in I_s(\overline{x}). & & & \end{array} \right\}$$

In the case of B-stationary points where MPCC-SC holds, the cone of critical directions becomes:

$$C_{\overline{x}} = \left\{ d \in \mathbb{R}^n \;\middle|\; \begin{array}{rcl} \nabla g_j(\overline{x})d &=& 0, \quad j \in J_0(\overline{x}), \\ \nabla r_i(\overline{x})d &=& 0, \quad i \in I_{rs}(\overline{x}), \\ \nabla s_i(\overline{x})d &=& 0, \quad i \in I_{rs}(\overline{x}), \\ \nabla r_i(\overline{x})d &=& 0, \quad i \in I_r(\overline{x}), \\ \nabla s_i(\overline{x})d &=& 0, \quad i \in I_s(\overline{x}) \end{array} \right\} \tag{4.3.7}$$

i.e., $C_{\overline{x}}$ coincides with $T_{\overline{x}\mathcal{M}_R}$, the tangent subspace of $\mathcal{M}_R$ in $\overline{x}$.

Note that if $\overline{x}$ is a non-degenerate critical point in the MPCC-sense, then $\mu_j \neq 0,\ j \in J_0(\overline{x})\ \rho_i,\ \sigma_i \neq 0,\ i \in I_{rs}(\overline{x})$ and $\nabla_x^2 L|_{T_{\overline{x}}\mathcal{M}_R}(\overline{x}, 1, \mu, \rho, \sigma)$ is nonsingular. This means $\overline{x}$ is a non-degenerate critical point of the relaxed problem $P_R(\overline{x})$ and thus also an (isolated) non-degenerate g.c. point of P_{CC}.

If $\overline{x}$ is a non-degenerate critical point in the MPCC-sense, such that $\mu_j > 0,\ j \in J_0(\overline{x})\ \rho_i,\ \sigma_i > 0,\ i \in I_{rs}(\overline{x})$ and $\nabla_x^2 L|_{T_{\overline{x}}\mathcal{M}_R}(\overline{x}, 1, \mu, \rho, \sigma) \succ 0$ are fulfilled, then $\overline{x}$ is a local minimizer of $P_R(\overline{x})$. As locally $\mathcal{M}_{CC} \subset \mathcal{M}_R$ holds, $\overline{x}$ will also be a local minimizer of P_{CC} in this case.

4.4 Optimality conditions based on the disjunctive structure

In this section we show how optimality conditions can simply be derived from the piecewise structure of the problems P_{CC}, see Section 4.2. These conditions are obtained under assumptions which are shown to be generic.

4.4.1 Genericity results for non-parametric P_{CC}

In this part we will review the genericity results of [53] needed in the next sections. By using the *piecewise description* (see Lemma 4.2.1) in principle, optimality conditions and genericity results for $P_I(\overline{x})$ lead directly to corresponding results for P_{CC}. Let, in the sequel, all functions f, g_j, s_i, r_i be in the space $[C^2]^1_n$ endowed with the C^2_S-topology, see Section 2.3. So, the set of problems P_{CC} can be identified with the set $\mathcal{P} := \{(f, g, s, r)\} = [C^2]^{1+q+2l}_n$. We say that a property for P_{CC} is generic if it holds for a subset of $\mathcal{P}$ which is dense and open with respect to the C^2_S-topology.

From the well-known genericity results for nonlinear programming problems, see [22], we obtain the following genericity results, see [53].

Theorem 4.4.1 (*cf.* **[53]**) *Generically for problems P_{CC} the following holds. For any feasible point $\overline{x} \in \mathcal{M}_{CC}$ the MPCC-LICQ condition is satisfied. Moreover any g.c. point is a non-degenerate critical point in the MPCC-sense, i.e., the conditions MPCC-LICQ, MPCC-SC and MPCC-SOC are fulfilled.*

Proof. Let us fix a partition (I_r, I_{rs}, I_s) of $\{1, \ldots, l\}$ corresponding to the active index triplet of a point $x \in \mathcal{M}_{CC}$. Let $P_R(x)$ be the associated relaxed problem. Generically by Theorem 2.2.2, at all feasible points of $P_R(x)$, LICQ holds and all its generalized critical points are non-degenerate, see Definition 2.2.5. We take the (finite) intersection of all these generic sets of functions, corresponding to all finitely many possible partitions (I_r, I_{rs}, I_s). Then the set of functions (f, g, r, s), such that for all corresponding possible problems $P_R(x)$ the LICQ condition holds at all its feasible points and all generalized critical points are non-degenerate, is generic.
Now, if $\overline{x} \in \mathcal{M}_{CC}$, then $\overline{x}$ is feasible for $P_R(\overline{x})$ associated with $I_r(\overline{x}), I_{rs}(\overline{x}), I_s(\overline{x})$. As LICQ holds generically for $\mathcal{M}_R$, the MPCC-LICQ condition is fulfilled for $\mathcal{M}_{CC}$. Moreover, if $\overline{x}$ is a g. c. point of P_{CC}, then it is also a generalized critical point of $P_R(\overline{x})$ (see Definition 2.2.3) and thus, generically, a non-degenerate critical point of P_{CC}.

□

4.4.2 Optimality conditions for problems P_{CC}

In this part we are interested in necessary and sufficient optimality conditions for minimizers of order one and two for problems P_{CC} in (4.1.2), see Definition 2.2.1. Using the *piecewise description*, see Lemma 4.2.1, all standard optimality conditions for $P_I(\overline{x})$ can directly be translated into corresponding results for P_{CC}.

We begin with characterizations for minimizers of order one.

Theorem 4.4.2 (Primal conditions of order 1) *For $\overline{x} \in \mathcal{M}_{CC}$:*

$$C_{\overline{x}} = \{0\} \quad \Rightarrow \quad \overline{x} \textit{ is a (isolated) local minimizer of order 1 of } P_{CC}.$$

If MPCC-LICQ holds at $\overline{x}$ also the converse is true.

Proof. It is well-known, see *e.g.* Still and Streng [Th.3.2,3.6] [62], that $C_I(\overline{x}) = \{0\}$ implies that $\overline{x}$ is an (isolated) local minimizer of order 1 of $P_I(\overline{x})$ and under MPCC-LICQ the converse holds. With regard to the definition of $C_{\overline{x}}$ in (4.3.4) the result follows from Lemma 4.2.1.

□

Theorem 4.4.3 (Dual conditions of order 1) *Let MPCC-LICQ hold at $\overline{x} \in \mathcal{M}_{CC}$. Then $\overline{x}$ is an (isolated) local minimizer of order 1 of P_{CC} if and only if one of the following equivalent conditions (a) or (b) holds:*

(a) $\nabla f(\overline{x}) \in \text{int } \mathcal{Q}_{\,\overline{x}}$, *where*

$$\mathcal{Q}_{\,\overline{x}} = \left\{ \begin{array}{c} d = \sum\limits_{j\in J(\overline{x})} \mu_j \nabla g_j(\overline{x}) + \sum\limits_{i\in I_{rs}(\overline{x})} \big(\rho_i \nabla r_i(\overline{x}) + \sigma_i \nabla s_i(\overline{x})\big) + \\ + \sum\limits_{i\in I_r(\overline{x})} \rho_i \nabla r_i(\overline{x}) + \sum\limits_{i\in I_s(\overline{x})} \sigma_i \nabla s_i(\overline{x}) \\ \mu_j \geq 0,\ j\in J_0(\overline{x}),\ \rho_i \geq 0, \sigma_i \geq 0,\ i \in I_{rs}(\overline{x}). \end{array} \right\}$$

(b) The vector $\overline{x}$ is a B-stationary point with (unique) multipliers $1, \mu, \rho, \sigma$ such that $|J_0(\overline{x})| + 2|I_{rs}(\overline{x})| + |I_r(\overline{x})| + |I_s(\overline{x})| = n$ and $\mu_j > 0,\ j \in J_0(\overline{x}),\ \rho_i > 0,$ $\sigma_i > 0,\ i \in I_{rs}(\overline{x})$.

Proof. It is well-known that the primal condition $C_I(\overline{x}) = \{0\}$ is equivalent with the condition $\nabla f(\overline{x}) \in \text{int } \mathcal{Q}_{\,I}(\overline{x})$, where

$$\mathcal{Q}_{\,I}(\overline{x}) \;=\; \left\{ \begin{array}{c} d \;=\; \sum\limits_{j\in J_0(\overline{x})} \mu_j \nabla g_j(\overline{x}) + \sum\limits_{i\in I_{rs}(\overline{x})} \big(\rho_i \nabla r_i(\overline{x}) + \sigma_i \nabla s_i(\overline{x})\big) + \\ + \sum\limits_{i\in I_r(\overline{x})} \rho_i \nabla r_i(\overline{x}) + \sum\limits_{i\in I_s(\overline{x})} \sigma_i \nabla s_i(\overline{x}) \\ \mu_j \geq 0,\ j\in J_0(\overline{x}),\ \rho_i \geq 0,\ i \in I_{rs}(\overline{x}) \setminus I,\ \sigma_i \geq 0,\ i \in I \end{array} \right\}$$

Applying Lemma 4.2.1 under MPCC-LICQ, yields (a).

(b) We now prove that under MPCC-LICQ (a) $\Leftrightarrow$ (b). Note that the sufficiency is evident. To prove the converse let us assume that $\nabla f(\overline{x}) \in \mathcal{Q}_{\,I}(\overline{x})$ and $|J_0(\overline{x})| + 2|I_{rs}(\overline{x})| + |I_r(\overline{x})| + |I_s(\overline{x})| < n$. This means that there exists $d,\ d \in \mathbb{R}^n$, such that $d \notin S_0$, where S_0 is the subspace generated by the vectors $\nabla g_j(\overline{x}),\ j \in J_0(\overline{x}),\ \nabla r_i(\overline{x}),\ i \in I_{rs}(\overline{x}) \cup I_r(\overline{x}),\ \nabla s_i(\overline{x}),\ i \in I_{rs}(\overline{x}) \cup I_s(\overline{x})$. Note that, since $(\overline{x}, 1, \mu, \rho, \sigma)$ is a B-stationary solution, it follows that $S_0 =$ span $\{\{-\nabla f(\overline{x})\} \cup S_0\}$. Consequently, for any $\varepsilon > 0$, $\varepsilon d \notin$ span $\{\{-\nabla f(\overline{x})\} \cup S_0\}$ and thus $\nabla f(\overline{x}) + \varepsilon d \notin$ span S_0 in contradiction to (a). Let us now assume that MPCC-SC does not hold, say $\mu_1 = 0$. Then, by MPCC-LICQ, for any $\varepsilon > 0$ the vector $\nabla f(\overline{x}) - \varepsilon \nabla g_1(\overline{x})$ is not contained in $\mathcal{Q}_{\overline{x}}$, a contradiction.

$\square$

We now turn to sufficient and necessary optimality conditions of order two. For the sufficient part we refer also to [53].

Theorem 4.4.4 (Dual conditions of order 2) *Let MPCC-LICQ hold at $\overline{x} \in \mathcal{M}_{CC}$ and assume $C_{\overline{x}} \neq \{0\}$, i.e., $\overline{x}$ is not a local minimizer of order 1. Then $\overline{x}$ is an (isolated) local minimizer of order 2 of P_{CC} if and only if $\overline{x}$ is a B-stationary point of P_{CC} with (unique) multipliers $(1, \mu, \rho, \sigma)$ such that MPCC-SC is satisfied and the condition MPCC-SOC holds with $d^T \nabla_x^2 L(\overline{x}, 1, \mu, \rho, \sigma) d > 0$ $\forall d \in C_{\overline{x}} \setminus \{0\}$.*
Under this condition, $\overline{x}$ is locally a unique B-stationary point of P_{CC}.

Proof. $C_{\overline{x}} \neq \{0\}$ implies $C_I(\overline{x}) \neq \{0\}$ for (at least) one set $I \subset I_{rs}(\overline{x})$ and by [62, Th.3.6], under LICQ, $\overline{x}$ is a (isolated) local minimizer of order 2 of $P_I(\overline{x})$ if and only if $\overline{x}$ is a KKT point satisfying the second order sufficient condition at problem $P_I(\overline{x})$ for all $I \subset I_{rs}(\overline{x})$. Again the result follows from Lemma 4.2.1.

□

Note that in view of the genericity result in Theorem 4.4.1 the following is true:

Corollary 4.4.1 *Generically each local minimizer of P_{CC} is either of order 1 or of order 2.*

It is interesting to note that with respect to the relaxed problem $P_R(\overline{x})$ (4.3.3) the following holds:

Corollary 4.4.2 *Let MPCC-LICQ be satisfied at $\overline{x} \in \mathcal{M}_{CC}$. Then $\overline{x}$ is a local minimizer of order 1 or 2 of P_{CC} if and only if $\overline{x}$ is a local minimizer of order 1 or 2 of $P_R(\overline{x})$ (4.3.3).*

Proof. Under MPCC-LICQ any local minimizer $\overline{x}$ of P_{CC} must be a strongly stationary point of P_{CC} with unique multipliers $(1, \mu, \rho, \sigma)$. Note that, as the Lagrange functions of P_R and P_{CC} are the same, $(\overline{x}, 1, \mu, \rho, \sigma)$ is also a KKT solution of $P_R(\overline{x})$. Moreover the set of critical directions for $P_R(\overline{x})$ coincides with $C_{\overline{x}}$, see (4.3.7). So the first order optimality condition $C_{\overline{x}} = \{0\}$, *cf.* Theorem 4.4.2, and the second order conditions, *cf.* Theorem 4.4.4, of P_{CC} and $P_R(\overline{x})$ coincide.

□

4.5 A parametric solution method

As it has been already remarked, to solve P_{CC} (see (4.1.2)) a parametric approach can be used, see also *e.g.* [16]. Here we consider the perturbed problem

$$\begin{array}{lll} \mathrm{P}_\tau: & & \min\limits_{x} \ f(x) \\ & \text{s.t.} & x \in \mathcal{M}_\tau \end{array} \tag{4.5.1}$$

$$\mathcal{M}_\tau = \left\{ x \in \mathbb{R}^n \;\middle|\; \begin{array}{rcll} g_j(x) & \geq & 0, & j = 1, \ldots, q, \\ r_i(x)s_i(x) & = & \tau, & i = 1, \ldots, l, \\ r_i(x), s_i(x) & \geq & 0, & i = 1, \ldots, l \end{array} \right\}$$

where τ, $\tau > 0$, is a perturbation parameter. We expect that, by letting $\tau \to 0$, there is a solution $\overline{x}_\tau$ of P_τ converging to a solution $\overline{x}$ of P_{CC}. In the following we intend to analyze this approach. Under natural (generic) assumptions it will be shown that we generally can expect a rate

$$\|\overline{x}_\tau - \overline{x}\| = \mathcal{O}(\sqrt{\tau}).$$

Other regularizations of MPCC problems have been considered in the literature such as

$$
\begin{array}{llrcll}
\mathrm{P}_\tau^{\leqq}: & & \multicolumn{4}{c}{\min_x \; f(x)} \\
& \text{s.t.} & g_j(x) & \geq & 0, & j = 1, \ldots, q, \\
& & r_i(x)s_i(x) & \leq & \tau, & i = 1, \ldots, l, \\
& & r_i(x), s_i(x) & \geq & 0, & i = 1, \ldots, l.
\end{array}
\tag{4.5.2}
$$

$$
\begin{array}{llrcll}
\hat{P}_\tau^{\leqq}: & & \multicolumn{4}{c}{\min_x f(x)} \\
& \text{s.t.} & g_j(x) & \geq & 0, & j = 1, \ldots, q, \\
& & r^T(x)s(x) & \leq & \tau, & \\
& & r_i(x), s_i(x) & \geq & 0, & i = 1, \ldots, l.
\end{array}
\tag{4.5.3}
$$

In [51], assumptions under which stationary points $x(\tau)$ of $\mathrm{P}_\tau^{\leqq}$, $\tau \downarrow 0$, converge to a B-stationary point of P_{CC} are given. In [47] it is shown that (under natural conditions) the solutions $x(\tau)$ of $\mathrm{P}_\tau^{\leqq}$ converge to a (nearby) solution $\overline{x}$ of P_{CC} with order $\mathcal{O}(\tau)$. Similar results are stated for the problem $\hat{P}_\tau^{\leqq}$.

Remark 4.5.1 *We emphasize that these regularizations $P_\tau^{\leqq}$, $\hat{P}_\tau^{\leqq}$ structurally completely differ from the smoothing approach P_τ. For $P_\tau^{\leqq}$, e.g., the following is shown in [51, Th.3.1,Cor.3.2]: If $\overline{x}$ is a solution of P_{CC} where MPCC-LICQ and MPCC-SC hold then for the (nearby) minimizers $\hat{x}_\tau$ of $P_\tau^{\leqq}$ (for τ small enough) the complementarity constraints $r_i(x)\ s_i(x) \leq \tau$, $i \in I_{rs}(\overline{x})$, are not active. More precisely,*

$$
r_i(\hat{x}_\tau) = s_i(\hat{x}_\tau) = 0 \ , \quad \forall i \in I_{rs}(\overline{x}),
$$

is true. This fact also directly follows from Corollary 4.4.2. In particular, in the case $I_{rs}(\overline{x}) = \{1, \ldots, l\}$ (for all small $\tau > 0$) the solution $\hat{x}_\tau$ of $P_\tau^{\leqq}$ coincides with the solution $\overline{x}$ of P_{CC}.

Note that the perturbation $r_i(x)s_i(x) = \tau$ in P_τ is also different from the following parametric MPCC, considered in [24] (see also Section 4.7)

$$
\begin{array}{llrcll}
P(\tau): & & \multicolumn{4}{c}{\min_x f(x, \tau)} \\
& \text{s.t.} & g_j(x, \tau) & \geq & 0, & j = 1, \ldots, q, \\
& & r_i(x, \tau) \cdot s_i(x, \tau) & = & 0, & i = 1, \ldots, l, \\
& & r_i(x, \tau)\ ,\ s_i(x, \tau) & \geq & 0, & i = 1, \ldots, l.
\end{array}
$$

with $f, g_j, r_i, s_i \in C^2$ (w.r.t. all variables). Near a non-degenerate solution $\overline{x}$ of $P(\overline{\tau})$, the value function will depend smoothly on the parameter τ. However, for the problem P_τ in (4.5.1), the value function does not depend smoothly on τ, $\tau \approx 0$, even in the non-degenerate case, as can be seen in Example 4.5.1 later on.

Remark 4.5.2 *We recall that under appropriate LICQ conditions all results for 4.1.2 remain true for problems P_{CC} in (4.1.1), with additional equality constraints.*

Let in the following φ, φ_τ denote the marginal values and $\mathcal{M}_{CC}$, $\mathcal{M}_\tau$ the feasible sets of P_{CC}, P_τ respectively.

4.5.1 Motivating examples

We begin with some illustrative examples.

Example 4.5.1

$$\begin{array}{rrcl} & \min x_1 + x_2 & & \\ s.t. & x_1\, x_2 & = & 0, \\ & x_1, x_2 & \geq & 0. \end{array}$$

Here the set $\mathcal{M}_\tau$ *converges* to the set $\mathcal{M}_{CC}$ and the solutions $x_\tau = (\sqrt{\tau}, \sqrt{\tau})$ of P_τ converge to the solution $\overline{x} = 0$ of the original problem with a rate

$$\|x_\tau - \overline{x}\| = \sqrt{2} \cdot \sqrt{\tau}.$$

In general, we cannot expect that the solutions of P_τ converge to a solution of P_{CC}. As we can see in the next example, they may even not converge to a feasible point if the feasible set is non-compact:

Example 4.5.2 Consider

$$\begin{array}{rrcl} & \min\ (x_2 - 1)^2 & & \\ s.t. & x_2 e^{-x_1} & = & 0, \\ & x_2,\ e^{-x_1} & \geq & 0 \end{array}$$

with $\mathcal{M}_{CC} = \{(x_1, 0),\ x_1 \in \mathbb{R}\}$, and

$$\begin{array}{rrcl} & \min\ (x_2 - 1)^2 & & \\ s.t. & (x_1^2 + x_2^2) e^{-x_1} & = & 0, \\ & (x_1^2 + x_2^2),\ e^{-x_1} & \geq & 0 \end{array}$$

with $\mathcal{M}_{CC} = \{(0, 0)\}$.

In both examples the feasible set and the set of minimizers coincide.

Note that in the first example $x_2 = 0 < e^{-x_1}$ and $\nabla x_2 \neq 0$. So MPCC-LICQ holds. However, for all $\tau,\ \tau > 0$, the point $(-ln(\tau), 1) \in M_\tau$ is the global minimizer of P_τ, and these points do not approach to the set of minimizers of P_{CC}.

In the second example, $\mathcal{M}_{CC}$ is bounded but it can be proven that $\mathcal{M}_\tau$ is not.

The next example shows that to assure $\mathcal{M}_\tau \neq \varnothing$, the MPCC-LICQ condition cannot be relaxed by the fulfillment of $MFCQ$ in $\mathcal{M}_R$, see (4.3.3).

Example 4.5.3 Consider the feasible set defined by $r_1(x) = x,\ r_2(x) = e^x - 1$, and $s_1(x) = s_2(x) = 1$, *i.e.*,

$$\mathcal{M}_{CC} = \left\{ x \in \mathbb{R} \;\middle|\; \begin{array}{rcl} x \cdot 1 & = & 0, \\ (e^x - 1) \cdot 1 & = & 0, \\ x,\ (e^x - 1),\ 1 & \geq & 0. \end{array} \right\}$$

The unique solution of the complementarity problem is $\overline{x} = 0$. However, $M_\tau = \varnothing$, for any $\tau,\ \tau > 0$. Moreover MFCQ holds at $\overline{x}$ for $\mathcal{M}_R$

The last example shows that a bad convergence rate of the smoothing approach is possible.

Example 4.5.4 (*cf.* **[47])**

$$\begin{array}{rrcl} & \min x_1^k + x_2 & & \\ s.t. & x_1 x_2 & = & 0, \\ & x_1, x_2 & \geq & 0, \end{array}$$

with $k > 0$. Here the minimizer is $\overline{x} = (0,\ 0)$ and the solution of the corresponding problem P_τ is $\overline{x}_\tau = ((\frac{\tau}{k})^{\frac{1}{k+1}}, k^{\frac{1}{k+1}} \tau^{\frac{k}{k+1}})$. So the convergence rate is $O(\tau^{\frac{1}{k}})$ with arbitrary large $k > 0$. Note that for $k > 1$, the MPCC-SC condition fails at the minimizer $\overline{x} = (0, 0)$.

In the following we are interested in the convergence of the feasible sets $\mathcal{M}_\tau$, the value function φ_τ and the solutions $\overline{x}_\tau$ of P_τ, *i.e.*, we study the convergence

$$\mathcal{M}_\tau \to \mathcal{M}_{CC},\ \varphi_\tau \to \varphi \quad \text{and} \quad \overline{x}_\tau \to \overline{x} \quad \text{for } \tau \to 0.$$

We will also obtain the rate of this convergence.

Firstly, motivated by Example 4.5.2, we restrict the feasible set to a compact subset $X,\ X \subset \mathbb{R}^n$. Note that, in practice, this does not mean a restriction since it is advisable to add (if necessary) *e.g.* box constraints, $x_\nu \leq \pm K,\ \nu = 1, \ldots, n$, for some large number K. So, we will assume:

$$\mathbf{A}_0. \quad M_\tau \subset X \text{ for all } \tau \geq 0 \text{ where } X \text{ is a compact subset of } \mathbb{R}^n. \tag{4.5.4}$$

4.5.2 The convergence behavior of the feasible set

In this part, we consider the convergence behavior of the feasible set $\mathcal{M}_\tau$. We will show that, under natural assumptions on the problem P_{CC} for all $x \in \mathcal{M}_{CC}$, $d(x, \mathcal{M}_\tau) = \mathcal{O}(\sqrt{\tau})$ holds and $d(x_\tau, \mathcal{M}_{CC}) = \mathcal{O}(\sqrt{\tau})$, for $x_\tau \in \mathcal{M}_\tau,\ \tau \to 0$.

We define the active index sets $(J_0(x), I_r(x), I_s(x), I_{rs}(x))$ of $x \in \mathcal{M}_{CC}$, as in the Section 4.2. To avoid the bad behavior in Example 4.5.2 and Example 4.5.3

(*cf.* Section 4.5.1) we will assume that condition A_0 (see (4.5.4)), holds and that MPCC-LICQ is fulfilled for $\mathcal{M}_{CC}$, globally or locally. We emphasize that these assumptions are generically satisfied, see Theorem 4.4.1. We begin with a general lemma.

Lemma 4.5.1 *For any sequence $x_\tau \in \mathcal{M}_\tau$ and $\tau \to 0$ it follows that $d(x_\tau, \mathcal{M}_{CC}) \to 0$ uniformly: i.e., to any $\varepsilon > 0$ there exists $\tau_0,\ \tau_0 > 0$ such that for all $\tau,\ 0 < \tau \leq \tau_0$, and $x_\tau \in \mathcal{M}_\tau$, the bound $d(x_\tau, \mathcal{M}_{CC}) < \varepsilon$ holds.*

Proof. Assuming that the statement is not true, there must exist $\gamma,\ \gamma > 0$ and a sequence $x_\tau \in \mathcal{M}_\tau$, such that, for $\tau \to 0$,

$$d(x_\tau, \mathcal{M}_{CC}) \geq \gamma.$$

Due to the compactness assumption A_0 we can choose a convergent subsequence $x_\tau \to \overline{x} \in X$. The condition $r_i(x_\tau)s_i(x_\tau) = \tau$, together with the continuity of the functions r_i, s_i leads, for $\tau \to 0$, to $r_i(\overline{x})s_i(\overline{x}) = 0$, *i.e.*, $\overline{x} \in \mathcal{M}_{CC}$, a contradiction.
□

To prove our main results on the behavior of $\mathcal{M}_\tau$, we make use of a local diffeomorphism. Such a transformation has been applied in [53] to illustrate the local behavior of $\mathcal{M}_{CC}$. Here we present a complete global analysis. The use of this transformation makes the proofs of the main results technically much simpler, however this approach relies on the MPCC-LICQ assumption.

Consider a point $\overline{x} \in \mathcal{M}_{CC}$ with $|J_0(\overline{x})| = \hat{q},\ 2|I_{rs}(\overline{x})| + |I_r(\overline{x})| + |I_s(\overline{x})| = l + p$ where $p \leq l,\ \hat{q} \leq q$ and $l + p + \hat{q} \leq n$. W.l.o.g. we can assume:

$$J_0(\overline{x}) = \{1, \ldots, \hat{q}\},\ I_{rs}(\overline{x}) = \{1, \ldots, p\},\ I_r(\overline{x}) = \{p+1, \ldots, l\},\ \ I_s(\overline{x}) = \emptyset. \tag{4.5.5}$$

By the MPCC-LICQ condition, $\nabla g_1(\overline{x}), \ldots, \nabla g_{\hat{q}}(\overline{x})$, $\nabla r_i(\overline{x}),\ i = 1, \ldots, l$, $\nabla s_i(\overline{x})$, $i = 1, \ldots, p$ are linearly independent and we can complete these vectors to a basis of $\mathbb{R}^n$ by adding vectors $v_i,\ i = l + p + \hat{q} + 1, \ldots, n$. Now we define the transformation $y = T(x)$ by:

$$\begin{array}{rclllrcll} y_i &=& r_i(x), & i = 1, \ldots, l, & y_{i+l} &=& s_i(x), & i = 1, \ldots, p, \\ y_{l+p+j} &=& g_j(x), & j = 1, \ldots, \hat{q}, & y_i &=& v_i^T(x - \overline{x}), & i > l + p + \hat{q}. \end{array} \tag{4.5.6}$$

By construction the Jacobian $\nabla T(\overline{x})$ is regular. So T defines locally a diffeomorphism. This means that there exists $\varepsilon,\ \varepsilon > 0$, and neighborhoods $B_\varepsilon^n(\overline{x})$ of $\overline{x}$ and $U_\varepsilon(\overline{y}) := T(B_\varepsilon^n(\overline{x}))$ of $\overline{y} = 0$ such that $T : B_\varepsilon^n(\overline{x}) \to U_\varepsilon(\overline{y})$ is a bijective mapping with $T, T^{-1} \in C^1$, $T(\overline{x}) = \overline{y}$ and for $y = T(x)$ it follows that

$$x \in \mathcal{M}_\tau \cap B_\varepsilon^n(\overline{x}) \quad \Leftrightarrow \quad \begin{array}{rcll} y_{l+p+j} &\geq& 0, & j = 1, \ldots, \hat{q}, \\ y_i \cdot y_{l+i} &=& \tau, & i = 1, \ldots, p, \\ y_i \cdot \widetilde{s}_i(y) &=& \tau, & i = p+1, \ldots, l, \\ y_i &\geq& 0, & i = 1, \ldots, l+p, \\ y &\in& U_\varepsilon(\overline{y}). & \end{array}$$

where $\widetilde{s}_i(\overline{y}) = s_i(T^{-1}(\overline{y})) = s_i(\overline{x}) =: c_i^s,\ c_i^s > 0,\ i = p+1,\dots,l$ and $\widetilde{g}_j(\overline{y}) = g_j(T^{-1}(\overline{y})) = g_j(\overline{x}) =: c_j^g,\ c_j^g > 0,\ j = \hat{q}+1,\dots,q$.

In particular, since T is a diffeomorphism, the distance between two points remains equivalent in the sense that with some constants $0 < \kappa_- < \kappa_+$:

$$\kappa_-\|y_1-y_2\| \le \|x_1-x_2\| \le \kappa_+\|y_1-y_2\|, \quad \forall x_1,x_2 \in B_\varepsilon^n(\overline{x}),\ y_1 = T(x_1),\ y_2 = T(x_2).$$

So, after applying a diffeomorphism T, we may assume $\overline{x} = 0$ and that there is some $\varepsilon,\ \varepsilon > 0$, such that for $x \in B_\varepsilon^n(\overline{x})$:

$$x \in \mathcal{M}_\tau \quad \Leftrightarrow \quad \begin{array}{rcll} g_j(x) = x_{l+p+j} & \geq & 0, & j = 1,\dots,\hat{q}, \\ \hat{h}_i(x) = x_i \cdot x_{l+i} & = & \tau, & i = 1,\dots,p, \\ \hat{h}_i(x) = x_i \cdot s_i(x) & = & \tau, & i = p+1,\dots,l, \\ x_i & \geq & 0, & i = 1,\dots,l+p. \end{array} \tag{4.5.7}$$

Moreover, since $s_i(\overline{x}) =: c_i^s,\ c_i^s > 0,\ i = p+1,\dots,l$ and $g_j(\overline{x}) =: c_j^g,\ c_j^g > 0$, $j = \hat{q}+1,\dots,q$, by choosing ε small enough we also can assume:

$$s_i(x) \geq \frac{c_i^s}{2},\ i = p+1,\dots,l \ \text{ and } \ g_j(x) \geq \frac{c_j^g}{2},\ j = \hat{q}+1,\dots,q,\ \ \forall x \in B_\varepsilon^n(\overline{x}). \tag{4.5.8}$$

Lemma 4.5.2 *Let MPCC-LICQ hold at $\overline{x} \in \mathcal{M}_{CC}$.*
(a) Then there exist $\varepsilon, \tau_0, \alpha, \beta > 0$ such that for all $\tau,\ 0 < \tau \le \tau_0$ the following holds: there exists $\overline{x}_\tau \in \mathcal{M}_\tau$ with

$$\|\overline{x}_\tau - \overline{x}\| \le \alpha\sqrt{\tau} \tag{4.5.9}$$

and, for any $\overline{x}_\tau \in \mathcal{M}_\tau \cap B_\varepsilon^n(\overline{x})$, there exists a point $\hat{x}_\tau,\ \hat{x}_\tau \in \mathcal{M}_{CC} \cap B_\varepsilon^n(\overline{x})$ satisfying

$$\|\hat{x}_\tau - \overline{x}_\tau\| \le \beta\sqrt{\tau}. \tag{4.5.10}$$

Moreover, if SC holds at $\overline{x}$ (cf. Section 4.3), the statements are true with $\sqrt{\tau}$ replaced by τ.

(b) If the condition SC is not fulfilled at $\overline{x}$ then the convergence rate $\mathcal{O}(\sqrt{\tau})$ in (4.5.9) is optimal. More precisely, there is some $\gamma,\ \gamma > 0$, such that for all $\overline{x}_\tau \in \mathcal{M}_\tau$ the relation $\|\overline{x}_\tau - \overline{x}\| \geq \gamma\sqrt{\tau}$ holds for all small τ.

Proof. (a) Let MPCC-LICQ hold at $\overline{x} \in \mathcal{M}_{CC}$. As discussed before (after applying diffeomorphism) we can assume that $\overline{x} = 0$ and that in a neighborhood $B_\varepsilon^n(\overline{x})$ of $\overline{x}$ the set $B_\varepsilon^n(\overline{x}) \cap \mathcal{M}_\tau$ is described by (4.5.7). To construct an element $x^\tau \in \mathcal{M}_\tau$, we fix the components $x_i^\tau = x_{l+i}^\tau = \sqrt{\tau},\ i = 1,\dots,p$ and $x_i^\tau = 0$, $i = l+p+1,\dots,n$. From (4.5.7) we then find:

$$\begin{array}{rcll} g_j(x^\tau) & = & 0, & j = 1,\dots,\hat{q}, \\ \hat{h}_i(x^\tau) & = & \tau, & i = 1,\dots,p, \\ \hat{h}_i(x^\tau) = x_i^\tau \cdot s_i(x^\tau) & = & \tau, & i = p+1,\dots,l. \end{array}$$

We only need to consider the remaining relations

$$\hat{h}_i(\widetilde{x}) := x_i^\tau \cdot s_i(\widetilde{x}) = \tau, \quad i = p+1, \dots, l. \tag{4.5.11}$$

which (for fixed τ) only depend on the remaining variables $\widetilde{x} = (x_{p+1}^\tau, \dots, x_l^\tau)$. For $\widetilde{x} = 0$ the gradients $\nabla \hat{h}_i(0) = e_i s_i(0) = e_i c_i^s$, $i = p+1, \dots, l$, are linearly independent. So the function $H : \mathbb{R}^{l-p} \to \mathbb{R}^{l-p}$, $H = (\hat{h}_{p+1}, \dots, \hat{h}_l)$ with $H(0) = 0$ has locally near $\widetilde{x} = 0$ a C^1 inverse such that (for small τ) the vector $\widetilde{x}^\tau := H^{-1}(e\,\tau)$, with $e = (1, \dots, 1) \in \mathbb{R}^{l-p}$, defines a solution of (4.5.11). Because of $H^{-1}(0) = 0$ it follows that $\|\widetilde{x}^\tau\| = \mathcal{O}(\tau)$.

Altogether with the other fixed components, this vector $\widetilde{x}^\tau$ defines a feasible point $x_\tau \in \mathcal{M}_\tau$ which satisfies

$$\|x_\tau - \overline{x}\| \le \mathcal{O}(\sqrt{\tau}).$$

We now prove (4.5.10). As shown above, see (4.5.7), for some ε, $\varepsilon > 0$, the point $\overline{x}_\tau \in B_\varepsilon^n(\overline{x})$ is in $\mathcal{M}_\tau$ if and only if $x := \overline{x}_\tau$ satisfies the relations

$$\begin{array}{rcll} g_j(x) = x_{l+p+j} & \ge & 0, & j = 1, \dots, \hat{q}, \\ h_i(x) = x_i \cdot x_{l+i} & = & \tau, & i = 1, \dots, p, \\ h_i(x) = x_i \cdot s_i(x) & = & \tau, & i = p+1, \dots, l. \end{array}$$

Obviously, $\min\{x_i, x_{l+i}\} \le \sqrt{\tau}$, $i = 1, \dots, p$, so w.l.o.g. we assume $x_i \le \sqrt{\tau}$, $i = 1, \dots, p$. By (4.5.8) for $x = \overline{x}_\tau \in B_\varepsilon^n(\overline{x})$ it follows that

$$x_i = \frac{\tau}{s_i(x)} \le \frac{\tau}{c_i^s/2} \le \frac{\tau}{c^s}\,, \qquad i = p+1, \dots, l, \tag{4.5.12}$$

where $c_s = \min\{c_i^s/2,\ i = p+1, \dots, l\}$. Given this element $x = \overline{x}_\tau \in \mathcal{M}_\tau$ we now choose the point $\hat{x}_\tau$ of the form $\hat{x}_\tau = (0, \dots, 0, x_{l+1}, \dots, x_n)$ which is contained in $\mathcal{M}_{CC}$. By using (4.5.12) and $x_i \le \sqrt{\tau}$, $i = 1, \dots, p$, we find ($x = \overline{x}_\tau$)

$$\|\hat{x}_\tau - \overline{x}_\tau\| \le \sqrt{p\,\tau + (l-p)\,\frac{\tau^2}{c_s^2}} \le \mathcal{O}(\sqrt{\tau}).$$

Let now SC be satisfied at $\overline{x} \in \mathcal{M}_{CC}, \overline{x} = 0$. Then locally in $B_\varepsilon^n(\overline{x})$, see above, the set $\mathcal{M}_\tau$ is defined by

$$\begin{array}{rcll} g_j(x) = x_{l+p+j} & \ge & 0, & j = 1, \dots, \hat{q}, \\ x_i \cdot s_i(x) & = & \tau, & i = 1, \dots, l, \end{array} \tag{4.5.13}$$

and $s_i(x) \ge c_1^s/2$, for all $x \in B_\varepsilon^n(\overline{x})$. As in the first part of proof, we can fix the coefficients of $x = x_\tau$ by $x_{l+i} = \overline{x}_i (= 0)$, $i = l+1, \dots, n$ and find a solution $x \in \mathcal{M}_\tau$ by applying the inverse function theorem to the remaining l equations

$$\hat{h}_i(\widetilde{x}) := x_i s_i(\widetilde{x}) = \tau\ , \quad i = 1, \dots, l.$$

only depending on the remaining variables $\widetilde{x} := (x_1, \ldots, x_l)$. This provides us with a solution $x = x_\tau$ of (4.5.13) with

$$\|x_\tau - \overline{x}\| = \mathcal{O}(\tau).$$

On the other hand for any solution $x := \overline{x}_\tau$ of (4.5.13) in $B^n_\varepsilon(\overline{x})$ the point $\hat{x}_\tau = (0, \ldots, 0, x_{l+1}, \ldots, x_n)$ is an element in $\mathcal{M}_{CC}$ with $\|\hat{x}_\tau - \overline{x}_\tau\| = \mathcal{O}(\tau)$.

(b) Suppose that SC is not fulfilled at $\overline{x}$, *i.e.*, for some $i_0 \in \{1, \ldots, l\}$, see (a),

$$\hat{h}_i(\overline{x}) = \overline{x}_{i_0} \cdot \overline{x}_{l+i_0} = 0 \quad \text{with} \quad \overline{x}_{i_0} = \overline{x}_{l+i_0} = 0.$$

Then near $\overline{x}$ any point $\hat{x} := x_\tau \in \mathcal{M}_\tau$ must satisfy $\hat{x}_{i_0} \cdot \hat{x}_{l+i_0} = \tau$ which implies

$$\|x_\tau - \overline{x}\| \geq \max\{\hat{x}_{i_0}\ ,\ \hat{x}_{l+i_0}\} \geq \sqrt{\tau}.$$

Recall that, due to the applied diffeomorphism, this inequality only holds up to a positive constant. □

Lemma 4.5.2 yields the local convergence of $\mathcal{M}_\tau$, near a point $\overline{x} \in \mathcal{M}_{CC}$. We now are interested in the global convergence behavior.

Lemma 4.5.3 *Let $\mathbf{A}_0$ hold, see (4.5.4). Suppose MPCC-LICQ is satisfied at each point $\overline{x} \in \mathcal{M}_{CC}$. Then there are $\tau_0, \alpha, \beta > 0$ such that for all τ, $0 < \tau \leq \tau_0$, the following holds:*
For each $\overline{x} \in \mathcal{M}_{CC}$ there exists $\overline{x}_\tau$, $\overline{x}_\tau \in \mathcal{M}_\tau$ with

$$\|\overline{x}_\tau - \overline{x}\| \leq \alpha\sqrt{\tau}, \tag{4.5.14}$$

and for any $\overline{x}_\tau \in \mathcal{M}_\tau$ there exists a point $\hat{x}_\tau$, $\hat{x}_\tau \in \mathcal{M}_{CC}$ satisfying

$$\|\hat{x}_\tau - \overline{x}_\tau\| \leq \beta\sqrt{\tau}. \tag{4.5.15}$$

Moreover, if SC holds at all $\overline{x} \in \mathcal{M}_{CC}$ the statements are true with $\sqrt{\tau}$ replaced by τ.

Proof. We firstly prove (4.5.15). Recall the local transformation constructed above near any point $\overline{x} \in \mathcal{M}_{CC}$, see 4.5.7. The union $\cup_{\overline{x} \in \mathcal{M}_{CC}} B^n_{\varepsilon(\overline{x})}(\overline{x})$ forms an open cover of the compact feasible set $\mathcal{M}_{CC} \subset X$. Consequently we can choose a finite cover, *i.e.*, points $x_\nu \in \mathcal{M}_{CC}$, $\nu = 1, \ldots, N$, such that with $\varepsilon_\nu = \varepsilon(x_\nu)$ the set $\cup_{\nu=1,\ldots,N} B^n_{\varepsilon_\nu}(x_\nu)$ provides an open cover of $\mathcal{M}_{CC}$ and with $\beta_\nu > 0$ the corresponding condition (4.5.10) holds.
By defining $B_\varepsilon(\mathcal{M}_{CC}) = \{x \in X \mid d(x, \mathcal{M}_{CC}) < \varepsilon\}$ we can choose some ε_0, $\varepsilon_0 > 0$ (small) such that

$$B_{\varepsilon_0}(\mathcal{M}_{CC}) \subset \bigcup_{\nu=1,\ldots,N} B^n_{\varepsilon_\nu}(x_\nu).$$

By choosing $\varepsilon = \varepsilon_0$ and τ_0 in Lemma 4.5.1 we find for all $\tau,\ 0 \le \tau \le \tau_0$,

$$\mathcal{M}_\tau \subset B_{\varepsilon_0}(\mathcal{M}_{CC}) \subset \bigcup_{\nu=1,\ldots,N} B^n_{\varepsilon_\nu}(x_\nu).$$

The second convergence result (4.5.15) directly follows by combining the finite cover argument with the local convergence and by noticing that we can choose as convergence constant the number $\beta = \max\{\beta_\nu;\ \ \nu = 1,\ldots,N\}$.

To prove (4.5.14) we have to show that the following sharpening of the local bound (4.5.9) holds: For $\overline{x} \in \mathcal{M}_{CC}$ there exist $\tau_0,\ \tau_0 > 0$ and $\varepsilon,\ \varepsilon > 0$, such that for any $x \in \mathcal{M}_{CC} \cap B^n_\varepsilon(\overline{x})$ and for any $\tau,\ 0 \le \tau \le \tau_0$, there is a point $x_\tau,\ x_\tau \in \mathcal{M}_\tau$ with

$$\|x_\tau - x\| \le \alpha\sqrt{\tau}. \tag{4.5.16}$$

Then a finite cover argument as above yields the global relation (4.5.14).

We only sketch the proof of (4.5.16). Let $\overline{x} \in \mathcal{M}_{CC}$ be fixed. In the proof of Lemma 4.5.2(a) we made use of a local diffeomorphism $T_{\overline{x}}(x)$ leading to relation (4.5.9). This transformation $T_{\overline{x}}$ is constructed depending on the active index set $I_a(\overline{x}) := I_{rs}(\overline{x}) \cup I_r(\overline{x}) \cup I_s(\overline{x}) \cup J_0(\overline{x})$, see (4.5.6). For any x near $\overline{x}$ we have $I_a(x) \subset I_a(\overline{x})$ and there are only finitely many choices for I_a namely $I_a^1,\ldots,I_a^R$. So if we fix $I_a^i,\ I_a^i \subset I_a(\overline{x})$, any point $\hat{x}$ near $\overline{x}$ with $I_a(\hat{x}) = I_a^i$ yields a local diffeomorphism $T_{\hat{x}}$ which depends smoothly on $\hat{x}$, see the construction (4.5.6). So we find a common bound. There exist $\alpha_i, \varepsilon_i > 0$ such that for any $x \in \mathcal{M}_{CC} \cap B^n_{\varepsilon_i}(\overline{x})$, with $I_a(x) = I_a^i$, there is a point $x_\tau,\ x_\tau \in \mathcal{M}_\tau$ such that, for all small τ

$$\|x_\tau - x\| \le \alpha_i\sqrt{\tau}.$$

Then, by choosing $\varepsilon = \min\{\varepsilon_i \mid i = 1,\ldots,R\}$, and $\alpha = \max\{\alpha_i \mid i = 1,\ldots,R\}$, we have shown the relation (4.5.16).

□

Remark 4.5.3 *Lemma 4.5.3 proves that the convergence in the Hausdorff distance*

$$d_H(\mathcal{M}_\tau, \mathcal{M}_{CC}) := \max\{\max_{x_\tau \in \mathcal{M}_\tau} d(x_\tau, \mathcal{M}_{CC})\ ,\ \max_{x \in \mathcal{M}_{CC}} d(x, \mathcal{M}_\tau)\}$$

between $\mathcal{M}_\tau$ and $\mathcal{M}_{CC}$ satisfies $d_H(\mathcal{M}_\tau, \mathcal{M}_{CC}) = \mathcal{O}(\sqrt{\tau})$.

The above convergence results for the feasible sets have been derived under MPCC-LICQ. Example 4.5.3 shows that some constraint qualification is needed. However the MPCC-LICQ condition can be slightly weakened. We show this on a simple special instance of a feasible set, defined by only one complementarity constraint:

$$\mathcal{M}^1 = \left\{ x \in \mathbb{R}^n \;\middle|\; \begin{array}{rcl} r(x)s(x) & = & 0, \\ r(x) & \geq & 0, \\ s(x) & \geq & 0. \end{array} \right\}$$

The corresponding perturbed feasible set is then

$$\mathcal{M}^1_\tau = \left\{ x \in \mathbb{R}^n \;\middle|\; \begin{array}{rcl} r(x)s(x) & = & \tau, \\ r(x) & \geq & 0, \\ s(x) & \geq & 0. \end{array} \right\}$$

We say that MFCQ' holds at $x \in \mathcal{M}^1$ if the common MFCQ condition holds at $\overline{x}$ for $\mathcal{M}^1_R = \left\{ x \in \mathbb{R}^n \;\middle|\; \begin{array}{rcl} r(x) & \geq & 0, \\ s(x) & \geq & 0 \end{array} \right\}$. Under this hypothesis we can prove the following result.

Proposition 4.5.1 *Let MFCQ' hold at a point $\overline{x} \in \mathcal{M}^1$. Then the following is true:*

(a) For any $\tau,\ \tau > 0$, small enough, there exists a point $x_\tau,\ x_\tau \in \mathcal{M}^1_\tau$ such that $\|\overline{x} - x_\tau\| = O(\sqrt{\tau})$.

(b) There are some $\varepsilon, \tau_0 > 0$ such that for any $x_\tau \in \mathcal{M}^1_\tau \cap B^n_\epsilon(\overline{x})$, $0 < \tau \leq \tau_0$, there exists $x^\tau_0,\ x^\tau_0 \in \mathcal{M}^1$, satisfying $\|x^\tau_0 - x_\tau\| = O(\sqrt{\tau})$.

Proof. (a) Let $\overline{x} \in \mathcal{M}^1$ be fixed. First we consider the case in which SC holds, *i.e.*, $r(\overline{x}) > 0,\ s(\overline{x}) = 0$ or $s(\overline{x}) > 0,\ r(\overline{x}) = 0$. Then the MFCQ' condition means $\nabla s(\overline{x}) \neq 0$ or $\nabla r(\overline{x}) \neq 0$, respectively, *i.e.*, MPCC-LICQ holds. So the proof is a consequence of Lemma 4.5.2.

If $r(\overline{x}) = s(\overline{x}) = 0$, due to the MFCQ' there is a vector $\xi,\ \xi \in \mathbb{R}^n$ such that:

$$\xi^T \nabla r(\overline{x}) > 0, \qquad \xi^T \nabla s(\overline{x}) > 0.$$

W.l.o.g. we assume $\|\xi\| = 1$. Note that there is some $t_0,\ t_0 > 0$, small enough such that $r(\overline{x} + t\xi) > 0$ and $s(\overline{x} + t\xi) > 0$, for all $t \leq t_0$. Let $r(\overline{x} + t_0\xi)s(\overline{x} + t_0\xi)$ be equal to τ_0. We will construct, for any $\tau,\ 0 < \tau \leq \tau_0$, a point $x_\tau = \overline{x} + t_\tau\xi$ such that $x_\tau \in \mathcal{M}^1_\tau$ and $t_\tau = O(\sqrt{\tau})$. Applying the Taylor expansion we find:

$$r(\overline{x} + t\xi)s(\overline{x} + t\xi) = ct^2 + o(t^2),$$

where $c = (\nabla r(\overline{x})^T\xi)(\nabla s(\overline{x})^T\xi) > 0$.

For any τ, $0 < \tau < \tau_0$, (τ_0 chosen small), by the Mean Value Theorem, there exists $t_\tau,\ 0 \leq t_\tau \leq t_0$, such that $r(\overline{x} + t_\tau\xi)s(\overline{x} + t_\tau\xi) = \tau$. As $ct^2_\tau + o(t^2_\tau) = \tau$, we have $t_\tau = \mathcal{O}(\sqrt{\tau})$ and $\overline{x} + t_\tau\xi \in \mathcal{M}^1_\tau$.

(b) W.l.o.g. we assume $r(\overline{x}) = 0, s(\overline{x}) \geq 0$. If $r(\overline{x}) = 0$ and $s(\overline{x}) > 0$ then MFCQ' implies MPCC-LICQ and the result is a consequence of Lemma 4.5.2.
If $r(\overline{x}) = s(\overline{x}) = 0$, the MFCQ' condition implies $\nabla r(\overline{x}) \neq 0$ and $\nabla s(\overline{x}) \neq 0$. We assume that $\nabla_{x_1} r(\overline{x}) \neq 0$ and $\nabla_{x_j} s(\overline{x}) \neq 0$ (j may be equal to 1). Then, we can choose $\varepsilon > 0$ such that for all $x \in B^n_\epsilon(\overline{x})$, $|\nabla_{x_1} r(x)| > \frac{1}{M_r} > 0$ and $\left|\nabla_{x_j} s(x)\right| > \frac{1}{M_s} > 0$ hold for some $M_r,\ M_s > 0$. This means that $\frac{1}{|\nabla_{x_1} r(x)|}$ and $\frac{1}{|\nabla_{x_j} s(x)|}$ are bounded.

Now, we define the following diffeomorphisms, both similar to that used in Lemma 4.5.2:

$$T_r : U \to V_r, \ \ T_r(x) = (r(x), x_2, \ldots, x_n) \, .$$

and

$$T_s : U \to V_s, \ \ T_s(x) = (x_1, \ldots x_{j-1}, s(x), x_{j+1}, \ldots, x_n) \, .$$

T_r (T_s) defines a local C^1-diffeomorphism from U to V_r (V_s respectively), where $U \subset B^n_\epsilon(\overline{x})$ and V_r (V_s respectively) are open balls centered at $\overline{x}$ and $(0, \overline{x}_2, \ldots, \overline{x}_n)$ ($(\overline{x}_1, \ldots \overline{x}_{j-1}, 0, \overline{x}_{j+1}, \ldots, \overline{x}_n)$ respectively).

Of course, for τ small enough, $x_\tau \in U$. We now take $\hat{x}^r_\tau = T_r^{-1}(0, (x_2, \ldots, x_n)_\tau)$. By construction, this point is in U and $r(\hat{x}^r_\tau) = 0$. Moreover,

$$\begin{aligned}
\|\hat{x}^r_\tau - x_\tau\| &= \|T^{-1}(0, (x_2, \ldots, x_n)_\tau) - T^{-1}(r(x_\tau), (x_2, \ldots, x_n)_\tau)\| \\
&= \mathcal{O}\left(\| \begin{pmatrix} \nabla r(x_\tau) \\ 0|I_{n-1} \end{pmatrix}^{-1} \begin{pmatrix} r(x_\tau) \\ 0 \end{pmatrix} \| \right) \\
&= \mathcal{O}\left(|\frac{r(x_\tau)}{\nabla_{x_1} r(x_\tau)}| \right) = \mathcal{O}\left(|r(x_\tau)|\right) .
\end{aligned}$$

Analogously if we define the point $\hat{x}^s_\tau = ((x_1, \ldots x_{j-1})_\tau, 0, (x_{j+1}, \ldots, x_n)_\tau)$, we find $\|\hat{x}^s_\tau - x_\tau\| = \mathcal{O}\left(|s(x_\tau)|\right)$ and $s(\hat{x}^s_\tau) = 0$. But as $r(x_\tau)s(x_\tau) = \tau$, one of the factors should be smaller that or equal to $\sqrt{\tau}$. If $r(x_\tau) \leq \sqrt{\tau}$ we put $\hat{x}_\tau = \hat{x}^r_\tau$ and if $s(x_\tau) \leq \sqrt{\tau}$, $\hat{x}_\tau = \hat{x}^s_\tau$. In both cases it holds that $\|\hat{x}_\tau - x_\tau\| \leq \mathcal{O}\left(\sqrt{\tau}\right)$ and $r(\hat{x}_\tau)s(\hat{x}_\tau) = 0$. However, possibly $r(\hat{x}_\tau) < 0$ or $s(\hat{x}_\tau) < 0$ may occur. So, we take $x^\tau_0 = x_\tau + t^* d$, where $d = \frac{(\hat{x}_\tau - x_\tau)}{\|\hat{x}_\tau - x_\tau\|}$ and

$$t^* = \min\left\{t \mid r(x_\tau + td)s(x_\tau + td) = 0, \ t \geq 0\right\} .$$

Note that now $x^\tau_0 \in \mathcal{M}_{CC}$. As

$$\|x^\tau_0 - x_\tau\| = t^* \leq \|\hat{x}_\tau - x_\tau\| \leq \mathcal{O}(\sqrt{\tau}),$$

we have $\|x^\tau_0 - x_\tau\| = \mathcal{O}(\sqrt{\tau})$. So x^τ_0 has the desired property. □

Remark 4.5.4 *This lemma presents a local convergence result. A global version can also be proven by compactness arguments as in Lemma 4.5.3.*

4.5.3 Convergence results for minimizers

We now discuss the implications of the results so far. In this section $\overline{x} \in \mathcal{M}_{CC}$ denotes a (candidate) minimizer of P_{CC} and $\overline{x}_\tau$ a nearby solution of P_τ, $\tau > 0$. We begin with the following assumptions:

A$_1$. There exists a solution $\overline{x}$ of P_{CC} and a continuous function $\alpha : [0,\infty) \to [0,\infty)$, $\alpha(0) = 0$ such that for any τ, small enough, $\tau > 0$, we can find a point $x_\tau \in \mathcal{M}_\tau$ satisfying

$$\|x_\tau - \overline{x}\| \leq \alpha(\tau).$$

A$_2$. There exists a continuous function $\beta : [0,\infty) \to [0,\infty)$, $\beta(0) = 0$ such that for any τ, small enough, $\tau > 0$, the following holds: we can find a solution $\overline{x}_\tau$ of P_τ and a corresponding point $\hat{x}_\tau \in \mathcal{M}_{CC}$ such that

$$\|\hat{x}_\tau - \overline{x}_\tau\| \leq \beta(\tau).$$

As we shall see, A$_1$ leads to the upper semi-continuity and A$_2$ to the lower semi-continuity of the value function φ_τ.

As $f \in C^1(X)$, due to the compactness of X, we can say that the function f is Lipschitz continuous on X with Lipschitz constant L:

$$|f(\hat{x}) - f(x)| \leq L \cdot \|\hat{x} - x\| \;, \quad \forall \hat{x}, x \in X. \tag{4.5.17}$$

Lemma 4.5.4 *Let the assumptions A_0 (see (4.5.4)), A_1 and A_2 hold and let $f \in C^1(X)$. Then with the constant L in (4.5.17) it follows for all τ, $\tau > 0$:*

$$L\beta(\tau) \leq \varphi_\tau - \varphi \leq L\alpha(\tau).$$

Proof. With the solution $\overline{x}$ of P_{CC} and the points x_τ in A$_1$ we find, using the Lipschitz condition (4.5.17),

$$\varphi_\tau - \varphi \leq f(x_\tau) - f(\overline{x}) \leq L\;\alpha(\tau)$$

and in the same way under A$_2$,

$$\varphi - \varphi_\tau \leq f(\hat{x}_\tau) - f(\overline{x}_\tau) \leq L\;\beta(\tau).$$

□

To obtain results on the rate of convergence for the solutions $\overline{x}_\tau$ of P_τ, we have to assume some growth condition at a solution $\overline{x}$ of P_{CC}.

A$_3$. We assume $\overline{x}$ is a local minimizer of P_{CC} of order ω, *i.e.*, that $\overline{x} \in \mathcal{M}_{CC}$ and for some $\kappa, \varepsilon > 0$ and $\omega = 1$ or $\omega = 2$ the following relation holds:

$$f(x) - f(\overline{x}) \geq \kappa \|x - \overline{x}\|^\omega, \qquad \forall x \in \mathcal{M}_{CC} \cap B^n_\varepsilon(\overline{x}).$$

Sufficient and necessary conditions for this assumption are to be found in Subsection 4.4.2.

Corollary 4.5.1 *Let the assumptions $f \in C^1(X)$, A_0, A_1, A_2 and A_3 hold. Then with the minimizers $\overline{x}_\tau$ of P_τ in A_2 and with some $c > 0$ it follows that*

$$\|\overline{x}_\tau - \overline{x}\| \leq c \cdot \big(\alpha(\tau) + \beta(\tau)\big)^{1/\omega}.$$

Proof. With the points $\overline{x}, \hat{x}_\tau \in \mathcal{M}_{CC}$ and $\overline{x}_\tau, x_\tau \in \mathcal{M}_\tau$ in A_1, A_2 we obtain

$$f(\overline{x}) \leq f(\hat{x}_\tau) \leq f(\overline{x}_\tau) + L\,\beta(\tau) \leq f(x_\tau) + L\,\beta(\tau) \leq f(\overline{x}) + L\,\alpha(\tau) + L\,\beta(\tau)$$

and thus

$$0 \leq f(\hat{x}_\tau) - f(\overline{x}) \leq L\,\alpha(\tau) + L\,\beta(\tau). \tag{4.5.18}$$

Again by taking the point $\hat{x}_\tau \in \mathcal{M}_{CC}$ in A_2 and using the condition A_3, together with (4.5.18), it follows that

$$\begin{aligned}
\|\overline{x}_\tau - \overline{x}\| &\leq \|\overline{x}_\tau - \hat{x}_\tau\| + \|\hat{x}_\tau - \overline{x}\| \\
&\leq \beta(\tau) + \left(\frac{f(\hat{x}_\tau) - f(\overline{x})}{\kappa}\right)^{1/\omega} \\
&\leq \beta(\tau) + \frac{1}{\kappa^{1/\omega}}\left(L\alpha(\tau) + L\beta(\tau)\right)^{1/\omega},
\end{aligned}$$

which proves the statement.

□

Corollary 4.5.2 *Let A_0 and MPCC-LICQ hold for $\mathcal{M}_{CC}$. Let $\overline{x} \in \mathcal{M}_{CC}$ be a global (or local) minimizer of order $\omega = 1$ or $\omega = 2$ of P_{CC}. Then for any τ, small enough, $\tau > 0$, there exists a global (or local) minimizer $\overline{x}_\tau$ of P_τ and for (each of) these minimizers it follows that:*

$$\|\overline{x}_\tau - \overline{x}\| \leq \mathcal{O}(\sqrt{\tau}^{1/\omega})$$

or

$$\|\overline{x}_\tau - \overline{x}\| \leq \mathcal{O}(\tau^{1/\omega}), \qquad \textit{if SC holds at } \overline{x}.$$

Proof. By Lemma 4.5.2, $\mathcal{M}_\tau \neq \emptyset$. Then, due to A_0, the set $\mathcal{M}_\tau$ is compact, so P_τ has a global minimizer. Lemma 4.5.3 implies Conditions A_1 and A_2 with $\alpha(\tau) = \mathcal{O}(\sqrt{\tau})$ and $\beta(\tau) = \mathcal{O}(\sqrt{\tau})$. So, for the global minimizers, the inequality directly follows by combining the results of Lemma 4.5.3 and Corollary 4.5.1.

For a local minimizer $\overline{x}$ we can restrict the feasible sets $\mathcal{M}_{CC}$ and $\mathcal{M}_\tau$ to a closed neighborhood $\overline{B}_\varepsilon^n(\overline{x})$ (closed, to assure the existence of a minimizer $\overline{x}_\tau$). Then the results hold for the (global) minimizer $\overline{x}_\tau$ in $\mathcal{M}_\tau \cap \overline{B}_\varepsilon^n(\overline{x})$. But since $\overline{x}_\tau \to \overline{x}$ for $\tau \to 0$, the points $\overline{x}_\tau$ are also in the open set $B_\varepsilon^n(\overline{x})$, for τ small enough, *i.e.*, $\overline{x}_\tau$ are local minimizers.

□

Let us recall Example 4.5.4 with $k = 2$.

Example 4.5.5 (*cf.* **[47]**)

$$\begin{array}{rrcl} & \min x_1^2 + x_2, & & \\ s.t. & x_1 \cdot x_2 & = & 0, \\ & x_1,\ x_2 & \geq & 0 \end{array}$$

The minimizer $\overline{x} = (0, 0)$ is of order $\omega = 2$ and it is a critical point satisfying $\nabla f(\overline{x}) = 0\nabla r(\overline{x}) + 1\nabla s(\overline{x})$. So the MPCC-SC condition is not fulfilled. Here the minimizers of P_τ read $\overline{x}_\tau = \left(\left(\frac{\tau}{2}\right)^{\frac{1}{3}}, (2\tau^2)^{\frac{1}{3}}\right)$.

The preceding example (see also [47]) shows that at a local minimizer $\overline{x}$ of order two, even under MPCC-LICQ, the convergence rate for $\|\overline{x}_\tau - \overline{x}\|$ can be slower than $\mathcal{O}(\sqrt{\tau})$. Note, however, that this example is not a generic one since the MPCC-SC condition does not hold. We will now show that in the generic case this bad behavior is excluded. More precisely under the conditions MPCC-LICQ, MPCC-SC and MPCC-SOC at a (local) minimizer $\overline{x}$, we prove that the minimizers $\overline{x}_\tau$ of P_τ are (locally) unique and the *optimal* convergence rate $\|\overline{x}_\tau - \overline{x}\| = \mathcal{O}(\sqrt{\tau})$ takes place.

Theorem 4.5.1 *Let $\overline{x}$ be a local minimizer of P_{CC} such that MPCC-LICQ, MPCC-SC and MPCC-SOC hold. Then for any $\tau > 0$, small enough, the local minimizers $\overline{x}_\tau$ of P_τ (near $\overline{x}$) are uniquely determined and satisfy $\|\overline{x}_\tau - \overline{x}\| = \mathcal{O}(\sqrt{\tau})$.*
The same statement holds for the global minimizers $\overline{x}$ and $\overline{x}_\tau$ of P and P_τ, respectively.

Proof. To prove this statement we again consider the problem P_τ in standard form, see the problem given in (4.5.7),

$$\begin{array}{llrcll} P_\tau : & & \min\ f(x) & & & \\ & \text{s.t.} & \hat{h}_i(x) = x_i \cdot x_{l+i} & = & \tau, & i = 1, \ldots, p, \\ & & \hat{h}_i(x) = x_{i+p} \cdot s_{p+i}(x) & = & \tau, & i = 1, \ldots, l-p, \\ & & g_j(x) = x_{l+p+j} & \geq & 0, & j = 1, \ldots, \hat{q}, \\ & & x_i, x_{l+i} & \geq & 0, & i = 1, \ldots, p, \\ & & x_i, s_i(x) & \geq & 0, & i = p+1, \ldots, l. \end{array} \tag{4.5.19}$$

where $\overline{x} = 0$ is the local solution of P_0 with $s_{p+i}(0) = c_i^s > 0,\ i = 1, \ldots, l-p$. Under MPCC-LICQ, the KKT condition for $\overline{x}$ reads:

$$\nabla f(\overline{x}) - \sum_{i=1}^{p} \big(\rho_i e_i + \sigma_i e_{i+l}\big) - \sum_{i=p+1}^{l} \rho_i e_i - \sum_{j=1}^{\hat{q}} \mu_j e_{l+p+j} \ = \ 0. \tag{4.5.20}$$

with multiplier vector (μ, ρ, σ). Here e_i denotes the canonical i^{th}-canonical vector of $\mathbb{R}^n$. Due to the fulfillment of the MPCC-SC and MPCC-LICQ conditions at

$\overline{x}$, we have that $\mu_j, \rho_i, \sigma_i \neq 0$, for all $j = 1, \dots, \hat{q},\ i = 1, \dots, p$. Moreover, as $\overline{x}$ is a local minimizer these multipliers will be strictly positive. So, in (4.5.19) the objective function $f(x)$ has the form

$$f(\overline{x}) = \sum_{i=1}^{p} \big(\rho_i x_i + \sigma_i x_{i+l}\big) + \sum_{i=p+1}^{l} \rho_i x_{p+i} + \sum_{j=1}^{\hat{q}} \mu_j x_{l+p+j} + q(x), \tag{4.5.21}$$

where $|q(x)| = \mathcal{O}(\|x\|^2)$. For convenience we now introduce the abbreviations: $\rho^1 = (\rho_1, \dots, \rho_p)^T$, $\rho^2 = (\rho_{p+1}, \dots, \rho_l)^T$, $\sigma^1 = (\sigma_1, \dots, \sigma_p)^T$, $\mu^1 = (\mu_1, \dots, \mu_{\hat{q}})^T$, $x^1 = (x_1, \dots, x_p)^T$, $x^2 = (x_{l+1}, \dots, x_{l+p})^T$, $x^3 = (x_{p+1}, \dots, x_l)^T$, $x^4 = (x_{l+p+1}, \dots, x_{l+p+\hat{q}})^T$ $x^5 = (x_{l+p+\hat{q}+1}, \dots, x_n)^T$ and write $x = ([x^1]^T, \dots, [x^5]^T)$. In this setting, the tangent space at $\overline{x}$ becomes

$$T_{\overline{x}} = \text{span } \{e_i\ ,\ i = l + p + \hat{q} + 1, \dots, n\},$$

recall that $T_{\overline{x}} = C_{\overline{x}}$, *cf.* (4.3.7). MPCC-SOC, together with the local minimizer condition implies

$$\nabla_x^2 f(x) \text{ is positive definite on } T_{\overline{x}} \text{ or } \nabla_{x^5}^2 f(x) = \nabla_{x^5}^2 q(x) \succ 0, \tag{4.5.22}$$

and problem (4.5.19) reads

$$\begin{array}{lrcl} P_\tau: \ \min & [\rho^1]^T x^1 + [\sigma^1]^T x^2 + [\rho^2]^T x^3 + [\mu^1]^T x^3 + q(x) & & \\ \text{s.t.} \qquad x_i^1 \cdot x_i^2 & = & \tau, & i = 1, \dots, p, \\ x_i^3 \cdot s_{p+i}(x) & = & \tau, & i = 1, \dots, l - p, \\ x_j^4 & \geq & 0, & j = 1, \dots, \hat{q}. \end{array} \tag{4.5.23}$$

Note that, by the condition $\mu^1 > 0$, near $\overline{x}$ all inequalities $x_j^4 \geq 0,\ j = 1, \dots, \hat{q}$ must be active.

As LICQ holds for x near to $\overline{x} = 0$, the minimizers $\overline{x}_\tau$ of P_τ are solutions of the following KKT system of (4.5.23) in the variables $(x, \lambda, \gamma, \nu)$:

$$\begin{pmatrix} \rho^1 + \nabla_{x^1} q \\ \\ \sigma^1 + \nabla_{x^2} q \\ \\ \rho^2 + \nabla_{x^3} q \\ \mu^1 + \nabla_{x^4} q \\ \nabla_{x^5} q \end{pmatrix} = \begin{pmatrix} x_1^2 & & \\ & \ddots & \\ & & x_p^2 \\ x_1^1 & & \\ & \ddots & \\ & & x_p^1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \lambda + \begin{pmatrix} x_1^3 \nabla_{x^1} s_{p+1} & \cdot\cdot & x_{l-p}^3 \nabla_{x^1} s_l \\ \\ x_1^3 \nabla_{x^2} s_{p+1} & \cdot\cdot & x_{l-p}^3 \nabla_{x^2} s_l \\ \\ x_1^3 \nabla_{x^3} s_{p+1} + s_{p+1} e_1 & \cdot\cdot & x_{l-p}^3 \nabla_{x^3} s_l + s_l e_{l-p} \\ x_1^3 \nabla_{x^4} s_{p+1} & \cdot\cdot & x_{l-p}^3 \nabla_{x^4} s_l \\ x_1^3 \nabla_{x^5} s_{p+1} & \cdot\cdot & x_{l-p}^3 \nabla_{x^5} s_l \end{pmatrix} \gamma + \begin{pmatrix} 0 \\ \\ 0 \\ \\ 0 \\ \nu \\ 0 \end{pmatrix}$$

together with the constraints in (4.5.23). In this system the vectors $e_1, \dots, e_{l-p}$ are unit vectors in $\mathbb{R}^{l-p}$. For simplicity we omitted the variable x in the functions $q(x)$ and $s_{p+1}(x), \dots, s_l(x)$.

Now the trick is to eliminate the unknown λ and to simplify (regularize) the equations $x_i^1 \cdot x_i^2 = \tau$ as follows. We define:

$$\begin{aligned}\hat\rho^1(x,\gamma) &:= \rho^1 + \nabla_{x^1} q - [x_1^3 \nabla_{x^1} s_{p+1} \,..\, x_{l-p}^3 \nabla_{x^1} s_l]\gamma \\ \hat\sigma^1(x,\gamma) &:= \sigma^1 + \nabla_{x^2} q - [x_1^3 \nabla_{x^2} s_{p+1} \,..\, x_{l-p}^3 \nabla_{x^2} s_l]\gamma.\end{aligned}$$

As $\rho^1, \sigma^1 > 0$ and $|q(x)| = \mathcal{O}(\|x\|^2)$, near $\overline{x} = 0$ we have $\hat\rho^1 = \rho^1 + \mathcal{O}(x)\gamma + \mathcal{O}(\|x\|^2)$ and $\hat\sigma^1 = \sigma^1 + \mathcal{O}(x)\gamma + \mathcal{O}(\|x\|^2)$. This means that both functions are strictly positive in a neighborhood of $(\overline{x}, \overline{\gamma})$, where $\overline{x} = 0$ and $\overline{\gamma}_i = \frac{\rho_i^2}{s_{p+i}(0)}$, $i = 1, \dots, l-p$. Recall that $s_{p+i}(0) > 0$. So near this point the functions $\sqrt{\frac{\hat\rho^1}{\hat\sigma^1}}$ and $\sqrt{\frac{\hat\sigma^1}{\hat\rho^1}}$ are C^1-functions of (x,γ). From the system we deduce $\hat\rho_i^1 = x_i^2 \lambda_i$, $\hat\sigma_i^1 = x_i^1 \lambda_i$. But since $x_i^1 x_i^2 = \tau$ it follows that $\hat\rho_i^1 \hat\sigma_i^1 = \tau \lambda_i^2$ or $\lambda_i = \sqrt{\frac{\hat\rho_i^1 \hat\sigma_i^1}{\tau}}$. So

$$x_i^1 = \sqrt{\hat\sigma_i^1 / \hat\rho_i^1} \cdot \sqrt{\tau} \quad \text{and} \quad x_i^2 = \sqrt{\hat\rho_i^1 / \hat\sigma_i^1} \cdot \sqrt{\tau}.$$

This means that the system above can be written equivalently as:

$$\begin{array}{ccccc} x_i^1 & - & \sqrt{\hat\sigma_i^1/\hat\rho_i^1} \cdot \sqrt{\tau} & = 0, \\ x_i^2 & - & \sqrt{\hat\rho_i^1/\hat\sigma_i^1} \cdot \sqrt{\tau} & = 0, \\ \left[x_1^3 \nabla_{x^3} s_{p+1} + s_{p+1} e_1 \ \dots \ x_{l-p}^3 \nabla_{x^3} s_l + s_{p+1} e_{l-p}\right]\gamma & - & \rho^2 - \nabla_{x^3} q & = 0, \\ (x_1^3 \nabla_{x^5} s_{p+1} \dots x_{l-p}^3 \nabla_{x^5} s_l)\gamma & - & \nabla_{x^5} q & = 0, \\ x_i^3 s_{p+i} & - & \tau & = 0, \\ x_i^4 & & & = 0, \end{array} \tag{4.5.24}$$

and the system corresponding to the multiplier ν:

$$-(x_1^3 \nabla_{x^5} s_{p+1} \dots x_{l-p}^3 \nabla_{x^5} s_l)\gamma \quad + \quad \mu^1 + \nabla_{x^4} q \ \ = \nu. \tag{4.5.25}$$

The relation (4.5.24) represents a system $F(x,\gamma,\tau) = 0$ of $n+l-p$ equations in $n+l-p+1$ variables (x,γ,τ). The point $(0,\overline{\gamma},0)$, $\overline{\gamma} = (\rho_1^2/s_{p+1}(0), \dots, \rho_{l-p}^2/s_l(0))$, solves system (4.5.24), recall that $s_i(0) > 0$, $i = p+1, \dots, l$. The Jacobian with respect to (x,γ) at this point $(0,\overline{\gamma},0)$ has the form:

$$\begin{array}{cccccc} x^1 & x^2 & x^3 & x^4 & x^5 & \gamma \\ \hline I & 0 & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 & 0 \\ X & X & X & X & X & \begin{bmatrix} s_{p+1} & & \\ & \ddots & \\ & & s_l \end{bmatrix} \\ 0 & 0 & X & 0 & \nabla_{x^5}^2 q & 0 \\ 0 & 0 & \begin{bmatrix} s_{p+1} & & \\ & \ddots & \\ & & s_l \end{bmatrix} & 0 & 0 & 0 \\ 0 & 0 & 0 & I & 0 & 0 \end{array} \tag{4.5.26}$$

where X is some matrix of appropriate dimension. Recall that $\nabla_{x^i} q(0) = 0$. The matrix in (4.5.26) is regular since $s_i(0) > 0,\ i = p+1, \ldots, l$ and $\nabla^2_{x^5} q(0) \succ 0$. Then we can apply the Implicit Function Theorem to the system in (4.5.24). As a consequence, near $\tau = 0$ it yields a unique solution $(x(\tau), \gamma(\tau))$ differentiable in the parameter $\sqrt{\tau}$. This implies $(x(\tau), \gamma(\tau)) = (\overline{x} + \mathcal{O}(\sqrt{\tau}), \overline{\gamma} + \mathcal{O}(\sqrt{\tau}))$. Substituting this solution $(x(\tau), \gamma(\tau))$ into the equation (4.5.25) the variable $\nu(\tau)$ is determined. Since the (local) minimizers $\overline{x}_\tau$ of P_τ (whose existence was shown in Corollary 4.5.2) must solve the systems (4.5.24)-(4.5.25), clearly $\overline{x}_\tau = x(\tau)$ is uniquely determined. The unique multipliers w.r.t. P_τ are $\nu_i(\tau)$ corresponding to $x_i^4 = 0$, $\gamma_i(\tau)$ corresponding to $x_i^3 s_{p+i}(x) = \tau$ and $\sqrt{\frac{\hat{\rho}_i^1 \hat{\sigma}_i^1}{\tau}}$ belonging to the constraint $x_i^1 \cdot x_i^2 = \tau$. This proves the statement for the local minimizers.

If $\overline{x}$ is a global minimizer we can argue as in the second part of the proof of Corollary 4.5.2. Firstly by restricting the minimization to a neighborhood $\overline{B}^n_\delta(\overline{x})$ the result follows as above. The compactness assumption for $\mathcal{M}_\tau$ and the fact that $\overline{x}$ is a global minimizer (of order $\omega = 2$) exclude global minimizers $\overline{x}_\tau$ of P_τ outside $\overline{B}^n_\delta(\overline{x})$.

□

In the next Corollary we indicate that the result of Theorem 4.5.1 is also true for C-stationary points satisfying the conditions MPCC-LICQ, MPCC-SC and MPCC-SOC.

Corollary 4.5.3 *Let $\overline{x}$ be a C-stationary point where MPCC-LICQ, MPCC-SC and MPCC-SOC hold. Then for any $\tau > 0$, small enough, there exists uniquely determined critical points x_τ of P_τ, converging to $\overline{x}$ according to $\|x_\tau - \overline{x}\| = \mathcal{O}(\sqrt{\tau})$.*

Proof. Note that since $\overline{x}$ is a C-stationary point and MPCC-SC holds, it follows that $\rho_i \cdot \sigma_i > 0,\ i \in I_{rs}(\overline{x})$ and $\mu > 0$. The condition (4.5.22) will then be replaced by the fact that $\nabla^2_{x^5} q(x)$ is non-singular. The rest follows as in the proof of Theorem 4.5.1. □

Remark 4.5.5 *Note that the hypotheses used in Theorem 4.5.1 and Corollary 4.5.3 are generic, see Theorem 4.4.1.*

Remark 4.5.6 *The results above guarantee the existence of a sequence of local minimizers (stationary points) of P_τ converging to the non-degenerate local minimizers (stationary points) of P_{CC}. For computing this sequence numerically, the constraints $r_i(x)s_i(x) = \tau$, and $r_i(x),\ s_i(x) \geq 0$ can be modeled equivalently with a unique equality $\psi_\tau(r_i(x), s_i(x)) = 0$ where $\psi_\tau : \mathbb{R}^2 \to \mathbb{R}$ is a so-called parameterized NCP-function satisfying $\psi_\tau(x, y) = 0 \Leftrightarrow x, y \geq 0\ xy = \tau$, see e.g. Chen and Mangasarian [9].*

Let us note that from the results of this part we also can deduce the convergence results for the relaxation $P_\tau^{\leq}$, see (4.5.2) and [47], under the stronger MPCC-LICQ condition. Let us reconsider this relaxation $P_\tau^{\leq}$. Suppose we have given a local solution $\overline{x}$ of P_{CC} such that MPCC-LICQ holds and, with corresponding Lagrange multipliers, MPCC-SC, MPCC-SOC are satisfied, *i.e.*, by Theorem 4.4.4, $\overline{x}$ is a minimizer of order $\omega = 2$. In view of Corollary 4.4.2, $\overline{x}$ is also a solution of the relaxed problem $P_R(\overline{x})$ in (4.3.3) and by using MPCC-SC it follows that for any $\tau > 0$, small enough, for the solutions $\hat{x}_\tau$ (near $\overline{x}$) of $\mathrm{P}_\tau^{\leq}$ (see Remark 4.5.1), the conditions $r_i(x)\ s_i(x) \leq \tau$, $i \in I_{rs}(\overline{x})$, are not active but that

$$r_i(\hat{x}_\tau) = s_i(\hat{x}_\tau) = 0, \quad \forall i \in I_{rs}(\overline{x}), \tag{4.5.27}$$

holds. So, to analyze the behavior of the solution $\hat{x}_\tau$ the whole analysis can be done under the condition (4.5.27), *i.e.*, we are in the situation as for the case that the *SC* condition holds. Consequently, instead of the convergence $\mathcal{O}(\sqrt{\tau})$ (*cf., e.g.* Lemma 4.5.2), we obtain a rate $\mathcal{O}(\tau)$ and in the same way the analysis in Section 4.5 simplifies resulting in a convergence behavior $\|\hat{x}_\tau - \overline{x}\| = \mathcal{O}(\tau)$.

Remark 4.5.7 *We wish to emphasize that the convergence results of this section can be generalized in a straightforward way to problems P_m, $m \geq 3$, containing constraints of the product form:*

$$r_1(x) r_2(x) \cdots r_m(x) = 0, \quad r_1(x), r_2(x), \ldots, r_m(x) \geq 0.$$

Here at a solution $\overline{x}$ of P_m where all constraints r_i are active, i.e.,

$$r_1(\overline{x}) = r_2(\overline{x}) = \ldots = r_m(\overline{x}) = 0,$$

a perturbation $r_1(x) r_2(x) \cdots r_m(x) = \tau$ will lead to a convergence rate

$$\|\overline{x}_\tau - \overline{x}\| \approx \mathcal{O}(\tau^{1/m}),$$

for the solutions $\overline{x}_\tau$ of the perturbed problem. Also all other results can be extended in a straightforward way to this generalization.

4.6 JJT-regularity results for P_τ

In this section we are going to consider P_τ (see (4.5.1)) as a 1-parametric problem. Using the JJT-regularity, we will proof that for a generic P_{CC} (see (4.1.2)) with $(f, r_1, \ldots, r_l, s_1, \ldots, s_l, g_1, \ldots, g_q) \in [C^3]_n^{1+l+l+q}$, the one-parametric problem P_τ, is regular for $\tau \in (0, 1]$ in the sense of Definition 2.4.7.
We firstly deal with the LICQ condition. For this part we can weaken the differentiability hypothesis and assume $(r_1, \ldots, r_l, s_1, \ldots, s_l, g_1, \ldots, g_q) \in [C^2]_n^{l+l+q}$.

Note that, for all $x \in \mathcal{M}_\tau,\ \tau > 0$, the values of the functions $r_i(x), s_i(x)$ are strictly positive. So the active index set is $J_0(x,\tau) = \{j \mid g_j(x) = 0\}$. By definition LICQ fails at a point $x \in \mathcal{M}_\tau,\ 0 < \tau$, if

$$\sum_{i=1}^{l} [s_i(x)\nabla r_i(x) + r_i(x)\nabla s_i(x)]\lambda_i + \sum_{j \in J_0(x,\tau)} \mu_j \nabla g_j(x) = 0, \quad \text{for some } (\lambda,\mu) \neq 0. \tag{4.6.1}$$

We define the set $\mathcal{M}^0 = \{(x,\tau) \mid 0 < \tau \leq 1,\ x \in \mathcal{M}_\tau$ such that (4.6.1) holds$\}$, and $\mathcal{M}^1 = \{(x,\tau) \in \mathcal{M}^0 \mid$ (4.6.1) holds with $\mu_j \neq 0,\ \forall j \in J_0(x,\tau)\}$. Then the following is true.

Proposition 4.6.1 *Let $(r_1,\dots,r_l,s_1,\dots,s_l,g_1,\dots,g_q) \in [C^2]_n^{l+l+q}$ be given. Then for almost every linear perturbation of $r_i, g_j,\ i = 1,\dots,l,\ j = 1,\dots,q$, the set $\mathcal{M}^0$ is a discrete set and $\mathcal{M}^0 = \mathcal{M}^1$.*

The set of functions $(r_1,\dots,r_l,s_1,\dots,s_l,g_1,\dots,g_q) \in [C_S^2]_n^{l+l+q}$, see Definition 2.3.1, such that $\mathcal{M}^0$ is a discrete set and $\mathcal{M}^0 = \mathcal{M}^1$, is generic in $[C_S^2]_n^{l+l+q}$.

Proof. As $g_j,\ j = 1,\dots,q$, are non-parametric functions, not depending on τ, from the standard genericity results we can conclude that for almost every linear perturbation of g_j, the vectors $\nabla g_j(x),\ \ j \in J_0(x,\tau)$ are linearly independent at all points $(x,\tau),\ x \in \mathcal{M}_\tau,\ 0 < \tau \leq 1$. So, in the rest of the proof, we will assume that the latter property holds. Now we will prove the statements of the proposition for $\tau \in [\frac{1}{k}, 1]$, where $k \in \mathbb{N}$ is fixed.
If LICQ is violated for P_τ then, by normalizing and possibly changing indices, the following system is solvable:

$$\begin{array}{rcl}\nabla r_1(x)s_1(x) + \nabla s_1(x)r_1(x) + \sum\limits_{j \in J_0(x,\tau)} \nabla g_j(x)\mu_j & & \\ + \sum\limits_{i=2}^{l} \lambda_i(\nabla r_i s_i + \nabla s_i r_i)(x) & = & 0, \\ r_i(x)s_i(x) & = & \tau, \quad i = 1,\dots,l, \\ g_j(x) & = & 0, \quad j \in J_0(x,\tau), \\ \mu_j & = & 0, \quad j \in J_B \subset J_0(x,\tau). \end{array} \tag{4.6.2}$$

For the perturbation $(r_i + b_i^T x + c_i),\ i = 1,\dots,l$, the Jacobian of the system with respect to (x,τ,λ,μ), $\lambda = (\lambda_2,\dots,\lambda_l))$, and the parameters b_1, c is of the form:

$$
\begin{array}{ccccc}
\partial x & \partial \tau & \partial \lambda, \mu & \partial b_1 & \partial c \\
\hline
\bigotimes_{xx} & 0 & \bigotimes & \begin{pmatrix} s_1 & 0 & \dots \\ 0 & \ddots & 0 \\ 0 & \dots & s_1 \end{pmatrix} + \nabla s_1 \, [x_1, \dots, x_n] & \nabla s_1 \mid \bigotimes \\
\bigotimes & \begin{pmatrix} -1 \\ \vdots \\ -1 \end{pmatrix} & 0 & s_1 \begin{bmatrix} x_1, \dots, x_n \\ 0 \end{bmatrix} & \begin{pmatrix} s_1 & & \\ & \ddots & \\ & & s_l \end{pmatrix} \\
\nabla^T g_{J_0(x,\tau)} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 \mid I_{|J_B|} & 0 & 0
\end{array}
\tag{4.6.3}
$$

The matrix : $\begin{pmatrix} diag(s_1 \dots s_1)_n + \nabla s_1 \, [x_1, \dots, x_n] & \nabla s_1 \\ s_1(x)[x_1, \dots, x_n] & s_1(x) \end{pmatrix}$ is non-singular, since $s_1(x) > 0$ for $\tau \in [\frac{1}{k}, 1]$. Then the first n+1 rows in (4.6.3) are linearly independent. So the matrix (4.6.3) has full row rank if and only if the following matrix has also full row rank:

$$
\begin{array}{ccccc}
\partial x & \partial \tau & \partial \lambda, \mu & \partial b_1 & \partial c \\
\hline
\bigotimes & \begin{pmatrix} -1 \\ \vdots \\ -1 \end{pmatrix} & 0 & 0 & \begin{pmatrix} s_2 & & \\ & \ddots & \\ & & s_l \end{pmatrix} \\
\nabla_x^T g_{J_0(x,\tau)} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 \mid I_{|J_B|} & 0 & 0
\end{array}
$$

But the latter is a direct consequence of the facts that $s_i(x) > 0, \ \forall x \in \mathcal{M}_\tau, \ \tau > 0$, and the linear independence of $\nabla_x g_j(x), \ j \in J_0(x, \tau)$.
Applying the Parameterized Sard Lemma (*cf.* Lemma 2.3.1) it follows that for almost every linear perturbation of $r_i, \ i = 1, \dots, l$, the sub-matrix of (4.6.3) corresponding to $\partial(x, \tau, \lambda, \mu)$ has full row rank at all solutions (x, τ, λ, μ) of (4.6.2), $\tau \in [\frac{1}{k}, 1]$. So the number of rows, $n + l + |J_0(x,t)| + |J_B|$, must be smaller than or equal to $n + 1 + l - 1 + |J_0(x, \tau)|$, the number of unknowns (x, τ, λ, μ). This implies $|J_B| \leq 0$, *i.e.*, $\mu_j \neq 0$ for all $j \in J_0(x, \tau)$. Consequently, the full row rank condition for (4.6.3) implies that for almost all linear perturbations, the following matrix is non-singular at all solutions (x, τ, λ, μ) of (4.6.2), $\tau \in [\frac{1}{k}, 1]$:

$$
\begin{array}{ccc}
\partial x & \partial \tau & \partial \lambda, \mu \\
\hline
\bigotimes_{xx} & 0 & \nabla_x(r_2 s_2) \dots \nabla_x(r_l s_l) \nabla_x^T g_{J_0(x,\tau)} \\
\nabla_x(r_1 s_1)^T & -1 & 0 \\
\vdots & -1 & 0 \\
\nabla_x(r_l s_l)^T & -1 & 0 \\
\nabla_x g_{J_0(x,\tau)} & 0 & 0
\end{array}
\tag{4.6.4}
$$

Here $\otimes_{xx} = \nabla_x^2\left(r_1(x)s_1(x)\right) + \sum\limits_{i=2}^{l}\lambda_i\nabla_x^2\left(r_i(x)s_i(x)\right) + \sum\limits_{j\in J_0(x,\tau)}\mu_j\nabla_x^2 g_j(x)$.

So, as the number of equations and variables of the system (4.6.2) (now without the equations $\mu_j \neq 0$) are the same, it describes a *0-dimensional* manifold, *i.e.*, the solution set is discrete.

The perturbation result for $\tau \in (0,1]$ follows by taking $k = 2, 3, \ldots$, and by intersecting the corresponding sets of parameters.

Now we will prove the genericity statement for $\tau \in [\frac{1}{k}, 1]$, where $k \in \mathbb{N}$ is fixed.
The dense part follows directly from the above perturbation argument as in the proof of Theorem 2.4.1, *cf.* Theorem 6.21 of [21].

For the open part, we assume that for the functions $(r^*(x), s^*(x), g^*(x)) \in [C_S^2]_n^{l+l+q}$ the property of the proposition holds for all $\tau \in [\frac{1}{k}, 1]$. We have to find a neighborhood $V_{\hat{\epsilon}(x)}$ of $r^*(x), s^*(x), g^*(x)$, given by a continuous positive function $\hat{\epsilon}(x)$, where this property remains stable. (For simplicity we often skip the variable x).
By Theorem 2.4.1, for the parametric problem, there is a neighborhood defined by a function $\epsilon_0(x,\tau)$ such that for $(\hat{h}_0, r_0, s_0, g_0) \in V_{\epsilon_0(x,\tau)}(r^*s^* - \tau, r^*, s^*, g^*)$, the corresponding feasible set is regular, *e.g.*, in the feasible set

$$\left\{ (x,\tau) \in \mathbb{R}^n \times \mathbb{R} \;\middle|\; \begin{array}{lcll} g_{0j}(x,\tau) & \geq & 0, & j = 1,\ldots,q, \\ r_{0i}(x,\tau) & \geq & 0, & i = 1,\ldots,l, \\ s_{0i}(x,\tau) & \geq & 0, & i = 1,\ldots,l, \\ \hat{h}_{0i}(x,\tau) & = & 0, & i = 1,\ldots,l \end{array} \right\}$$

LICQ fails at most in a discrete set. Moreover, at the points (x,τ) where LICQ fails we have:

for any non-trivial linear combination of the active constraints, the coefficients corresponding to the active inequality constraints are no-zero. (4.6.5)

In case of g.c. points of type 4 or 5, this property corresponds to the fulfillment of condition (4*b*) and (5*b*) in Definition 2.4.5 and Definition 2.4.6 respectively.
As in the proof of Proposition 3.4.1, we define the continuous and positive function $\epsilon(x) = \min\limits_{\tau\in[\frac{1}{k},1]} \epsilon_0(x,\tau)$.

Now let the functions $(\hat{h}_i, r_i, s_i,\ g_j) : \mathbb{R}^n \to \mathbb{R},\ i = 1,\ldots,l,\ j = 1,\ldots,q$, be such that $|r_i(x) - r_i^*(x)| < \epsilon(x)$, $|s_i(x) - s_i^*(x)| < \epsilon(x)$, $|\hat{h}_i(x) - r_i^*(x)s_i^*(x)| < \epsilon(x)$, $|g_j(x) - g_j^*(x)| < \epsilon(x)$ and the partial derivatives satisfy an analogous inequality. It is clear that $(\hat{h} - \tau, r, s, g) \in V_{\epsilon_0(x,\tau)}(r^*s^* - \tau, r^*, s^*, g^*)$. So, by assumption, for all $(\hat{h}-\tau, r, s, g) \in V_{\epsilon(x)}(r^*s^*, r^*, s^*, g^*)$, at the associated parametric set:

$$\left\{ (x,\tau) \in \mathbb{R}^n \times \mathbb{R} \;\middle|\; \begin{array}{rcll} g_j(x) & \geq & 0, & j=1,\ldots,q, \\ r_i(x) & \geq & 0, & i=1,\ldots,l, \\ s_i(x) & \geq & 0, & i=1,\ldots,l, \\ \hat{h}_i(x)-\tau & = & 0, & i=1,\ldots,l \end{array} \right\}$$

LICQ fails in at most a discrete set and condition (4.6.5) holds.

Now we will shrink the function $\epsilon(x)$ to obtain a neighborhood $V_{\hat{\epsilon}(x)}$ of (r^*, s^*, g) such that $(r, s, g) \in V_{\hat{\epsilon}(x)}$ means $(rs, r, s, g) \in V_{\epsilon(x)}(r^*s^*, r^*, s^*, g^*)$. Let us consider the C^2_S-neighborhood defined by the function:

$$\hat{\epsilon}(x) = \min\left\{\epsilon(x), \frac{\epsilon(x)}{c_0(x)}\right\},$$

where

$$\begin{aligned} c_0(x) \quad = \quad & 4\epsilon(x) + \sum_{i=1}^{l}\left[|r_i^*(x)| + |s_i^*(x)| + \sum_{j=1}^{n}\left[\left|\frac{\partial r_i^*(x)}{\partial x_j}\right| + \left|\frac{\partial s_i^*(x)}{\partial x_j}\right|\right]\right] \\ & + \sum_{i=1}^{l}\sum_{j=1}^{n}\sum_{k=j}^{n}\left[\left|\frac{\partial^2 r_i^*(x)}{\partial x_j \partial x_k}\right| + \left|\frac{\partial^2 s_i^*(x)}{\partial x_j \partial x_k}\right|\right]. \end{aligned}$$

As $\hat{\epsilon}(x)$ is the minimum of two continuous and positive functions, $V_{\hat{\varepsilon}(x)}$ defines a strong neighborhood.

If $r_i \in V_{\hat{\epsilon}(x)}(r_i^*)$, then $|r_i(x) - r_i^*(x)| < \hat{\epsilon}(x) \leq \epsilon(x)$. A similar relation holds for the functions s_i and g_i and their first and the second derivatives. Then, $r_i(x) \in V_{\epsilon(x)}(r_i^*),\ s_i(x) \in V_{\epsilon(x)}(s_i^*),\ g_j(x) \in V_{\epsilon(x)}(g_j^*)$ for $i=1,\ldots,l,\ j=1,\ldots,q$. We now prove that if $(r_i, s_i) \in V_{\hat{\epsilon}(x)}(r_i^*) \times V_{\hat{\epsilon}(x)}(s_i^*)$, then $r_i(x)s_i(x) \in V_{\epsilon(x)}(r_i^*s_i^*)$. So, for all $g_j \in V_{\epsilon(x)}(g_j^*),\ r_i \in V_{\epsilon(x)}(r_i^*),\ s_i \in V_{\epsilon(x)}(s_i^*),\ i=1,\ldots,l,\ j=1,\ldots,q$, the corresponding set $\mathcal{M}_\tau$ will satisfy the properties of the proposition. First we have:

$$|s_i(x)r_i(x) - r_i^*(x)s_i^*(x)| \leq |s_i(x)||r_i(x) - r_i^*(x)| + |r_i^*(x)||s_i(x) - s_i^*(x)|$$

$$< \hat{\epsilon}(x)\left[|s_i(x)| + |r_i^*(x)|\right].$$

But as $|s_i(x) - s_i^*(x)| < \hat{\epsilon}(x) \leq \epsilon(x)$, so $|s_i(x)| \leq \epsilon(x) + |s_i^*(x)|$ and

$$|s_i(x)r_i(x) - r_i^*(x)s_i^*(x)| < \hat{\epsilon}(x)\left[\epsilon(x) + |s_i^*(x)| + |r_i^*(x)|\right]$$

$$\leq \frac{\epsilon(x)\left[\epsilon(x) + |s_i^*(x)| + |r_i^*(x)|\right]}{\epsilon(x) + |r_i^*(x)| + |s_i^*(x)|} \leq \epsilon(x).$$

For the first order derivatives $\partial = \frac{\partial}{\partial x_j}$, we find:

$$\begin{array}{rcl} |\partial(r_i s_i) - \partial(r_i^* s_i^*)| & = & |s_i \partial r_i + r_i \partial s_i - s_i^* \partial r_i^* - r_i^* \partial s_i^*| \\ & \leq & |s_i \partial r_i - s_i^* \partial r_i^*| + |r_i \partial s_i - r_i^* \partial s_i^*| \\ & \leq & |s_i||\partial r_i - \partial r_i^*| + |\partial r_i^*||s_i - s_i^*| + \\ & & +|r_i||\partial s_i - r_i^* \partial s_i^*| + |\partial s_i^*||r_i - r_i^*| \\ & \leq & \hat{\epsilon}(x)[\epsilon(x) + s^*] + |\partial r_i^*|\hat{\epsilon}(x) + \hat{\epsilon}(x)[r_i^* + \epsilon(x)] + |\partial s|\hat{\epsilon}(x) \\ & = & \hat{\epsilon}(x)[|s_i^*| + |r_i^*| + |\partial r_i^*| + |\partial s_i^*| + 2\epsilon] \leq \hat{\epsilon}(x) c_0(x). \end{array}$$

Then, using $\hat{\epsilon}(x) \leq \frac{\epsilon(x)}{c_0(x)}$, we also obtain:

$$|\partial(r_i s_i) - \partial(r_i^* s_i^*)| \leq \epsilon(x).$$

For the second derivatives we can similarly prove the inequality

$$|\partial^2(r_i s_i) - \partial^2(r_i^* s_i^*)| \leq \epsilon(x).$$

Then for all functions $(r, s, g) \in V_{\hat{\epsilon}(x)}$, it holds that $(rs, r, s, g) \in V_{\epsilon(x)}(r^* s^*, r^*, s^*, g^*)$ which proves the open part.

Together with the density we have shown the genericity result for $\tau \in [\frac{1}{k}, 1]$. Now taking the intersection of the corresponding open and dense sets in $[C_S^2]_n^{l+l+q}$ for $k = 2, 3, \ldots$, we see that generically $\mathcal{M}^0$ is a discrete set and $\mathcal{M}^0 = \mathcal{M}^1$.

$\square$

Remark 4.6.1 *Note that, as a consequence of the Parameterized Sard Lemma, it also holds that for almost every linear perturbation of (r, g) the matrix (4.6.4) is non-singular for all solutions (x, τ, λ, μ) of system (4.6.1)*

Remark 4.6.2 *A similar result can be proven if instead of C_S^2, we consider C_S^3. The only difference is that in the open part of the proof the function $\hat{\epsilon}(x)$ will be smaller in order to assure $|\frac{\partial^3}{\partial x_{k_1}, x_{k_2}, x_{k_3}} r_i s_i(x) - \frac{\partial^3}{\partial x_{k_1}, x_{k_2}, x_{k_3}} r_i^* s_i^*(x)| < \hat{\epsilon}(x)$ for all $i = 1, \ldots, l,\ k_1, k_2, k_3 = 1, \ldots, n$.*

Proposition 4.6.2 *Let $(f, r_1, \ldots, r_l, s_1, \ldots, s_l, g_1, \ldots, g_q) \in [C_S^3]_n^{1+l+l+q}$, be fixed. Then for almost every perturbation, linear in $(r_1, \ldots, r_l, s_1, \ldots, s_l, g_1, \ldots, g_q)$ and quadratic in f, the corresponding parametric problem P_τ, $\tau \in (0, 1]$ is regular on $(0, 1]$, see Definition 2.4.7.*
In particular, the set of functions $(f, r_1, \ldots, r_l, s_1, \ldots, s_l, g_1, \ldots, g_q) \in \mathcal{F}|_{(0,1]}$, is generic in $[C_S^3]_n^{1+l+l+q}$.

Proof. By Proposition 4.6.1 and Remark 4.6.1, for almost all linear perturbation of $(r_1, \ldots, r_l, g_1, \ldots, g_q)$, LICQ is not fulfilled only in a discrete set (*i.e.*, $\mathcal{M}^0$ is a discrete set) $\mathcal{M}^0 = \mathcal{M}^1$ and the matrix (4.6.4) is non-singular at all solutions (x, τ, λ, μ) of system (4.6.1) with $(x, \tau) \in \mathcal{M}^0$. Let us fix an arbitrary linear perturbation $(C_r, d_r, C_g, d_g) \in \mathbb{R}^{ln+l+qn+q}$ of the constraints satisfying the previous statements and denote the associated (perturbed) functions by $(\hat{r}, \hat{s}, \hat{g})$. Now the corresponding feasible set $\mathcal{M}_\tau$ is fixed and we again consider the two cases: the generalized critical points (x, τ) where the LICQ condition holds and where it fails.

We now can show that, for almost every quadratic perturbation $\hat{f}$ of $f(x)$, given by $(A, b) \in \mathbb{R}^{\frac{n(n+1)}{2}+n}$, the generalized critical points of P_τ, $\tau \in [\frac{1}{k}, 1]$ (corresponding to $(\hat{f}, \hat{r}, \hat{s}, \hat{g})$) where LICQ holds are of type 1, 2 or 3. This can be done following the lines of the proof of Theorem 2.4.1 (see Theorem 6.18 of [21]).

Now we consider the case where LICQ fails. Let $(\overline{x}, \overline{\tau})$ be an arbitrarily fixed point where LICQ fails, *i.e.*, $(\overline{x}, \overline{\tau}) \in \mathcal{M}^0$. Note that such a point is a g.c. point. As can be derived using the proof of Theorem 2.4.1 (see Theorem 6.18, [21]) $(\overline{x}, \overline{\tau})$ is a g.c. point of type 4 or 5 of the perturbed problem, for almost all quadratic perturbations $\hat{f}$ of $f(x)$ given by $(A, b) \in \mathbb{R}^{\frac{n(n+1)}{2}+n}$. As $\mathcal{M}^0$ has only a discrete number of elements, if we intersect for all $(\overline{x}, \overline{\tau}) \in \mathcal{M}^0$ the sets of parameters (A, b) such that, for the corresponding perturbed problem, $(\overline{x}, \overline{\tau})$ is a g.c. point of type 4 or 5, we obtain that, for almost all quadratic perturbations of $f(x)$, all points where LICQ fails are g.c. points of type 4 or 5. If we intersect this set of parameters (A, b) with the set of (A, b) such that the g.c. points satisfying LICQ are of type 1, 2 or 3, we have proven that for almost all arbitrarily fixed (C_r, d_r, C_g, d_g), for almost every (A, b) the g.c. points of the corresponding perturbed problem P_τ are of type 1, 2, 3, 4 or 5. Now the first statement of the Proposition is a direct consequence of the Fubini theorem in the set $\mathbb{R}^{ln+l+qn+q} \times \mathbb{R}^{\frac{n(n+1)}{2}+n}$.

For the genericity result, we have to show that the set of functions $(f, r_1, \ldots, r_l, s_1, \ldots, s_l, g_1, \ldots, g_q) \in [C_S^3]_n^{1+l+l+q}$ such that the corresponding problem P_τ, is JJT-regular on $[\frac{1}{k}, 1]$, is open and dense in $[C_S^3]_n^{1+l+l+q}$.
The open part is proved in the same way as in the proof of Proposition 4.6.1 and the density is obtained as in the classical case by using partitions of unity. $\square$

Remark 4.6.3 *By using ideas similar to those of the proof of Theorem 4.5.1 it can be seen that if P_{CC} is regular in the MPCC-sense then, for all (x_τ, τ) satisfying $(x_\tau, \tau) \in \Sigma_{gc}$ and $x_\tau \to \overline{x} \in \mathcal{M}_{CC}$, when $\tau \to 0^+$, it follows that for $\tau \ll 1$, (x_τ, τ) is a non-degenerate critical point.*

The previous results enable us to apply a pathfollowing algorithm for solving P_τ, $\tau \to 0$, because, at least locally, a program like PAFO, by Gollmer, Kausmann, Nowack, Wendler and Bacallao [18], will be able to perform a continuation proces on the set of g.c. points around points of type 1 and to handle the singularities in the generic case for $\tau \in (0, 1]$.

4.7 Parametric problem

In this section, we consider one-parametric Mathematical Programs with Complementarity Constraints

$$\begin{array}{lrl} P_{CC}(t): & & \min f(x,t) \\ & \text{s.t.} & x \in \mathcal{M}_{CC}(t) \end{array} \tag{4.7.1}$$

$$\mathcal{M}_{CC}(t) = \left\{ x \in \mathbb{R}^n \;\middle|\; \begin{array}{rcl} h_k(x,t) & = & 0, \quad k = 1, \ldots, q_0, \\ g_j(x,t) & \geq & 0, \quad j = 1, \ldots, q, \\ r_i(x,t) & \geq & 0, \quad i = 1, \ldots, l, \\ s_i(x,t) & \geq & 0, \quad i = 1, \ldots, l, \\ r_i(x,t)s_i(x,t) & = & 0, \quad i = 1, \ldots, l. \end{array} \right\}$$

$t \in \mathbb{R}$. For simplicity we again use the abbreviations $h = (h_1, h_2, \ldots, h_{q_0})$, $g = (g_1, g_2, \ldots, g_q)$, $r = (r_1, r_2, \ldots, r_l)$ and $s = (s_1, s_2, \ldots, s_l)$ and assume at least $(f, h, g, r, s) \in [C^2]^{1+q_0+q+l+l}_{n+1}$.

The section has two parts. We begin with a brief study of the feasibility problem *find* $x \in \mathcal{M}_{CC}(t)$, $t \in [0, 1]$, when $n = l, q = q_0 = 0$. The non-parametric case is also considered.

In the second subsection, the local behavior of the solution set of general problems $P_{CC}(t)$ is analyzed under generical assumptions. We also describe the singularities that may appear in the generic case.

4.7.1 Structure for the special case $n = l, q = q_0 = 0$

Let us consider the problem:

$$\text{for each } t, \text{ find } x \text{ such that } x \in \mathcal{M}_{CC}(t) \tag{4.7.2}$$

where

$$\mathcal{M}_{CC}(t) = \left\{ x \in \mathbb{R}^n \;\middle|\; \begin{array}{rcl} r_i(x,t) & \geq & 0, \quad i = 1, \ldots, n, \\ s_i(x,t) & \geq & 0, \quad i = 1, \ldots, n, \\ r_i(x,t)s_i(x,t) & = & 0, \quad i = 1, \ldots, n. \end{array} \right\}$$

A particular case is the so-called 1-parametric nonlinear complementarity problems (NLCP) (for the non-parametric case see (3.4.4)):

$$\text{for each } t \text{ find } x \in \mathbb{R}^n \text{ such that } \left\{ \begin{array}{rcl} x & \geq & 0, \\ \Phi(x,t) & \geq & 0, \\ \Phi(x,t)^T x & = & 0, \end{array} \right. \tag{4.7.3}$$

where $\Phi \in [C^3]^n_{n+1}$. Problem (4.7.3) can also be seen as the 1-parametric VI problem $VIP(t; \Phi(x,t), \mathbb{R}^n_+)$, see (3.2.1).

For the present case, $n = l$, $q = q_0 = 0$, we will prove that, generically, the set $\{(x,t) \mid x \in \mathcal{M}_{CC}(t)\}$ may locally be described by a curve $(x(t),t)$, or may bifurcate into two branches or may have quadratic turning points. In the non-parametric case we show that generically $\mathcal{M}_{CC}$ is a discrete set.

As $r_i(x,t)s_i(x,t) = 0$ must hold at feasible points, at least one of the functions $r_i(x,t)$ or $s_i(x,t)$ is equal to 0. So, to simplify the presentation we will often assume, w.l.o.g., that at a feasible point $(\overline{x},\overline{t})$, the condition $r(\overline{x},\overline{t}) = 0, s(\overline{x},\overline{t}) \geq 0$ holds. As usual, we define the active index set as:

$$\begin{array}{rcl} I_r(x,t) & = & \{i \mid r_i(x,t) = 0,\ s_i(x,t) > 0\}\,, \\ I_s(x,t) & = & \{i \mid s_i(x,t) = 0,\ r_i(x,t) > 0\}\,, \\ I_{rs}(x,t) & = & \{i \mid r_i(x,t) = s_i(x,t) = 0\}\,. \end{array} \tag{4.7.4}$$

The non-parametric case.

In this part we study the non-parametric version of problem (4.7.2):

$$\text{Find } x \in \mathcal{S}_0, \tag{4.7.5}$$

$$\mathcal{S}_0 = \left\{ x \in \mathbb{R}^n \;\middle|\; \begin{array}{rcl} r_i(x)s_i(x) & = & 0,\quad i = 1,\ldots,n, \\ r_i(x) & \geq & 0,\quad i = 1,\ldots,n, \\ s_i(x) & \geq & 0,\quad i = 1,\ldots,n. \end{array} \right\}$$

The following genericity result holds for $(r,s) \in [C^\infty]^{n+n}_n$.

Proposition 4.7.1 *Generically, with respect to the $[C^2_S]$-topology on $[C^\infty]^{n+n}_n$, the functions (r,s) define a set $\mathcal{S}_0$ which is a 0-dimensional manifold, i.e., it is a discrete set. Besides, generically, at each $x \in \mathcal{S}_0$ it follows $I_{rs}(x) = \varnothing$ and $[\nabla r_{I_r(x)}(x), \nabla s_{I_s(x)}(x)]$ is a regular matrix.*

Proof. By definition, at all feasible points, at least n constraints are equal to zero. W.l.o.g., and possibly after changing the roles of r_i and s_i, we assume $r(x) = 0$ and $s_1(x) = s_2(x) = \ldots s_{l_0}(x) = 0,\ 0 \leq l_0 \leq n$. The result now is a direct consequence of the Jet Transversality Theorem, Theorem 2.3.1, applied to the functions (r,s) with respect to the jet-manifold $\{(x,z,u) \in \mathbb{R}^{n+n+n} \mid r(x) - z = 0,\ s(x) - u = 0\}$ and the manifold $\mathbb{R}^n \times 0_{n+l_0} \times \mathbb{R}^{n-l_0}$.

$\square$

The one-parametric case.

Now we turn back to problem (4.7.2) in order to study the structure of the set

$$\mathcal{M} = \left\{ (x,t) \in \mathbb{R}^n \times \mathbb{R} \;\middle|\; \begin{array}{rcl} r_i(x,t)s_i(x,t) & = & 0, \quad i = 1,\ldots,n, \\ r_i(x,t) & \geq & 0, \quad i = 1,\ldots,n, \\ s_i(x,t) & \geq & 0, \quad i = 1,\ldots,n \end{array} \right\}$$

We will always assume $(r,s) \in [C^\infty]^{n+n}_{n+1}$. Firstly, we will discuss the different generic types of solution points (x,t) that may appear using first order information.

Proposition 4.7.2 *Generically, with respect to the $[C^2_S]$-topology on $[C^\infty]^{n+n}_{n+1}$, the corresponding set $\mathcal{M}$ is such that if $(\overline{x},\overline{t}) \in \mathcal{M}$, then one of the following three conditions is fulfilled:*

(a) $I_{rs}(\overline{x},\overline{t}) = \varnothing$ and $\nabla_x[r_{I_r(\overline{x},\overline{t})}(\overline{x},\overline{t}), s_{I_s(\overline{x},\overline{t})}(\overline{x},\overline{t})]$ is non-singular. Then locally around $(\overline{x},\overline{t})$, $\mathcal{M}$ is defined by a curve $(x(t),t)$, $x(\overline{t}) = \overline{x}$, $t \in [\overline{t}-\varepsilon, \overline{t}+\varepsilon]$, for some ε, $\varepsilon > 0$.

(b) $I_{rs}(\overline{x},\overline{t}) = \varnothing$, $\nabla_x[r_{I_r(\overline{x},\overline{t})}(\overline{x},\overline{t}), s_{I_s(\overline{x},\overline{t})}(\overline{x},\overline{t})]$ has rank $n-1$ and the matrix $\nabla_{(x,t)}[r_{I_r(\overline{x},\overline{t})}(\overline{x},\overline{t}), s_{I_s(\overline{x},\overline{t})}(\overline{x},\overline{t})]$ has rank n.

(c) $I_{rs}(\overline{x},\overline{t})$ is a singleton, i.e., $I_{rs}(\overline{x},\overline{t}) = \{i^*\}$, and the matrix $\nabla_{(x,t)}\left[r_{I_r(\overline{x},\overline{t})}(\overline{x},\overline{t}), s_{I_s(\overline{x},\overline{t})}(\overline{x},\overline{t}), r_{i^*}(\overline{x},\overline{t}), s_{i^*}(\overline{x},\overline{t})\right]$ is non-singular.

Proof. Let the sets $I_r(\overline{x},\overline{t}), I_{rs}(\overline{x},\overline{t}), I_s(\overline{x},\overline{t})$ be fixed. We will assume, w.l.o.g., that $I_s(\overline{x},\overline{t}) = \varnothing$ and $I_{rs}(\overline{x},\overline{t}) = \{1,\ldots,l_0\}$, *i.e.*, $r(\overline{x},\overline{t}) = 0$ and $s_1(\overline{x},\overline{t}) = \ldots = s_{l_0}(\overline{x},\overline{t}) = 0$. By the Jet Transversality Theorem, *cf.* Theorem 2.3.1, generically in (r,s), the jet manifold $(x,t,r(x,t),s(x,t))$ and the manifold $\mathbb{R}^n \times \mathbb{R} \times 0_{n+l_0} \times \mathbb{R}^{n-l_0}$ intersect transversally.
In other words, if we define

$$M_1 = \left\{ y = (x,t,z,u) \in \mathbb{R}^{n+1+n+n} \;\middle|\; \begin{array}{rcl} r(x,t) - z & = & 0, \\ s(x,t) - u & = & 0 \end{array} \right\}$$

and

$$M_2 = \left\{ y = (x,t,z,u) \in \mathbb{R}^{n+1+n+n} \;\middle|\; \begin{array}{rcl} z & = & 0, \\ u_i & = & 0, \quad i = 1,\ldots,l_0 \end{array} \right\}$$

then at all $\overline{y} \in M_1 \cap M_2$, the gradients of the functions defining M_1 and M_2 are linearly independent. This means that, the following matrix has full row rank at all $\overline{y} \in M_1 \cap M_2$,

$$\begin{array}{cccc} \partial x & \partial t & \partial z & \partial u \\ \hline \nabla_x r & \nabla_t r & -I_n & 0 \\ \nabla_x s & \nabla_t s & 0 & -I_n \\ 0 & 0 & I_n & 0 \\ 0 & 0 & 0 & I_{l_0}|0 \end{array}$$

Then, the number of columns $(n+1+2n)$ must be greater than or equal to the number of rows $2n+n+l_0$, *i.e.*, $l_0 \leq 1$. So, if $l_0 = 0$, due to the rank conditions, $\nabla_{(x,t)} r(\overline{x},\overline{t})$ has rank n. In this case either $\nabla_x r(\overline{x},\overline{t})$ is non-singular or $\nabla_x r(\overline{x},\overline{t})$ has rank $n-1$, corresponding to conditions (a) and (b). If $l_0 = 1$ the non-singularity of $\nabla_{(x,t)} \left[r_{I_r(\overline{x},\overline{t})}(\overline{x},\overline{t}), r_{i^*}(\overline{x},\overline{t}), s_{i^*}(\overline{x},\overline{t}) \right]$ follows from the rank condition. The result is now a consequence of intersecting the generic sets resulting for all finitely many possible combinations of $I_r(\overline{x},\overline{t}), I_{rs}(\overline{x},\overline{t}), I_s(\overline{x},\overline{t})$.

□

Remark 4.7.1 *If we consider the 1-parametric NLCP, see (4.7.3), as the variational inequality problem $VI(t; \Phi(x,t), \mathbb{R}^n_+)$, the generic possibilities of Proposition 4.7.2 can be identified with the following cases of 1-parametric VI g.c. points, see Definition 3.2.2 and [19]:*
Points satisfying (*a*) *are non-degenerate critical points of $VI(t; \Phi(x,t), \mathbb{R}^n_+)$.*
Case (*b*) *corresponds to points where the condition VI-1c in Definition 3.2.2 is violated. Finally for points satisfying* (*c*) *the condition VI-1b, see Definition 3.2.2, fails.*

Now we will specify the conditions given in Proposition 4.7.2 (b) and (c) and define generic singularities in order to determine the local behavior of $\mathcal{M}$. In the case of NLCP, these singularities will coincide with those described for the 1-parametric problem VI in Section 3.3.

For a point $(\overline{x},\overline{t}) \in \mathcal{M}$ of type VI-2 we need the fulfilment of condition (c) of Proposition 4.7.2 and the regularity of the matrices :

$$\begin{aligned} &[\nabla_x r_{I_r(\overline{x},\overline{t})}(\overline{x},\overline{t}),\ \nabla_x s_{I_s(\overline{x},\overline{t})}(\overline{x},\overline{t}), \nabla_x r_{i^*}(\overline{x},\overline{t})] \\ &[\nabla_x r_{I_r(\overline{x},\overline{t})}(\overline{x},\overline{t}),\ \nabla_x s_{I_s(\overline{x},\overline{t})}(\overline{x},\overline{t}), \nabla_x s_{i^*}(\overline{x},\overline{t})]. \end{aligned} \tag{4.7.6}$$

Proposition 4.7.3 *Generically, with respect to the $[C^3_S]$-topology on $[C^\infty]^{n+n}_{n+1}$, the matrices given in (4.7.6) are regular at the points $(\overline{x},\overline{t})$ satisfying (c) in Proposition 4.7.2.*

Under these assumptions, for the case $I_{rs}(\overline{x},\overline{t}) = \{i^*\}$, the set $\mathcal{M}$ bifurcates into two branches: one corresponding to the solutions of $r_{I_r(\overline{x},\overline{t})}(x,t) = 0$, $s_{I_s(\overline{x},\overline{t})}(x,t) = 0$, $r_{i^*}(x,t) = 0$, $s_{i^*}(x,t) \geq 0$ and the other given by $r_{I_r(\overline{x},\overline{t})}(x,t) = 0$, $s_{I_s(\overline{x},\overline{t})}(x,t) = 0$, $s_{i^*}(x,t) = 0$, $r_{i^*}(x,t) \geq 0$.

Proof. Let $(\overline{x},\overline{t}) \in \mathcal{M}$ satisfy the condition given in Proposition 4.7.2-c. W.l.o.g., we assume $I_s(\overline{x},\overline{t}) = \varnothing$ and $I_{rs}(\overline{x},\overline{t}) = \{1\}$. First note that $\nabla_x r(\overline{x},\overline{t})$ will be singular if and only if, up to some permutations of columns and rows, the following relation holds for some $n_0 > 0$:

$$r_a = r_c r_d^{-1} r_b,$$

where $\nabla_x r$ is decomposed as $\nabla_x r = \begin{pmatrix} r_a & r_c \\ r_b & r_d \end{pmatrix}$, $rank(\nabla_x r) = rank(r_d)$ and $r_a \in \mathbb{R}^{n_0 \times n_0}$.

Now we fix a possible decomposition and consider the manifold

$$M_2 = \left\{ y = (x,t,z,u,z^1,u^1) \in \mathbb{R}^{n+1+n+1+n(n+1)+(n+1)} \;\middle|\; \begin{array}{rcl} z &=& 0, \\ u &=& 0, \\ z_a^1 - z_c^1[z_d^1]^{-1}z_b^1 &=& 0 \end{array} \right\}$$

and the jet manifold

$$M_1 = \left\{ y = (x,t,z,u,z^1,u^1) \in \mathbb{R}^{n+1+n+1+n(n+1)+(n+1)} \;\middle|\; \begin{array}{rcl} r(x,t) - z &=& 0, \\ s_1(x,t) - u &=& 0, \\ \nabla_{(x,t)} r(x,t) - z^1 &=& 0, \\ \nabla_{(x,t)} s_1(x,t) - u^1 &=& 0 \end{array} \right\}$$

in the corresponding jet space. Here $z^1 = \begin{pmatrix} z_a^1 & z_c^1 & z_e^1 \\ z_b^1 & z_d^1 & z_f^1 \end{pmatrix}$, the sub-matrix $\begin{pmatrix} z_a^1 & z_c^1 \\ z_b^1 & z_d^1 \end{pmatrix}$ corresponds to $\nabla_x r(x,t)$ and the vector $\begin{pmatrix} z_e^1 \\ z_f^1 \end{pmatrix}$ to $\nabla_t r(x,t)$. Using the Jet Transversality Theorem, Theorem 2.3.1, generically in (r,s), it follows $M_1 \pitchfork M_2$. This means that the following matrix has generically full row rank at $y \in M_1 \cap M_2$.

$$\begin{array}{cccccc}
\partial x & \partial t & \partial z & \partial u & \partial z^1 & \partial u^1 \\
\hline
\nabla_x r & \nabla_t r & -I_n & 0 & 0 & 0 \\
\nabla_x s_1 & \nabla_t s_1 & 0 & -1 & 0 & 0 \\
\begin{matrix} \nabla_x^2 r \\ \nabla_t \nabla_x r \end{matrix} & \begin{matrix} \nabla_t \nabla_x r \\ \nabla_t^2 r \end{matrix} & 0 & 0 & -I_{n(n+1)} & 0 \\
\begin{matrix} \nabla_x^2 s_1 \\ \nabla_t \nabla_x s_1 \end{matrix} & \begin{matrix} \nabla_t \nabla_x s_1 \\ \nabla_t^2 s_1 \end{matrix} & 0 & 0 & 0 & -I_{n+1} \\
0 & 0 & I_n & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & I_{n_0^2} | \bigotimes & 0
\end{array}$$

This implies that the number of columns is greater than or equal to the number of rows, *i.e.*, $n+1+n+1+(n+1)(n+1) \geq n+1+(n+1)(n+1)+n+1+n_0^2$. So $n_0 = 0$ or, equivalently, $\nabla_x r$ has full rank. Taking all possible decompositions of $\nabla_x r$, it follows that, generically, $\nabla_x r(\overline{x},\overline{t})$ is regular as desired. Analogously, generically the matrix $(\nabla_x r_2(\overline{x},\overline{t}), \ldots, \nabla_x r_n(\overline{x},\overline{t}) \nabla_x s_1(\overline{x},\overline{t}))$ is also regular.

We have proven that, generically, the matrices given in (4.7.6) are non-singular for $(\overline{x},\overline{t}) \in \mathcal{M}$ satisfying (c) in Proposition 4.7.2. Under this regularity condition, we can guarantee the existence of curves $(x_1(t),t)$ and $(x_2(t),t)$ satisfying $r(x_1(t),t) = 0$, $t \in [\overline{t}-\epsilon, \overline{t}+\epsilon]$, $r_i(x_2(t),t) = 0$, $i = 2,\ldots,n$, $s_1(x_2(t),t) = 0$, $t \in [\overline{t}-\epsilon,\overline{t}+\epsilon]$ and $x_1(\overline{t}) = x_2(\overline{t}) = \overline{x}$ for some $\epsilon > 0$.

The relation $r(x_1(t),t)=0$ implies $\nabla_x r\dot{x}_1+\nabla_t r=0$. Since the matrix $\nabla_{(x,t)}(r,s_1)$ is non-singular, it must follow that $\nabla_x s_1\dot{x}_1+\nabla_t s_1\neq 0$. This means that around $\bar{t}$, $s_1(x_1(t),t)$ is either an increasing or a decreasing function of t. In any case, by moving along the curve, in which $r(x_1(t),t)=0$ holds, one and only one of the branches, either for $t\geq\bar{t}$ or $t\leq\bar{t}$, will be (locally) feasible with $s_1(x_1(t),t)\geq 0$. A similar situation occurs if we consider the curve $(x_2(t),t)$ given by $s_1(x_2(t),t)=0$, $r_i(x_2(t),t)=0, i=2,\ldots,n$.

□

In the case of points $(\overline{x},\bar{t})$ of type VI-3, see Section 3.3, higher order conditions should be added to condition (b) in Proposition 4.7.2. We give the following result without proof. For the proof, which can be done along the lines of the proof of Proposition 4.7.3, a 2-jet manifold will also be required.

Proposition 4.7.4 *Generically in $[C^3_S]^{n+n}_{n+1}$, the points $(\overline{x},\bar{t})\in\mathcal{M}$, where condition (b) in Proposition 4.7.2, holds are non-degenerate critical points of:*

$$\begin{array}{rrcl} & \min t & & \\ s.t. & r_{I_r(\overline{x},\bar{t})}(x,t) & = & 0, \\ & s_{I_s(\overline{x},\bar{t})}(x,t) & = & 0. \end{array}$$

At points $(\overline{x},\bar{t})$ where the conditions of the previous proposition hold, the set $\mathcal{M}$ has a turning point.

Concluding, we have obtained the generic types of solution points for problem (4.7.2). In the case of NLCP, these types coincide with the singular points defined in [19] for one-parametric VI.

4.7.2 One parametric P_{CC}

In this section we extend the concepts and regularity results of standard 1-parametric finite problems (2.2.1) to 1-parametric mathematical programs with complementarity constraints. We will also present the singularities that may appear generically at g.c. points.

Let us consider the parametric problem $P_{CC}(t),\ t\in\mathbb{R}$, where

$$P_{CC}(t): \qquad \begin{array}{rc} & \min f(x,t) \\ \text{s.t.} & x\in\mathcal{M}_{CC}(t) \end{array} \tag{4.7.7}$$

$$\mathcal{M}_{CC}(t)=\left\{ x\in\mathbb{R}^n \;\middle|\; \begin{array}{rcll} h_k(x,t) & = & 0, & k=1,\ldots,q_0, \\ g_j(x,t) & \geq & 0, & j=1,\ldots,q, \\ r_i(x,t) & \geq & 0, & i=1,\ldots,l, \\ s_i(x,t) & \geq & 0, & i=1,\ldots,l, \\ r_i(x,t)s_i(x,t) & = & 0, & i=1,\ldots,l. \end{array} \right\}$$

These problems have been studied in [24] from a local viewpoint. The continuity and differentiability of the value function with respect to the parameters was analyzed around a non-degenerate stationary point. Here we will study the generic behavior of the whole set of solution points.

To cover all possible singularities, we analyze the problem with additional equality constraints. As in the non-parametric case, the active index sets are:

$$\begin{array}{rcl} J_0(\overline{x},\overline{t}) & = & \left\{j \mid g_j(\overline{x},\overline{t}) = 0\right\}, \\ I_r(\overline{x},\overline{t}) & = & \left\{i \mid r_i(\overline{x},\overline{t}) = 0,\ s_i(\overline{x},\overline{t}) > 0\right\}, \\ I_s(\overline{x},\overline{t}) & = & \left\{i \mid s_i(\overline{x},\overline{t}) = 0,\ r_i(\overline{x},\overline{t}) > 0\right\}, \\ I_{rs}(\overline{x},\overline{t}) & = & \left\{i \mid r_i(\overline{x},\overline{t}) = s_i(\overline{x},\overline{t}) = 0\right\}. \end{array}$$

The definitions of the Lagrange function, multipliers, stationarity, MPCC-SC and MPCC-SOC can also be easily extended from the non-parametric (see Section 4.3) to the parametric case.

Definition 4.7.1 *Let*

$$L(x,t,\lambda_0,\lambda,\gamma,\mu,\rho,\sigma) = \begin{array}{c} \lambda_0 f(x,t) - \sum\limits_{k=1}^{q_0} \lambda_k h_k(x,t) - \sum\limits_{i\in I_{rs}(\overline{x},\overline{t})} (\rho_i r_i(x,t) + \sigma_i s_i(x,t)) \\ - \sum\limits_{j\in J_0(x,t)} \mu_j g_j(x,t) - \sum\limits_{i\in I_r(\overline{x},\overline{t})} \rho_i r_i(x,t) - \sum\limits_{i\in I_s(\overline{x},\overline{t})} \sigma_i s_i(x,t) \end{array}$$

be the Lagrange function. Then for problem $P_{CC}(t)$:

- *MPCC-LICQ (MPCC-MFCQ) holds at $(\overline{x},\overline{t})$ if it is satisfied at $\overline{x} \in \mathcal{M}_{CC}(\overline{t})$.*
- *$(\overline{x},\overline{t})$ is a Fritz John, (weakly, C-, M-, A-, B-, strongly) stationary point if $\overline{x}$ is a Fritz John, (weakly, C-, M-, A-, B-, strongly) stationary point of problem $P_{CC}(\overline{t})$. The set of strongly stationary points will be denoted as Σ_{stat}.*
- *MPCC-SC(SOC) holds at $(\overline{x},\overline{t})$ if MPCC-SC(SOC) holds at $\overline{x}$ for problem $P_{CC}(\overline{t})$.*

As already done in (4.3.3), we will define the relaxed problem corresponding to $P_{CC}(t)$ at a feasible point $(\overline{x},\overline{t})$

$$P_R^{(\overline{x},\overline{t})}(t): \qquad \begin{array}{c} \min f(x,t) \\ \text{s.t} \quad x \in \mathcal{M}_R^{(\overline{x},\overline{t})}(t) \end{array} \tag{4.7.8}$$

$$\mathcal{M}_R^{(\overline{x},\overline{t})}(t) = \left\{ x \in \mathbb{R}^n \;\middle|\; \begin{array}{rcl} h_k(x,t) & = & 0, \quad k = 1,\dots,q_0, \\ g_j(x,t) & \geq & 0, \quad j = 1,\dots,q, \\ r_i(x,t) = 0, & & s_i(x,t) \geq 0, \quad i \in I_r(\overline{x},\overline{t}), \\ s_i(x,t) = 0 & & r_i(x,t) \geq 0, \quad i \in I_s(\overline{x},\overline{t}), \\ r_i(x,t) \geq 0, & & s_i(x,t) \geq 0, \quad i \in I_{rs}(\overline{x},\overline{t}). \end{array} \right\}$$

The next lemma establishes the relationship between $\mathcal{M}_R^{(\overline{x},\overline{t})}(t)$ and the set of feasible solutions of the original problem.

Lemma 4.7.1 *Let $\overline{x} \in \mathcal{M}_{CC}(\overline{t})$ be given. Then there is a neighborhood $V_x \times V_t$ of $(\overline{x},\overline{t})$ such that for all $t \in V_t$, it follows that $V_x \cap \mathcal{M}_{CC}(t) \subset \mathcal{M}_R^{(\overline{x},\overline{t})}(t)$. Moreover, if $t \in V_t$ and $x \in V_x \cap \mathcal{M}_{CC}(t)$, then $I_{rs}(x,t) \subset I_{rs}(\overline{x},\overline{t})$.*

Proof. By continuity, there is a neighborhood $V_0 = V_x \times V_t$ of $(\overline{x},\overline{t})$ such that $(x,t) \in V_0$ implies $s_i(x,t) > 0,\ i \in I_r(\overline{x},\overline{t})$ and $r_i(x,t) > 0,\ i \in I_s(\overline{x},\overline{t})$. If $(x,t) \in V_0$ and $x \in \mathcal{M}_{CC}(t)$ then $r_i(x,t) = 0,\ i \in I_r(\overline{x},\overline{t})$ and $s_i(x,t) = 0$, $i \in I_s(\overline{x},\overline{t})$ while for $i \in I_{rs}(\overline{x},\overline{t}),\ r_i(x,t)s_i(x,t) = 0,\ r_i(x,t),\ s_i(x,t) \geq 0$. This means $x \in \mathcal{M}_R^{(\overline{x},\overline{t})}(t)$, the feasible set of the parametric relaxed problem $P_R^{(\overline{x},\overline{t})}(t)$. As a direct consequence $I_{rs}(x,t) \subset I_{rs}(\overline{x},\overline{t})$ for all (x,t) such that $t \in V_t$, $x \in \mathcal{M}_{CC}(t) \cap V_x$.

□

As usual, instead of studying the structure of the set of local minimizers or stationary points of $P_{CC}(t)$, it is more natural to analyze the larger set of generalized critical points.

Definition 4.7.2 We say that $(\overline{x},\overline{t})$ is *a generalized critical point of* $P_{CC}(t)$*, i.e.,* $(\overline{x},\overline{t}) \in \Sigma_{gc}(P_{CC}(t))$, if $\overline{x} \in \mathcal{M}_{CC}(\overline{t})$ and $\nabla_x f(\overline{x},\overline{t}), \nabla_x h_1(\overline{x},\overline{t}), \ldots, \nabla_x h_{q_0}(\overline{x},\overline{t})$, $\nabla_x g_{J_0(\overline{x},\overline{t})}(\overline{x},\overline{t})$, $\nabla_x r_{I_r(\overline{x},\overline{t})}(\overline{x},\overline{t})$, $\nabla_x r_{I_{rs}(\overline{x},\overline{t})}(\overline{x},\overline{t}), \nabla_x s_{I_{rs}(\overline{x},\overline{t})}(\overline{x},\overline{t}), \nabla_x s_{I_s(\overline{x},\overline{t})}(\overline{x},\overline{t})$ are linearly dependent.
It is *a critical point (i.e. $(\overline{x},\overline{t}) \in \Sigma_{crit}(P_{CC}(t))$)* if in addition MPCC-LICQ holds.
Let $(1,\lambda,\mu,\sigma,\rho)$ be a vector of multipliers associated to a g.c. point $(\overline{x},\overline{t})$, *i.e.,* $\nabla_x L(\overline{x},\overline{t},1,\lambda,\mu,\sigma,\rho) = 0$. We will say that *MPCC-SC holds* if $\mu_j \neq 0$, $\forall j \in J_0(\overline{x},\overline{t})$ and $\sigma_i,\ \rho_i \neq 0,\ \forall i \in I_{rs}(\overline{x},\overline{t})$.

From the definition it directly follows that $(\overline{x},\overline{t})$ is a generalized critical point (respectively a critical point) of $P_{CC}(t)$ if and only if it is a g.c point (respectively a critical point) of $P_R^{(\overline{x},\overline{t})}(t)$. Together with Lemma 4.7.1 we will see that, even more, locally around $(\overline{x},\overline{t})$ any g.c. point of $P_{CC}(t)$ is an element of $\Sigma_{gc}(P_R^{(\overline{x},\overline{t})}(t))$.

Lemma 4.7.2 *If $(\overline{x},\overline{t}) \in \Sigma_{gc}(P_{CC}(t))$, then around this point*

$$\Sigma_{gc}(P_{CC}(t)) \subset \Sigma_{gc}(P_R^{(\overline{x},\overline{t})}(t)).$$

Moreover, locally, $\Sigma_{gc}(P_{CC}(t)) = \Sigma_{gc}(P_R^{(\overline{x},\overline{t})}(t)) \cap \{(x,t) \in \mathbb{R}^n \times \mathbb{R} \mid x \in \mathcal{M}_{CC}(t)\}$.

Proof. It is a direct consequence of the definitions and Lemma 4.7.1. □

These facts enable us to reduce locally the whole genericity analysis for $P_{CC}(t)$ to the analysis of the relaxed problems $P_R^{(\overline{x},\overline{t})}(t)$.

Lemma 4.7.3 *The set of functions $(f, h, g, r, s) \in [C^3_S]^{1+q_0+q+2l}_{n+1}$ such that for all possible choices of $I_r(\overline{x}, \overline{t}), I_s(\overline{x}, \overline{t}), I_{rs}(\overline{x}, \overline{t})$, the associated problems $P_R^{(\overline{x},\overline{t})}(t)$ are JJT-regular, contains an open and dense subset of $[C^3_S]^{1+q_0+q+2l}_{n+1}$.*

Proof. We apply the same idea as in the proof of Theorem 4.4.1. For a fixed triplet I_r, I_s, I_{rs}, the associated relaxed problem $P_R^{(\overline{x},\overline{t})}(t)$ is a common parametric nonlinear program. Using the classical JJT-theorem, see Theorem 2.4.1, $P_R^{(\overline{x},\overline{t})}(t)$ is JJT-regular for an open and dense set $\mathcal{S}(I_r, I_s, I_{rs})$ of functions $(f, h, g, r, s) \in [C^3_S]^{1+q_0+q+2l}_{n+1}$.
Now we consider all the finitely many possible combinations of I_r, I_s, I_{rs} and define the intersection $\mathcal{S} = \bigcap \mathcal{S}(I_r, I_s, I_{rs})$. As $\mathcal{S}$ is a finite intersection of open and dense sets, it is an open and dense set, and the elements in $\mathcal{S}$ have the desired properties.

$\square$

We have proven that generically the generalized critical points of $P_R^{(\overline{x},\overline{t})}(t)$ are of type 1, 2, 3, 4 or 5. In view of Definition 4.7.2 and Lemma 4.7.3 we can introduce the different types of g.c. points of $P_{CC}(t)$.

Definition 4.7.3 A g.c. point $(\overline{x}, \overline{t})$ of problem $P_{CC}(t)$ is *of type i*, $i = 1, 2, \ldots, 5$, if it is a g.c. point of type i for the corresponding problem $P_R^{(\overline{x},\overline{t})}(t)$.
A problem $P_{CC}(t)$ is called *MPCC-regular* if all its generalized critical points are of type $1, \ldots, 5$.

As in the general case, the singular points of type $i = 2, \ldots, 5$ are isolated singularities. The following result is a direct consequence of the previous definition, Lemma 4.7.2 and Lemma 4.7.3.

Theorem 4.7.1 *The set of functions $(f, h, g, r, s) \in [C^3_S]^{1+q_0+q+2l}_{n+1}$ such that $P_{CC}(t)$ is MPCC-regular contains an open and dense subset of $[C^3_S]^{1+q_0+q+2l}_{n+1}$.*

In the following we describe the behavior of the set $\Sigma_{gc}(P_{CC}(t))$ around any of the 5 types of g.c. points in detail. We will see that, locally, $\Sigma_{gc}(P_{CC}(t))$ can be written as the union of the set of g.c. points of certain relaxed problems. We will always consider $(\overline{x}, \overline{t})$ as an element of $\Sigma_{gc}(P_R^{(\overline{x},\overline{t})}(t))$, even if eventually it can be considered as a g.c. point of some other relaxed problem. Note that for MPCC-problems the active index sets do not change in the same way as for nonlinear programs without complementarity constraints. In fact, the active index sets I_r, I_{rs}, I_s should always form a partition of $\{1, \ldots, l\}$. Of course, $J_0(x, t)$ may change freely.

For the analysis of the behavior of the set of g.c. points, we again introduce the so-called linear and quadratic indexes. For a g.c. point $(\overline{x}, \overline{t})$ of $P_{CC}(t)$ we define the indices LI, (LNI), as the number of positive (negative) multipliers

$\mu_j, \sigma_i, \rho_i,\ j \in J_0(\overline{x},\overline{t}),\ i \in I_{rs}(\overline{x},\overline{t})$ of the Lagrangean function associated with $g_j, j \in J_0(\overline{x},\overline{t}),\ r_i, s_i,\ i \in I_{rs}(\overline{x},\overline{t})$. The quadratic indices QI, (QNI), represent the number of positive (negative) eigenvalues of the corresponding matrix $\nabla_x^2 L|_{T_{\overline{x}}\mathcal{M}_R^{(\overline{x},\overline{t})}(\overline{t})}$. Evidently, if $(\overline{x},\overline{t})$ is a strongly stationary point then LNI=0 must hold and at a local minimizer where *MPCC-LICQ* holds, LNI=QNI=0. Changes in these indexes imply leaving or entering the set of local minimizers and/or the set of stationary points.

G.C. points of type 1. If $(\overline{x},\overline{t}) \in \Sigma_{gc}^1(P_{CC}(t))$, *i.e.*, $(\overline{x},\overline{t}) \in \Sigma_{gc}^1(P_R^{(\overline{x},\overline{t})}(t))$, the set of active indexes will not change locally in $\Sigma_{gc}^1(P_R^{(\overline{x},\overline{t})}(t))$. So around $(\overline{x},\overline{t})$, $\Sigma_{gc}^1(P_R^{(\overline{x},\overline{t})}(t)) \subset \{(x,t) \mid\ x \in \mathcal{M}_{CC}(t)\}$ and, by Lemma 4.7.2, $\Sigma_{gc}(P_{CC}(t)) = \Sigma_{gc}(P_R^{(\overline{x},\overline{t})}(t))$, locally. Consequently the following holds in a neighborhood $V_{\overline{x}} \times V_{\overline{t}}$ of $(\overline{x},\overline{t})$:

- For any $(x,t) \in \Sigma_{gc}(P_{CC}(t)) \cap (V_x \times V_t)$, the active index sets do not change and are $J_0(x,t) = J_0(\overline{x},\overline{t}),\ I_r(x,t) = I_r(\overline{x},\overline{t}),\ I_s(x,t) = I_s(\overline{x},\overline{t})$, and $I_{rs}(x,t) = I_{rs}(\overline{x},\overline{t})$.
- $\Sigma_{gc}(P_{CC}(t)) \cap (V_x \times V_t) = \Sigma_{gc}^1(P_R^{(\overline{x},\overline{t})}) \cap (V_x \times V_t)$.
- There is a continuous function $x : [\overline{t}-\epsilon, \overline{t}+\epsilon] \to \mathbb{R}^n,\ x(\overline{t}) = \overline{x},\ \epsilon > 0$, such that $\Sigma_{gc}(P_{CC}(t)) \cap (V_x \times V_t) = \{(x(t),t),\ [\overline{t}-\epsilon, \overline{t}+\epsilon]\}$.

Around a g.c. point $(\overline{x},\overline{t})$ of type 1, the multipliers $\rho_i, \sigma_i, \mu_j\ i \in I_{rs}(\overline{x},\overline{t})$, $j \in J_0(\overline{x},\overline{t})$ and the eigenvalues of $\nabla_x^2 L|_{T_x\mathcal{M}_R^{(\overline{x},\overline{t})}(\overline{t})}$ do not change their sign. In particular the indices $(LI,\ LNI,\ QI,\ QNI) = (a,\ b,\ c,\ d)$ are constant, see Figure 4.1. So all stationarity types remain locally stable.

The LICQ condition for (x,t) in $P_R^{(\overline{x},\overline{t})}(t)$ is equivalent to MPCC-LICQ in $P_{CC}(t)$. So B-stationarity and strong stationarity are equivalent. The fulfillment of the MPCC-SC condition implies the equivalence between M-stationarity and strong stationarity.

G.C. points of type 2. For the g.c. points of type 2 we have to distinguish between two cases.

Definition 4.7.4 A g.c. point $(\overline{x},\overline{t})$ is *a g.c. point of type 2a for* $P_{CC}(t)$ if it is a g.c. point of type 2 for $P_R^{(\overline{x},\overline{t})}(t)$, with $\mu_j = 0$, for some $j \in J_0(\overline{x},\overline{t})$.
It is *a g.c. point of type 2b for* $P_{CC}(t)$ if it is a g.c. point of type 2 for $P_R^{(\overline{x},\overline{t})}(t)$, with either $\rho_{i^*} = 0$ or $\sigma_{i^*} = 0$, for some $i^* \in I_{rs}(\overline{x},\overline{t})$.

In the case of a g.c. point of type 2a, the local structure is the same as in classical nonlinear programming around a g.c. point of type 2, and locally

$$\Sigma_{gc}(P_{CC}(t)) = \Sigma_{gc}(P_R^{(\overline{x},\overline{t})}(t)).$$

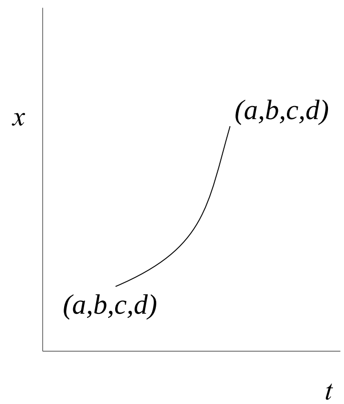


Figure 4.1: G.C. points of type 1

Let us consider a g.c. point $(\overline{x}, \overline{t})$ of type 2b. W.l.o.g. we assume $\sigma_{i^*} = s_{i^*}(\overline{x}, \overline{t}) = 0$ for some $i^* \in I_{rs}(\overline{x}, \overline{t})$. As it is a g.c. point of type 2 for $P_R^{(\overline{x},\overline{t})}(t)$, the possible active constraints for a g.c. point (x, t) of $P_R^{(\overline{x},\overline{t})}(t)$ near $(\overline{x}, \overline{t})$, are given by:

$$B_1\colon \begin{cases} g_j(x,t) &= 0, \quad j \in J_0(\overline{x}, \overline{t}), \\ r_i(x,t) &= 0, \quad i \in I_r(\overline{x}, \overline{t}) \cup I_{rs}(\overline{x}, \overline{t}), \\ s_k(x,t) &= 0, \quad k \in I_s(\overline{x}, \overline{t}) \cup I_{rs}(\overline{x}, \overline{t}). \end{cases}$$

$$B_2\colon \begin{cases} g_j(x,t) &= 0, \quad j \in J_0(\overline{x}, \overline{t}), \\ r_i(x,t) &= 0, \quad i \in I_r(\overline{x}, \overline{t}) \cup I_{rs}(\overline{x}, \overline{t}), \\ s_k(x,t) &= 0, \quad k \in I_s(\overline{x}, \overline{t}) \cup I_{rs}(\overline{x}, \overline{t}) \setminus \{i^*\}\,. \end{cases}$$

As $(\overline{x}, \overline{t}) \in \Sigma_{gc}^2(P_R^{(\overline{x},\overline{t})}(t))$, around $t = \overline{t}$, the set $\Sigma_{gc}(P_R^{(\overline{x},\overline{t})}(t))$ bifurcates at $(\overline{x}, \overline{t})$ into two parts. One part corresponds to the case B_1 and is described by a differentiable curve $\mathcal{C}^1 = \{(x^1(t), t) \mid t \in [\overline{t} - \epsilon, \overline{t} + \epsilon],\ x^1(\overline{t}) = \overline{x}\}$. The part associated to B_2, denoted as $\mathcal{C}^0$, has exactly one feasible branch: if $s_{i^*}(x, t) > 0$, for $t > \overline{t}$, the set $\mathcal{C}^0$ is $\{(x^0(t), t) \mid t \in [\overline{t}, \overline{t} + \epsilon],\ x^0(\overline{t}) = \overline{x}\}$ and if $s_{i^*}(x, t) > 0$, for $t < \overline{t}$, we have $\mathcal{C}^0 = \{(x^0(t), t) \mid t \in [\overline{t} - \epsilon, \overline{t}],\ x^0(\overline{t}) = \overline{x}\}$. These cases correspond to $D_t[s_{i^*}(x^0(t), t)](\overline{t}) > 0$ and $D_t[s_{i^*}(x^0(t), t)](\overline{t}) < 0$ respectively. Recall that since $(\overline{x}, \overline{t})$ is a g.c. point of type 2 of $P_R^{(\overline{x},\overline{t})}(t)$, $D_t[s_{i^*}(x^0(t), t)](\overline{t}) \neq 0$. Moreover on the curve $(x^1(t), t)$, $t \in [\overline{t} - \epsilon, \overline{t} + \epsilon]$, the multiplier $\sigma_{i^*}^1(t)$ corresponding to the active constraint $s_{i^*}(x, t)$ changes its sign around $\overline{t}$.

Note that, around $(\overline{x}, \overline{t})$, for the generalized critical points (x, t) of $P_R^{(\overline{x},\overline{t})}(t)$ it holds $x \in \mathcal{M}_{CC}(t)$. So, locally, $\Sigma_{gc}(P_{CC}(t)) = \Sigma_{gc}(P_R^{(\overline{x},\overline{t})}(t))$.
Summarizing, around a g.c. point of type 2b with $\sigma_{i^*} = s_{i^*}(\overline{x}, \overline{t}) = 0$ we have:

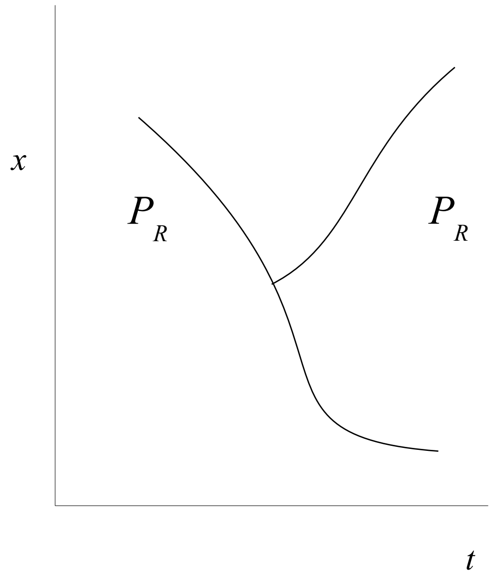


Figure 4.2: G.C. point of type 2a

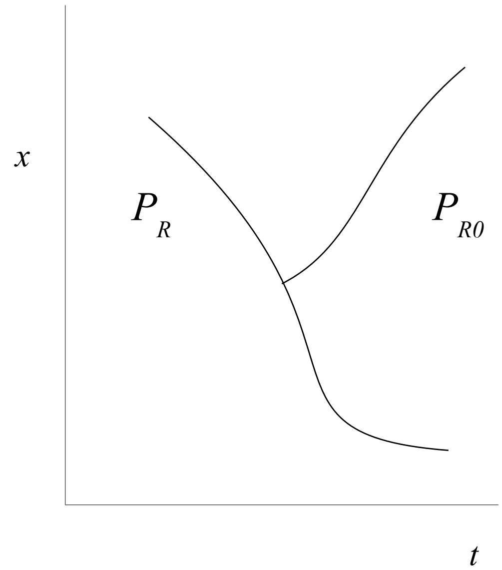


Figure 4.3: G.C. point of type 2b

- The possible combination of active indices sets for $(x,t) \in \Sigma_{gc}(P_{CC}(t))$ near $(\overline{x},\overline{t})$, are
  - $J_0(x,t) = J_0(\overline{x},\overline{t})$, $I_r(x,t) = I_r(\overline{x},\overline{t})$, $I_s(x,t) = I_s(\overline{x},\overline{t})$, $I_{rs}(x,t) = I_{rs}(\overline{x},\overline{t})$
  - $J_0(x,t) = J_0(\overline{x},\overline{t})$, $I_r(x,t) = I_r(\overline{x},\overline{t}) \cup \{i^*\}$, $I_s(x,t) = I_s(\overline{x},\overline{t})$, $I_{rs}(x,t) = I_{rs}(\overline{x},\overline{t}) \setminus \{i^*\}$.
- $\Sigma_{gc}(P_{CC}(t)) = \mathcal{C}^0 \cup \mathcal{C}^1 = \Sigma^1_{gc}(P_{CC}(t)) \cup \{\overline{x},\overline{t}\}$.
- Around $\overline{t}$ the multiplier $\sigma^1_{i^*}(t)$ changes its sign.

We want to point out that for $(x,t) \in \mathcal{C}^1$ the corresponding relaxed problem is given by $P_R^{(\overline{x},\overline{t})}(t)$, but it changes for $(x,t) \in \mathcal{C}^0$. We will denote it as $P_{R0}^{(\overline{x},\overline{t})}(t)$. The picture of the set of generalized critical points around a g.c. point of type 2a and 2b is depicted, respectively, in the Figures 4.2 and 4.3.
Now we are going to describe the local behavior with regard to stationarity. Around g.c. points of type 2a, a multiplier $\mu_j,\ j \in J_0(\overline{x},\overline{t})$ changes its sign. So, trivially no stationarity type is locally stable.

Let us consider the g.c. points of type 2b, with $\sigma_{i^*} = s_{i^*}(\overline{x},\overline{t}) = 0$. We will denote by $\sigma^1_i(t), \rho^1_i(t),\ i \in I_{rs}(\overline{x},\overline{t}),\ \mu^1_j(t),\ j \in J_0(\overline{x},\overline{t})$, the multipliers corresponding to $(x^1(t),t) \in \mathcal{C}^1$ with respect to the functions $s_i, r_i, g_j,\ i \in I_{rs}(\overline{x},\overline{t})$,

$j \in J_0(\overline{x},\overline{t})$. Analogously $\sigma_i^0(t), \rho_i^0(t),\ i \in I_{rs}(\overline{x},\overline{t}) \setminus \{i^*\}, \rho_{i^*}^0(t)\ \mu_j^0(t),\ j \in J_0(\overline{x},\overline{t})$ are the multipliers associated to $(x^0(t),t) \in \mathcal{C}^0$. Recall $s_{i^*}(x,t) > 0$, for all $(x,t) \in \mathcal{C}^0 \setminus \{(\overline{x},\overline{t})\}$. W.l.o.g. we assume:

$$\text{The multiplier } \sigma_{i^*}^1(t) \text{ is positive for } t > \overline{t} \text{ and negative for } t < \overline{t}. \tag{4.7.9}$$

Due to the conditions for g.c. points of type 2, it follows that the signs of $\sigma_i^1(t), \rho_i^1(t),\ \mu_j^1(t),\ i \in I_{rs}(\overline{x},\overline{t}) \setminus \{i^*\},\ j \in J_0(\overline{x},\overline{t})$ and $\rho_{i^*}^1(t)$ coincide with the signs of $\sigma_i^0(t), \rho_i^0(t), \mu_j^0(t),\ i \in I_{rs}(\overline{x},\overline{t}) \setminus \{i^*\},\ j \in J_0(\overline{x},\overline{t})$ and $\rho_{i^*}^0(t)$ respectively.

As *MPCC-LICQ* holds, *B*-stationarity and strong stationarity are equivalent around $(\overline{x},\overline{t})$. The fulfillment of the *MPCC-SC* condition (except for $(\overline{x},\overline{t})$) implies that *M*-stationarity and *B*-stationarity are also equivalent (locally) for $(x,t) \neq (\overline{x},\overline{t})$. So, let us analyze the behavior of strong stationarity.

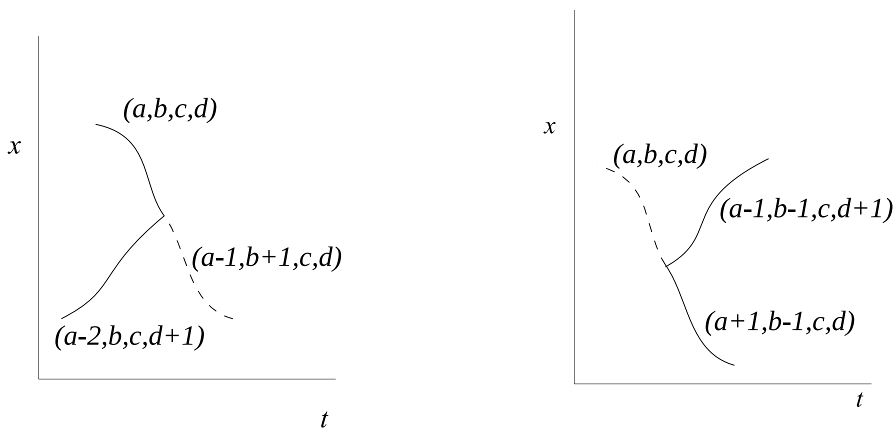
For simplicity, we assume that the points of $\mathcal{C}^0 \setminus \{\overline{x},\overline{t}\}$ are strongly stationary points. This means that $\sigma_i^0(t) > 0,\ \rho_i^0(t) > 0,\ \forall i \in I_{rs}(\overline{x},\overline{t}) \setminus \{i^*\},\ \mu_j^0(t) > 0,$ $j \in J_0(\overline{x},\overline{t})$ and the sign of $\rho_{i^*}^0(t)$ is free. Then, if $\rho_{i^*}^0(t) < 0$, the points of $\mathcal{C}^0 \setminus \{\overline{x},\overline{t}\}$ are strongly stationary points but as $i^* \in I_{rs}(x^1(t),t),\ t \in [\overline{t}-\epsilon, \overline{t}+\epsilon]$, and $\rho_{i^*}^1(t) < 0$, it follows that the elements of $\mathcal{C}^1$ are not strongly stationary points. If $\rho_{i^*}^0(t) > 0$, due to assumption (4.7.9), $\sigma_{i^*}^1(t) > 0$ only for $t > \overline{t}$. So, $(x^1(t),t)$ is a strongly stationary point if and only if $t > \overline{t}$.

As for the case 2b the multipliers $\mu_j,\ j \in J_0(\overline{x},\overline{t})$, do not change their sign, weakly stationarity locally remains stable.

Now let us assume that the points of $\mathcal{C}^0$ are *A*-stationary points. If $\rho_{i^*}^0(t) > 0$, *A*-stationarity is stable in $\mathcal{C}^1 \cup \mathcal{C}^0$. However, if $\rho_{i^*}^0(t) < 0$, at $(x^1(t),t) \in \mathcal{C}^1$, *A*-stationarity is equivalent with $\sigma_{i^*}^1(t) > 0$, so it will be fulfilled if and only if $t > \overline{t}$, see condition (4.7.9).

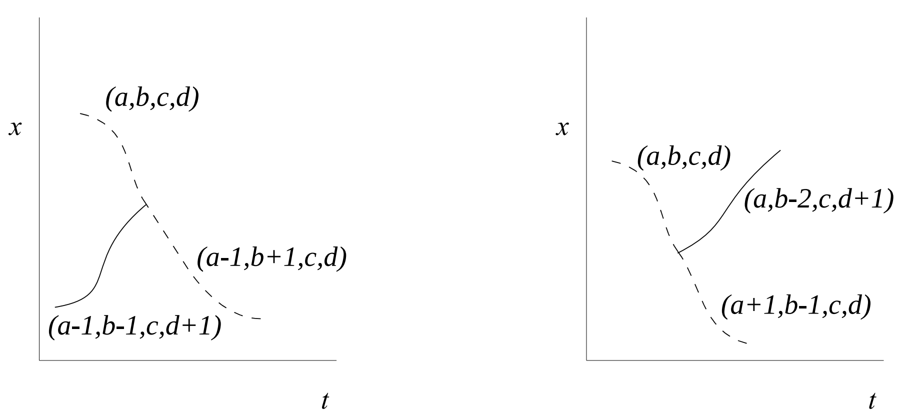
In case that C-stationarity holds at $\mathcal{C}^0$, again by condition (4.7.9), if $\rho_{i^*}^0(t) > 0$, the *C*-stationarity holds for $(x^1(t),t) \in \mathcal{C}^1$ if and only if $t > \overline{t}$. If $\rho_{i^*}^0(t) < 0$, it will be satisfied for $t < \overline{t}$. In both cases only one branch of $\mathcal{C}^1$, either for $t > \overline{t}$ or for $t < \overline{t}$ will contain C-stationary points.
With respect to the linear and quadratic indices, we can see that around a g.c. point of type 2a, they behave as around a g.c. point of type 2 in the nonlinear program $P_R^{(\overline{x},\overline{t})}(t)$. In the case of g.c. points of type 2b, we show all possible combinations of indices (LI, LNI, QI, QNI) in the Figures 4.4, 4.5, 4.6, and 4.7. There, all possible cases for the set of strongly stationary points are also considered. Assuming that the g.c. points in $\mathcal{C}^0$, $t \neq \overline{t}$, are strongly stationary points, the continuous line represents the set Σ_{stat} while the dotted line corresponds to the set $\Sigma_{gc} \setminus \Sigma_{stat}$.

G.C. points of type 3. At a g.c. point $(\overline{x},\overline{t})$ of type 3, the second order condition *MPCC-SOC* is violated. With the same analysis as for g.c. points of type 1, it can be seen that, locally, the same relaxed problem $P_R^{(\overline{x},\overline{t})}(t)$ remains



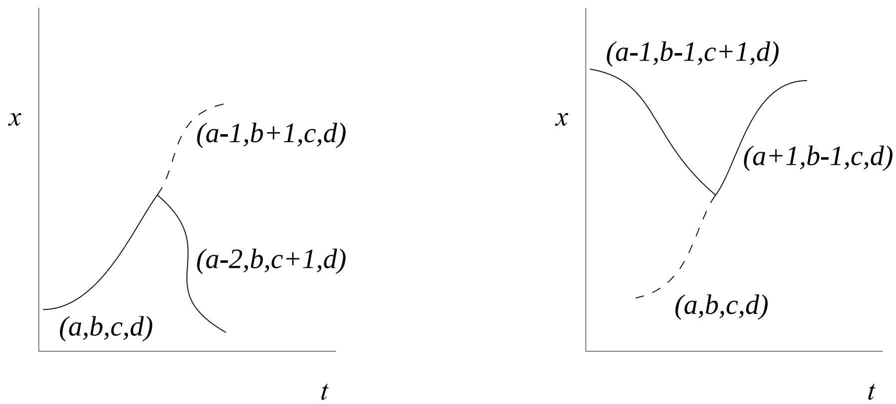
Case $\frac{d}{dt}[s_{i^*}(x^0(\bar{t}),\bar{t})] < 0$, $D_t\sigma^1_{i^*}(\bar{t}) < 0$. Case $\frac{d}{dt}[s_{i^*}(x^0(\bar{t}),\bar{t})] > 0$, $D_t\sigma^1_{i^*}(\bar{t}) > 0$.

Figure 4.4: Type 2b, with $\rho_{i^*}(\bar{t}) > 0$.



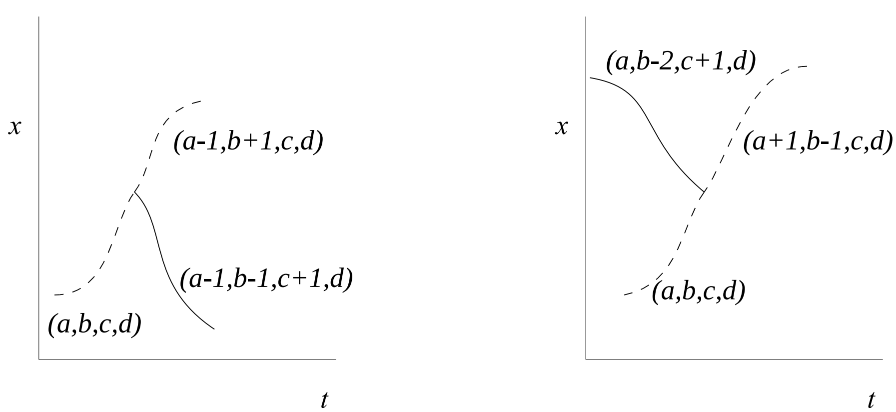
Case $\frac{d}{dt}[s_{i^*}(x^0(\bar{t}),\bar{t})] < 0$, $D_t\sigma^1_{i^*}(\bar{t}) < 0$. Case $\frac{d}{dt}[s_{i^*}(x^0(\bar{t}),\bar{t})] > 0$, $D_t\sigma^1_{i^*}(\bar{t}) > 0$.

Figure 4.5: Type 2b, with $\rho_{i^*}(\bar{t}) < 0$.



Case $D_t\sigma^1_{i^*}(\bar{t}) < 0$, $\frac{d}{dt}[s_{i^*}(x^0(\bar{t}), \bar{t})] > 0$. Case $D_t\sigma^1_{i^*}(\bar{t}) > 0$, $\frac{d}{dt}[s_{i^*}(x^0(\bar{t}), \bar{t})] < 0$.

Figure 4.6: Type 2b, with $\rho_{i^*}(\bar{t}) > 0$.



Case $\frac{d}{dt}[s_{i^*}(x^0(\bar{t}), \bar{t})] > 0$, $D_t\sigma^1_{i^*}(\bar{t}) < 0$. Case $\frac{d}{dt}[s_{i^*}(x^0(\bar{t}), \bar{t})] < 0$, $D_t\sigma^1_{i^*}(\bar{t}) > 0$.

Figure 4.7: Type 2b, with $\rho_{i^*}(\bar{t}) < 0$.

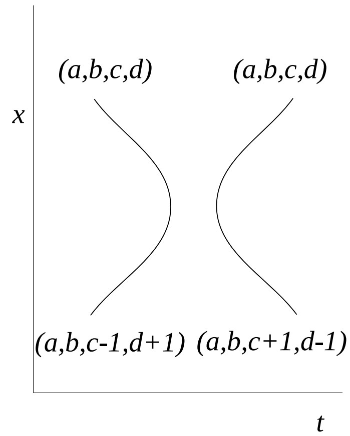


Figure 4.8: G.C. points of type 3

valid and that $\Sigma_{gc}(P_{CC}(t)) = \Sigma_{gc}(P_R^{(\overline{x},\overline{t})}(t))$. So Σ_{gc} has a quadratic turning point as at a g.c. point of type 3 for the standard parametric nonlinear problem $P_R^{(\overline{x},\overline{t})}(t)$. Moreover, this singularity is isolated and locally $\Sigma_{gc}(P_{CC}(t)) = \Sigma^1_{gc}(P_{CC}(t)) \cup \{(\overline{x},\overline{t})\}$ (see Figure 4.8).

With respect to stationarity, the situation is as in a g.c. point of type 1. Locally the stationarity type remains unchanged as the sign of the multipliers does not change.

Now we will consider the singular points in which *MPCC-LICQ* is not satisfied. In these cases, B-stationarity and strong stationarity are not necessarily equivalent.

G.C. points of type 4. Around a g.c. point $(\overline{x},\overline{t})$ of type 4, the active indices do not change in $\Sigma_{gc}(P_R^{(\overline{x},\overline{t})}(t))$, so locally:

$$\Sigma_{gc}(P_{CC}(t)) = \Sigma_{gc}(P_R^{(\overline{x},\overline{t})}(t)) = \Sigma^1_{gc}(P_{CC}(t)) \cup \{\overline{x},\overline{t}\}.$$

This means that $\Sigma_{gc}(P_{CC}(t))$ has a quadratic turning point around $(\overline{x},\overline{t})$. Moreover, all multipliers change their sign at $(\overline{x},\overline{t})$ and $\nabla_x f(\overline{x},\overline{t})$ is not a linear combination of the active constraints. In particular this means that $(\overline{x},\overline{t})$ cannot be a stationary point.

Now we will proceed with the analysis of the stationarity types. If $I_{rs}(x,t) = J_0(x,t) = \varnothing$ ($LI = LNI = 0$), all stationarity types are equivalent. Moreover, if (x,t) is a critical point then it will also be strongly stationary. As, locally, $I_{rs}(x,t) = J_0(x,t) = I_{rs}(\overline{x},\overline{t}) = J_0(\overline{x},\overline{t}) = \varnothing$, and MPCC-LICQ holds for all $(x,t) \neq (\overline{x},\overline{t})$, the g.c. points with $t \neq \overline{t}$ are strongly stationary points.

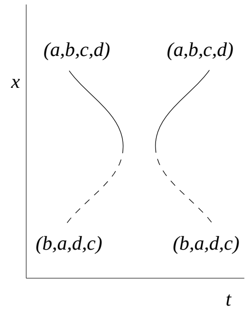


Figure 4.9: G.C. points of type 4 with $I_{rs}(\overline{x},\overline{t}) \cup J_0(\overline{x},\overline{t}) \neq \varnothing$

Now, consider the case $I_{rs}(\overline{x},\overline{t}) \neq \varnothing$ and $J_0(\overline{x},\overline{t}) = \varnothing$. As MPCC-LICQ and MPCC-SC holds except for $(\overline{x},\overline{t})$, B-, M- and strong stationarity are equivalent for $\Sigma_{gc}(P_{CC}(t)) \setminus \{\overline{x},\overline{t}\}$. So we only analyze the behavior of strong stationarity. But clearly the later is not stable, since around a g.c. point $(\overline{x},\overline{t})$ of type 4, the multipliers $\sigma_i, \rho_i,\ i \in I_{rs}(\overline{x},\overline{t})$, change their sign and at strongly stationary points they should be both positive. With respect to C-stationarity, only the product $\sigma_i \cdot \rho_i,\ i \in I_{rs}(\overline{x},\overline{t})$ has to be positive, so, locally, C-stationarity remains stable in $\Sigma_{gc}(P_{CC}(t)) \setminus \{\overline{x},\overline{t}\}$. As $J_0(\overline{x},\overline{t}) = \varnothing$, all points of $\Sigma_{gc}(P_{CC}(t)) \setminus \{\overline{x},\overline{t}\}$ are weakly stationary points. If there are A-stationary points, converging to $(\overline{x},\overline{t})$ and for some index $i \in I_{rs}(\overline{x},\overline{t})$ the associated multipliers σ_i, ρ_i are positive, then, when passing $(\overline{x},\overline{t})$, both will be negative. Consequently in this case the A-stationarity is not stable. As for A-stationary points either ρ_i or σ_i should be positive for all $i \in I_{rs}(\overline{x},\overline{t})$, when passing the singular point of type 4, A-stationarity is stable if and only if $\sigma_i \cdot \rho_i < 0,\ \forall i \in I_{rs}(\overline{x},\overline{t})$.

Finally, if $J_0(\overline{x},\overline{t}) \neq \varnothing$, then all stationarity types are not stable.

In Figure 4.9 we can see the local behavior of $\Sigma_{gc}(P_{CC}(t))$ around such a point. In the case that $b = 0, a > 0$, the full line represents the set Σ_{stat} and the dotted line, $\Sigma_{gc} \setminus \Sigma_{stat}$.

G.C. points of type 5. Let us begin with the characteristics of such a point and the consequences for the set of g.c. points of the associated relaxed problem $P_R^{(\overline{x},\overline{t})}(t)$.

For the problem $P_R^{(\overline{x},\overline{t})}(t)$, the point $(\overline{x},\overline{t}) \in \Sigma_{gc}^5(P_R^{(\overline{x},\overline{t})}(t))$ is an isolated singular point. Moreover, the set $\Sigma_{gc}(P_R^{(\overline{x},\overline{t})}(t))$ bifurcates around $(\overline{x},\overline{t})$. Each bifurcation branch corresponds to g.c. points where $\nabla_x f$ is a linear combination of the gradients of all active constraints, except for one active inequality constraint. The set of g.c. points is composed by the feasible points of $P_R^{(\overline{x},\overline{t})}(t)$ which are g.c. points of one of the nonlinear programs $(P^{g_{j^*}}(t))$, $(P^{s_{i^*}}(t))$, $(P^{r_{i^*}}(t))$ for $i^* \in I_{rs}(\overline{x},\overline{t})$, $j^* \in J_0(\overline{x},\overline{t})$. Here

$$
\begin{array}{llrcl}
P^{g_{j^*}}(t): & & & \min f(x,t) & \\
& \text{s.t.} & h_k(x,t) & = & 0, \quad k=1,\dots,q_0, \\
& & g_j(x,t) & = & 0, \quad j \in J_0(\overline{x},\overline{t}) \setminus \{j^*\}, \\
& & r_i(x,t) & = & 0, \quad i \in I_r(\overline{x},\overline{t}) \cup I_{rs}(\overline{x},\overline{t}), \\
& & s_i(x,t) & = & 0, \quad i \in I_s(\overline{x},\overline{t}) \cup I_{rs}(\overline{x},\overline{t}),
\end{array}
$$

corresponds to the case where the constraint $g_{j^*}(x,t),\ j^* \in J_0(\overline{x},\overline{t})$ is not longer active. In the same way for $s_{i^*}(x,t),\ i^* \in I_{rs}(\overline{x},\overline{t})$, we define $P^{s_{i^*}}(t)$ and for $r_{i^*}(x,t),\ i^* \in I_{rs}(\overline{x},\overline{t}),\ P^{r_{i^*}}(t)$.

Now we consider one of these problems, taking $P^{g_{j^*}}(t)$ w.l.o.g. Due to the condition

$$
rank\left(\nabla_{(x,t)}\left[h_1,\dots,h_{q_0},g_{J_0(\overline{x},\overline{t})},r_{I_r(\overline{x},\overline{t})},r_{I_{rs}(\overline{x},\overline{t})},s_{I_{rs}(\overline{x},\overline{t})}s_{I_s(\overline{x},\overline{t})}\right](\overline{x},\overline{t})\right) = n+1,
$$

the eliminated constraint $(g_{j^*}(x,t))$ will change its sign around the point $(\overline{x},\overline{t})$ in the set of g.c. points of $P^{g_{j^*}}(t)$. So, exactly one branch, corresponding to $t > \overline{t}$ or $t < \overline{t}$, will consist of g.c. points of $P_R^{(\overline{x},\overline{t})}(t)$. On this branch the eliminated constraint will be strictly positive. The branch where this happens, *i.e.*, either for $t > \overline{t}$ or for $t < \overline{t}$, can be determined by using the coefficients of the following (non trivial) linear combination, see *e.g.*, [21]:

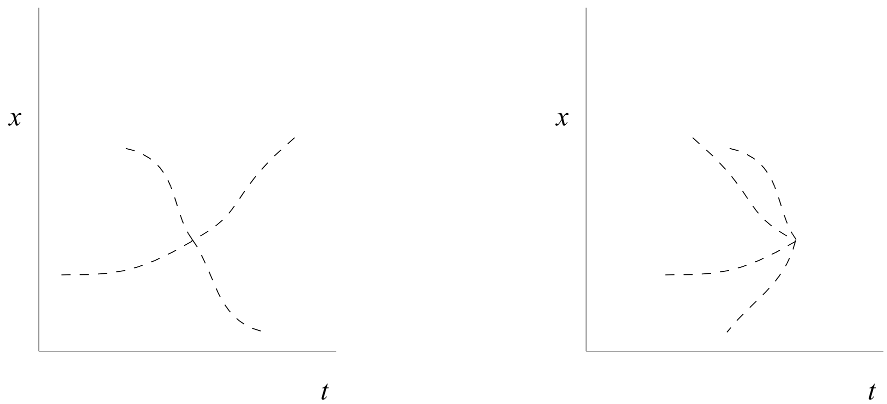
$$
\begin{aligned}
&\sum_{k=1}^{q_0} \lambda_k \nabla_x h_k(\overline{x},\overline{t}) + \sum_{j\in J_0(\overline{x},\overline{t})} \mu_j \nabla_x g_j(\overline{x},\overline{t}) + \\
+&\sum_{i\in I_r(\overline{x},\overline{t})\cup I_{rs}(\overline{x},\overline{t})} \rho_i \nabla_x r_i(\overline{x},\overline{t}) + \sum_{i\in I_s(\overline{x},\overline{t})\cup I_{rs}(\overline{x},\overline{t})} \sigma_i \nabla_x s_i(\overline{x},\overline{t}) \;=\; 0
\end{aligned}
\tag{4.7.10}
$$

Now we will present the meaning of these facts for the set of g.c. points of $P_{CC}(t)$.

Note that, for (x,t) near to $(\overline{x},\overline{t})$, any feasible point x of $P^{g_{j^*}}(t)$ with $g_{j^*}(x,t) \geq 0$, is also feasible for $\mathcal{M}_{CC}(t)$. The same holds in the case that x is a feasible point of $P^{s_{i^*}}(t)$ (respectively $P^{r_{i^*}}(t)$) and $s_{i^*}(x,t) \geq 0$ (respectively $r_{i^*}(x,t) \geq 0$). Consequently $\Sigma_{gc}(P_{CC}(t)) = \Sigma_{gc}(P_R^{(\overline{x},\overline{t})}(t))$ holds and $\Sigma_{gc}(P_{CC}(t))$ also bifurcates into $|J_0(\overline{x},\overline{t})| + 2|I_{rs}(\overline{x},\overline{t})|$ branches.

Now we will consider the relaxed problem that corresponds to each branch. On the branch where $g_{j^*}(x,t) > 0$ holds for some $j^* \in J_0(\overline{x},\overline{t})$, locally, the corresponding relaxed problem is given by $P_R^{(\overline{x},\overline{t})}(t)$. The branch in which $s_{i^*}(x,t)$, $i^* \in I_{rs}(\overline{x},\overline{t})$ is no longer active, corresponds to the active index sets, $J_0(\overline{x},\overline{t})$, $I_r(\overline{x},\overline{t}) \cup \{i^*\}, I_s(\overline{x},\overline{t}), I_{rs}(\overline{x},\overline{t}) \setminus \{i^*\}$. We denote the corresponding relaxed problem by $P_R^{s_{i^*}}(t)$. In the case of deactivating $r_{i^*},\ i^* \in I_{rs}(\overline{x},\overline{t})$ the corresponding active index sets are $J_0(\overline{x},\overline{t}), I_r(\overline{x},\overline{t}), I_s(\overline{x},\overline{t}) \cup \{i^*\}, I_{rs}(\overline{x},\overline{t}) \setminus \{i^*\}$ and the associated relaxed problem is denoted by $P_R^{r_{i^*}}(t)$.
Considering all possible combinations we can conclude that, locally, the following holds:

- $\Sigma_{gc}(P_{CC}(t)) = \Sigma_{gc}^1(P_{CC}(t)) \cup \{(\overline{x},\overline{t})\}$.
- $\Sigma_{gc}(P_{CC}(t))$ bifurcates into $|J_0(\overline{x},\overline{t})| + 2|I_{rs}(\overline{x},\overline{t})|$ branches.



If MFCQ holds If MFCQ fails

Figure 4.10: G.C. point of type 5

- $\Sigma_{gc}(P_{CC}(t)) = \bigcup_{i \in I_{rs}(\overline{x},\overline{t})} (\Sigma_{gc}(P_R^{s_i}(t)) \cup \Sigma_{gc}(P_R^{r_i}(t))) \bigcup \Sigma^0_{gc}(P_R^{(\overline{x},\overline{t})}(t))$, where $\Sigma^0_{gc}(P_R^{(\overline{x},\overline{t})}(t))$ denotes the set of g.c. points corresponding to all problems $P^{g_{j^*}}(t)$, which are the only g.c. points with associated relaxed problem $P_R^{(\overline{x},\overline{t})}(t)$.

Figure 4.10 shows the behavior of this set, when $I_{rs}(\overline{x},\overline{t}) = \{1\}$ and $J_0(\overline{x},\overline{t}) = \{1,2\}$.

As, locally, $\Sigma_{gc}(P_{CC}(t)) = \Sigma_{gc}(P_R^{(\overline{x},\overline{t})}(t))$ holds, the set of g.c. points will have a non-smooth turning point if and only if MFCQ fails at $\overline{x} \in \mathcal{M}_R^{(\overline{x},\overline{t})}(\overline{t})$.

We end with the analysis of the stationarity types. Let us suppose that there is a branch of (strongly ,M-, B-,C- or A-) stationary points. In this case, when moving to a new branch, the new multipliers will depend on the coefficients of the linear combination given in (4.7.10) and the multipliers at the original branch. In general all combinations of signs are possible. In any case, when calculating the new multipliers, we will be able to check whether the stationarity type remains stable or not.

Concluding, the analysis of these singularities allow us to implement a pathfollowing algorithm for generic parametric MPCC problems. Indeed at a g.c. point we can locally follow $\Sigma_{gc}(P_{CC}(t))$ by means of well determined nonlinear optimization problems. We can also check numerically when it is possible to leave or enter the set of C-, B-, strongly stationary points.

Chapter 5

Bilevel problems

5.1 Introduction

Bilevel problems represent another important class of optimization problems whose structure is

$$\begin{aligned} P_{BL}: \qquad & \min_{x,y} f(x,y) \\ & \text{s.t. } (x,y) \in M_{BL} \end{aligned} \tag{5.1.1}$$

$$M_{BL} = \left\{ (x,y) \in \mathbb{R}^{n+m} \;\middle|\; \begin{array}{c} g_j(x,y) \geq 0,\ j = 1, \ldots, q, \\ y \text{ solves } Q(x). \end{array} \right\}$$

Here $x \in \mathbb{R}^n,\ y \in \mathbb{R}^m$ and $Q(x)$ stands for the parametric optimization problem

$$\begin{aligned} Q(x): \qquad & \min_{y} \phi(x,y) \\ & \text{s.t. } \ y \in Y(x) \\ & Y(x) = \{ y \in \mathbb{R}^m \mid v_i(x,y) \geq 0,\ i = 1, \ldots, l \} \end{aligned}$$

and is called lower level problem. As usual $(f, g_1, \ldots, g_q,) \in [C^2]^{1+q+l}_{n+m}$. We will require $(\phi, v_1, \ldots, v_l) \in [C^3]^1_{n+m}$.

This kind of problem appears in many applications such as Cournot equilibrium problems, see (1.4.2), or when solving semi-infinite programs, see [59]. However, finding a solution is not an easy problem, because even in order to check feasibility one must solve a nonlinear optimization problem $Q(x)$. Bilevel problems have been investigated in a large number of papers and books, see *e.g.*, [3], [42], [11] and the references therein.

We may try to solve P_{BL} via a reduction approach as described in [60]. The idea is, roughly, the following: if we denote the solution of $Q(x)$ by $y(x)$ and if we suppose that at least near a point $\overline{x}$ this function behaves well, *i.e.*, $y(x) \in C^1$, then locally near $\overline{x}$ the problem can be written equivalently as the standard program:

$$\begin{aligned} & \min f(x, y(x)) \\ & \text{s.t. } g_j(x, y(x)) \geq 0,\ j = 1, \ldots, q. \end{aligned}$$

However, this assumption is a very strong hypothesis as can be seen in the following example.

Example 5.1.1

$$\begin{array}{rl} & \min 2x_1 + x_2 + y \\ \text{s.t.} & x_1 \geq 0, \\ y \text{ solves } Q(x): & \begin{array}{rrcl} & \min y & & \\ \text{s.t.} & x_1 + y & \geq & 0, \\ & x_2 + y & \geq & 0. \end{array} \end{array}$$

The solution of the lower level problem is $y = -x_1$ if $x_1 \leq x_2$ and $y = -x_2$ in the other case. So the function $y(x)$ is not C^1 around $x = (0, 0)$ and the point $(0, 0, 0)$ is a global minimizer of the problem.

An alternative solution is via a KKT approach. The constraint y *solves* $Q(x)$, is replaced by the KKT condition:

$$\begin{array}{rcll} \nabla_y \phi(x, y) - \sum_{i=1}^{l} \lambda_i \nabla_y v_i(x, y) & = & 0, & \\ v_i(x, y) & \geq & 0, & i = 1, \ldots, l, \\ \lambda_i & \geq & 0, & i = 1, \ldots, l, \\ \lambda_i v_i(x, y) & = & 0, & i = 1, \ldots, l. \end{array} \tag{5.1.2}$$

So instead of P_{BL}, we solve the complementarity constrained problem

$$\begin{array}{rl} P_{\text{KKTBL}}: & \min\limits_{x,y,\lambda} f(x, y) \\ & \text{s.t. } (x, y, \lambda) \in M_{\text{KKTBL}}, \end{array} \tag{5.1.3}$$

where

$$M_{\text{KKTBL}} = \left\{ (x, y, \lambda) \in \mathbb{R}^{n+m+l} \;\middle|\; \begin{array}{rcll} \nabla_y \phi(x, y) - \sum_{i=1}^{l} \lambda_i \nabla_y v_i(x, y) & = & 0, & \\ v_i(x, y) & \geq & 0, & i = 1, \ldots, l, \\ \lambda_i & \geq & 0, & i = 1, \ldots, l, \\ \lambda_i v_i(x, y) & = & 0, & i = 1, \ldots, l, \\ g_j(x, y) & \geq & 0, & j = 1, \ldots, q. \end{array} \right\}$$

The FJ necessary condition can also be used. It will lead to a MPCC problem with a similar structure.

Due to the additional variable λ, the program (5.1.3) can be seen as a lifting procedure. On the other hand it is a relaxation of the original problem P_{BL} (5.1.1) in the sense that, if $(x, y) \in M_{BL}$ and MFCQ holds at $y \in Y(x)$, then there is some $\lambda \in \mathbb{R}^l$ such that $(x, y, \lambda) \in M_{\text{KKTBL}}$. Moreover, if (x, y, λ) is a local solution of problem (5.1.3) such that y is a minimizer of $Q(x)$ and MFCQ holds at $y \in Y(x)$, then (x, y) must also be a local solution of P_{BL}. In the case that $Q(x)$ is a convex program and MFCQ is satisfied for all $y \in Y(x)$, both problems (5.1.1) and (5.1.3) are equivalent.

Obviously, P_{KKTBL} is a special instance of a complementarity constrained problem. So, we can apply the solution method of Chapter 4, *i.e.*, we can try to solve the perturbed problem P_τ (see problem (4.5.1) in Section 4.6) using standard software of nonlinear programming. For general MPCC problems, this approach has a favorable convergence behavior under natural assumptions, see Chapter 4. However, since P_{KKTBL} is a MPCC with a special structure, these assumptions may not be generically fulfilled. In the sequel, we will study the properties of P_{KKTBL} from a generical point of view. We will present an algorithm that under these generical conditions will find a critical point of P_{BL}.

The following question arises. How is the structure of the KKT model compared with the structure of the original bilevel program, especially at solutions $(\overline{x}, \overline{y})$ where the reduction approach fails as in Example 5.1.1. We will see that the singular behavior of P_{BL} will partially reappear in the KKT formulation.

The chapter is organized as follows. In the next section we study the problem (5.1.3) as a complementarity constrained program and examine its structure from a generical viewpoint. This analysis will be used to obtain in Section 5.3 a numerical procedure for solving bilevel problems. We end with some numerical examples.

5.2 Genericity analysis of the KKT approach

We consider the KKT formulation (5.1.3) of the bilevel problem P_{BL} . Since it is a complementarity constrained program with a special structure, the genericity results of Chapter 4, developed for the general case, are no more valid. We have to perform a modified genericity analysis. It appears that also for P_{KKTBL}, generically MPCC-LICQ is satisfied at all feasible points. But for the local minimizers $(\overline{x}, \overline{y}, \overline{\lambda})$, the situation may be more complicated than for the general MPCC problems. We will see that the conditions MPCC-SC and MPCC-SOC may fail at local minimizers $(\overline{x}, \overline{y}, \overline{\lambda})$ of P_{KKTBL} where, for the corresponding minimizer $(\overline{x}, \overline{y})$ of P_{BL}, the lower level problem $Q(x)$ does not allow a local reduction as in Example 5.1.1.

With respect to a feasible point $(\overline{x}, \overline{y}, \overline{\lambda})$ of problem P_{KKTBL}, we introduce the following active index sets:

$$
\begin{array}{rcl}
J_{0v}(\overline{x},\overline{y}) & = & \{i \mid v_i(\overline{x},\overline{y}) = 0\}, \\
J_0(\overline{x},\overline{y},\overline{\lambda}) & = & \left\{i \mid v_i(\overline{x},\overline{y}) = 0,\ \overline{\lambda}_i > 0\right\}, \\
J\Lambda_0(\overline{x},\overline{y},\overline{\lambda}) & = & \left\{i \mid v_i(\overline{x},\overline{y}) = \overline{\lambda}_i = 0\right\}, \\
\Lambda_0(\overline{x},\overline{y},\overline{\lambda}) & = & \left\{i \mid v_i(\overline{x},\overline{y}) > 0,\ \overline{\lambda}_i = 0\right\}, \\
J_{0g}(\overline{x},\overline{y}) & = & \{j \mid g_j(\overline{x},\overline{y}) = 0\}.
\end{array}
\tag{5.2.1}
$$

Note that $J_{0v}(\overline{x},\overline{y}) = J_0(\overline{x},\overline{y},\overline{\lambda}) \cup J\Lambda_0(\overline{x},\overline{y},\overline{\lambda})$. This union does not depend on $\overline{\lambda}$.

Theorem 5.2.1 *Let $(\hat{\phi},\hat{v}_1,\ldots,\hat{v}_l) \in [C^3_S]^{1+l}_{n+m}$ and $(\hat{g}_1,\ldots,\hat{g}_q) \in [C^2_S]^q_{n+m}$ be fixed. For almost all $(C^x_\phi, C^y_\phi, d_\phi, C_v, d_v, C_g, d_g) \in \mathbb{R}^{mn+\frac{m(m+1)}{2}+m+l(n+m)+l+q(n+m)+q}$, the condition MPCC-LICQ holds at all feasible points of problem (5.1.3) defined by $\phi(x,y) = \hat{\phi}(x,y) + x^T[C^x_\phi]^T y + \frac{y^T[C^y_\phi]y}{2} + d^T_\phi y$, $v(x,y) = \hat{v}(x,y) + C_v(x,y) + d_v$ and $g(x,y) = \hat{g}(x,y) + C_g(x,y) + d_g$. Here $\mathbb{R}^{\frac{m(m+1)}{2}}$ denotes the space of symmetric matrices of order m.*
Moreover, generically in the set $\{(\phi,v,g)\} = [C^3_S]^{1+l}_{n+m} \times [C^2_S]^q_{n+m}$, the condition MPCC-LICQ holds at all feasible points of the set M_{KKTBL} defined by (ϕ,v,g).

Proof. Note that a similar result under stronger conditions has been proven in [52] for mathematical programs with variational inequality constraints, see problem (1.1.8).
The idea of the proof is the following. For any feasible point $(\overline{x},\overline{y},\overline{\lambda})$ of the problem given by (ϕ,v,g), there is a partition $J_0 = J_0(\overline{x},\overline{y},\overline{\lambda})$, $J\Lambda_0 = J\Lambda_0(\overline{x},\overline{y},\overline{\lambda})$, $\Lambda_0 = \Lambda_0(\overline{x},\overline{y},\overline{\lambda})$ of $\{1,\ldots,l\}$ and a set $J_{0g} = J_{0g}(\overline{x},\overline{y}) \subset \{1,\ldots,q\}$, such that $(\overline{x},\overline{y},\overline{\lambda})$ solves the system:

$$
\begin{array}{rcll}
\nabla_y\phi(x,y) - \sum\limits_{i=1}^{l} \lambda_i \nabla_y v_i(x,y) & = & 0, & \\
v_i(x,y) & = & 0, & i \in J_0 \cup J\Lambda_0, \\
\lambda_i & = & 0, & i \in J\Lambda_0 \cup \Lambda_0, \\
g_j(x,y) & = & 0, & j \in J_{0g}.
\end{array}
\tag{5.2.2}
$$

If MPCC-LICQ fails, then the gradients of the active constraints in (5.2.2) are linearly dependent, *i.e.*, there exists a non-zero vector $(\alpha,\beta,\mu,\gamma) \in \mathbb{R}^\kappa$,

$\kappa = m + |J_0 \cup J\Lambda_0| + |J_{0g}| + |J\Lambda_0 \cup \Lambda_0|$, such that:

$$
\begin{array}{lcl}
\begin{bmatrix} \nabla_x[\nabla_y \phi(x,y)]^T - \sum\limits_{i=1}^{l} \lambda_i \nabla_x [\nabla_y v_i(x,y)]^T \\ \nabla_y^2 \phi(x,y) - \sum\limits_{i=1}^{l} \lambda_i \nabla_y^2 v_i(x,y) \end{bmatrix} \alpha + & & \\
+ \sum\limits_{i \in J_0 \cup J\Lambda_0} \beta_i \nabla_{(x,y)} v_i(x,y) + \sum\limits_{j \in J_{0g}} \mu_j \nabla_{(x,y)} g_j(x,y) & = & 0, \\
\nabla_y v_i(x,y)^T \alpha & = & 0, \qquad i \in J_0, \\
\nabla_y v_i(x,y)^T \alpha - \gamma_i & = & 0, \quad i \in J\Lambda_0 \cup \Lambda_0.
\end{array}
\tag{5.2.3}
$$

If $\alpha = 0$, then $\gamma = 0$, and thus at (x, y) the gradients $\nabla_{(x,y)} g_{J_{0g}}$, $\nabla_{(x,y)} v_{J_0 \cup J\Lambda_0}$ are linearly dependent. This means that in this case LICQ fails in M^0_{KKTBL}, where

$$
M^0_{\text{KKTBL}} = \left\{ (x,y) \in \mathbb{R}^{n+m} \;\middle|\; \begin{array}{lcll} v_i(x,y) & \geq & 0, & i = 1, \ldots, l, \\ g_j(x,y) & \geq & 0, & j = 1, \ldots, q. \end{array} \right\}
$$

But as M^0_{KKTBL} is the feasible set of a common nonlinear program, it is known that, for almost every linear perturbation of $(\hat{v}_1, \ldots, \hat{v}_l, \hat{g}_1, \ldots, \hat{g}_q)$, the LICQ condition holds for all $(x, y) \in M^0_{\text{KKTBL}}$. This means that, for almost every (C_v, d_v, C_g, d_g) there is no feasible point of M_{KKTBL} solving the system (5.2.3) with $\alpha = 0$.

So, we only have to consider the remaining case $\alpha \neq 0$. Let us fix the set of active constraints $(J_0, J\Lambda_0, \Lambda_0)$. W.l.o.g. we assume $\alpha_1 \neq 0$ and take the corresponding combination in (5.2.3):

$$
(1, \alpha_2^0, \ldots \alpha_m^0, \beta^0, \mu^0, \gamma^0) = (\frac{\alpha}{\alpha_1}, \frac{\beta}{\alpha_1}, \frac{\mu}{\alpha_1}, \frac{\gamma}{\alpha_1}).
$$

Note that $(\alpha_2^0, \ldots, \alpha_m^0, \beta^0, \mu^0, \gamma^0) \in \mathbb{R}^{\kappa-1}$. Then, if MPCC-LICQ fails for $\phi = \hat{\phi}(x,y) + x^T [C_\phi^x]^T y + \frac{y^T [C_\phi^y] y}{2} + d_\phi^T y$, $v = \hat{v}(x,y) + C_v(x,y) + d_v$ and $g = \hat{g}(x,y) + C_g(x,y) + d_g$, the condition (5.2.3) for the perturbed problem now reads:

$$
\begin{aligned}
\left[\nabla_{(x,y)}[\nabla_y\phi(x,y)]^T - \sum_{i=1}^{l}\lambda_i\nabla_{(x,y)}[\nabla_y v_i(x,y)]^T\right]\begin{pmatrix}1\\ \alpha_2^0\\ \vdots\\ \alpha_m^0\end{pmatrix} + \qquad & \\
\sum_{i\in J_0\cup J\Lambda_0}\beta_i^0\nabla_{(x,y)}[v_i(x,y)] + \sum_{j\in J_{0g}}\mu_j^0\nabla_{(x,y)}g_j(x,y) &= 0,\\
\nabla_y[v_i(x,y)]^T\begin{pmatrix}1\\ \alpha_2^0\\ \vdots\\ \alpha_m^0\end{pmatrix} &= 0, \quad i\in J_0,\\
\nabla_y[v_i(x,y)]^T\begin{pmatrix}1\\ \alpha_2^0\\ \vdots\\ \alpha_m^0\end{pmatrix} - \gamma_i^0 &= 0, \quad i\in J\Lambda_0\cup\Lambda_0,\\
\nabla_y\phi(x,y) - \sum_{i=1}^{l}\lambda_i\nabla_y v_i(x,y) &= 0,\\
v_i(x,y) &= 0, \quad i\in J_0\cup J\Lambda_0,\\
\lambda_i &= 0, \quad i\in\Lambda_0\cup J\Lambda_0,\\
g_j(x,y) &= 0, \quad j\in J_{0g}.
\end{aligned}
\tag{5.2.4}
$$

The Jacobian matrix of the system (5.2.4) with respect to the variables $(x, y, \lambda, \alpha_2^0, \dots, \alpha_m^0, \beta^0, \mu^0, \gamma^0)$ and the parameters $(C_\phi^x, C_\phi^y, d_\phi, C_v, d_v, C_g, d_g)$, is:

$\partial_{(x,y)}$	∂_λ	$\partial_{\alpha_2^0,\dots,\alpha_m^0,\beta^0,\mu^0}$	∂_{γ^0}	$\partial_{C_\phi^x,C_\phi^y}$	∂_{d_ϕ}	∂_{C_v}	∂_{d_v}	∂_{d_g}
$\otimes$	$\otimes$	$\otimes$	0	Ω_ϕ	0	$\otimes$	0	0
$\otimes$	0	$\otimes$	0	0	0	Ω_v	0	0
$\otimes$	0	$\otimes$	$0\|I_{\|J\Lambda_0\cup\Lambda_0\|}$	0	0	$\otimes$	0	0
$\otimes$	$\otimes$	0	0	$\otimes$	I_m	$\otimes$	0	0
$\otimes$	0	0	0	0	0	$\otimes$	$I_{\|J_0\cup J\Lambda_0\|}\|0$	0
0	$0\|I_{\|J\Lambda_0\cup\Lambda_0\|}$	0	0	0	0	0	0	0
$\otimes$	0	0	0	0	0	0	0	$I_{\|J_{0g}\|}\|0$

where $\Omega_v = \begin{pmatrix}(0,\dots,0,1,\alpha_2^0,\dots,\alpha_m^0) & \dots & 0\\ 0 & \ddots & 0\\ 0 & \dots & (0,\dots,0,1,\alpha_2^0,\dots,\alpha_m^0)\end{pmatrix}$ has $|J_0|$ rows and $\Omega_\phi = \begin{pmatrix}I_n|\otimes & \multicolumn{3}{c}{0}\\ \multirow{2}{*}{0} & 1 & \otimes & \otimes\\ & 0 & I_{m-1} & \otimes\end{pmatrix}$.

It can be seen that the Jacobian matrix has full row rank. By the Parameterized Sard Lemma, see Lemma 2.3.1, for almost every $(C_\phi^x, C_\phi^y, d_\phi, C_v, d_v, C_g, d_g)$,

the matrix formed by the columns of the Jacobian matrix corresponding to the variables $(x, y, \lambda, \alpha_2^0, \ldots, \alpha_m^0, \beta^0, \mu^0, \gamma^0)$ has full row rank. But the number of rows, $n + m + l + m + |J_0 \cup J\Lambda_0| + |J\Lambda_0 \cup \Lambda_0| + |J_{0g}|$, is larger than the number $n+m+l+m-1+|J_0 \cup J\Lambda_0|+|J_{0g}|+|J\Lambda_0 \cup \Lambda_0|$ of columns. This means that for almost every parameter, there is no solution of system (5.2.4), *i.e.*, MPCC-LICQ holds.
If we consider any possible combination of active index sets $J_0, J\Lambda_0, \Lambda_0$ and J_{0g} an analogous result is obtained. The perturbation result follows by intersecting the set of parameters where *MPCC-LICQ* holds at the solutions of system (5.2.4), for all possible combinations of active index sets.

We now turn to the proof of the genericity statement. First we will prove that, for fixed $N \in \mathbb{N}$, the set of functions (ϕ, v, g) where MPCC-LICQ holds at all feasible points (x, y, λ) of M_{KKTBL} with $\|\lambda\| \leq N$, is open and dense in $[C_S^3]^{1+l}_{n+m} \times [C_S^2]^q_{n+m}$.

The density part of the statement follows, as usual, directly from the perturbation result shown above, by using partitions of unity. For more details see [52].

Now we will prove the openness property. We proceed as in the proof of Proposition 4 in [53]. Consider $(\hat{\phi}, \hat{v}_1, \ldots, \hat{v}_l, \hat{g}_1, \ldots, \hat{g}_q)$ such that, for the resulting set M_{KKTBL}, MPCC-LICQ holds at all feasible points $(\overline{x}, \overline{y}, \overline{\lambda})$, with $\|\overline{\lambda}\| \leq N$. Let us fix $(\overline{x}, \overline{y}) \in \mathbb{R}^n \times \mathbb{R}^m$. By continuity, it can be seen that, if $(\overline{x}, \overline{y}, \lambda) \notin M_{\text{KKTBL}}$, then for all $\lambda \in \overline{B}^l_N(0)$ there is a neighborhood $V_{(\overline{x},\overline{y})}$ of $(\hat{\phi}, \hat{v}, \hat{g})$ and U of $(\overline{x}, \overline{y})$ such that for all $(\phi, v, g) \in V_{(\overline{x},\overline{y})}$

$$\left[U \times \overline{B}^l_N(0)\right] \cap M_{\text{KKTBL}} = \varnothing$$

holds for the feasible set M_{KKTBL} corresponding to (ϕ, v, g).

Let $(\overline{x}, \overline{y})$ be a point such that $(\overline{x}, \overline{y}, \overline{\lambda}) \in M_{\text{KKTBL}}$, for some $\overline{\lambda} \in \overline{B}^l_N(0)$. By assumption, at this point MPCC-LICQ is valid. Using a continuity argument, we can prove that there are neighborhoods U and $V_{(\overline{x},\overline{y})}$ of $(\overline{x}, \overline{y})$ and $(\hat{\phi}, \hat{v}, \hat{g})$, respectively, such that, for all $(x, y, \lambda) \in \left[U \times \overline{B}^l_N(0)\right] \cap M_{\text{KKTBL}}$, the MPCC-LICQ condition holds at the feasible set M_{KKTBL} corresponding to $(\phi, v, g) \in V_{(\overline{x},\overline{y})}$.

Using the neighborhoods $V_{(\overline{x},\overline{y})}$ and a locally finite cover w.r.t. the (x, y)-space, we can construct a neighborhood $\hat{V}$ of $(\hat{\phi}, \hat{v}, \hat{g})$ such that MPCC-LICQ holds at all feasible points (x, y, λ) of problem (5.1.3), defined by $(\phi, v, g) \in \hat{V}$ and with $\|\lambda\| \leq N$.

Altogether we have proven that the set of functions $(\hat{\phi}, \hat{v}, \hat{g})$ such that the MPCC-LICQ condition is satisfied, at all feasible points (x, y, λ) of M_{KKTBL} with $\|\lambda\| \leq N$, is open and dense.

Finally, if we intersect these open and dense sets for $N = 1, 2, \ldots$, we will obtain the generic set of functions where MPCC-LICQ holds at all feasible points. $\square$

We now study the structure of the critical points of problem (5.1.3) in the generic case. The critical points (x, y, λ) are feasible points with associated multipliers $(\alpha, \beta, \mu, \gamma)$ such that $(x, y, \lambda, \alpha, \beta, \mu, \gamma)$ solves the system:

$$\nabla_{(x,y)} f(x, y) - \sum_{i \in J_0 \cup J\Lambda_0} \beta_i \nabla_{(x,y)} v_i(x, y) - \sum_{j \in J_{0g}} \mu_j \nabla_{(x,y)} g_j(x, y) - \\ - \begin{bmatrix} \nabla_x [\nabla_y \phi(x, y)]^T - \sum_{i=1}^{l} \lambda_i \nabla_x [\nabla_y v_i(x, y)]^T \\ \nabla_y^2 \phi(x, y) - \sum_{i=1}^{l} \lambda_i \nabla_y^2 v_i(x, y) \end{bmatrix} \alpha = 0, \tag{5.2.5}$$

$$\nabla_y v_i(x, y)^T \alpha = 0, \quad i \in J_0,$$
$$\nabla_y v_i(x, y)^T \alpha - \gamma_i = 0, \quad i \in J\Lambda_0 \cup \Lambda_0$$

with $J_0, J\Lambda_0, \Lambda_0, J_{0g}$, the active index sets at (x, y, λ), as defined in (5.2.1).

Note that, as generically MPCC-LICQ holds, generically any solution of (5.1.3) must satisfy the KKT conditions (5.2.5). For numerical purposes, it is desirable that the conditions MPCC-SC and MPCC-SOC hold at critical points. However, as we shall show, MPCC-SC may fail generically. Let us consider the problem:

$$\begin{array}{ll} & \min -x - y \\ \text{s.t. } y \text{ solves } Q(x): & \begin{array}{lrcl} & & \min y & \\ \text{s.t.} & -x + y & \geq & 0, \\ & -y & \geq & 0. \end{array} \end{array} \tag{5.2.6}$$

It can easily be seen that in this problem, MFCQ fails at the solution $\overline{y} = 0$ of the lower level problem $Q(0)$. The point $(\overline{x}, \overline{y}) = (0, 0)$ is the minimizer of the bilevel problem and at $\overline{y}$ the condition MFCQ fails for $Q(\overline{x})$. In the Figure 5.1 the feasible set and the solution is depicted.

The KKT approach leads to the program:

$$\begin{array}{lrcl} & \min -x - y & & \\ \text{s.t.} & -x + y & \geq & 0, \\ & -y & \geq & 0, \\ & 1 - \lambda_1 + \lambda_2 & = & 0, \\ & \lambda_1, \lambda_2 & \geq & 0, \\ & (-x + y)\lambda_1 & = & 0, \\ & -y\lambda_2 & = & 0. \end{array} \tag{5.2.7}$$

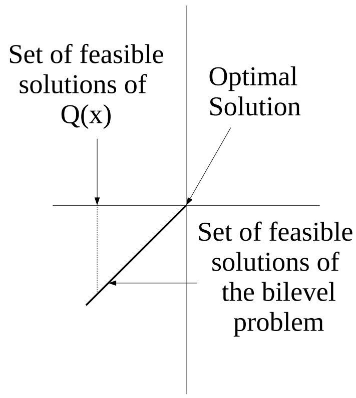


Figure 5.1: Feasible set of bilevel program (5.2.6)

and $(x, y, \lambda_1, \lambda_2) = (0, 0, \lambda_1, \lambda_1 - 1),\ \lambda_1 \geq 1$ are global minimizers. If we take the point $z := (0, 0, 1, 0)$, the critical point condition for the MPCC problem (5.2.7) is:

$$\begin{pmatrix} -1 \\ -1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \beta_1 + \begin{pmatrix} 0 \\ -1 \\ 0 \\ 0 \end{pmatrix} \beta_2 + \begin{pmatrix} 0 \\ 0 \\ -1 \\ 1 \end{pmatrix} \alpha + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \gamma$$

So the multipliers are $\alpha = \gamma = 0$ and $\beta_1 = 1,\ \beta_2 = 2$. As $J\Lambda_0(z) = \{2\}$, *MPCC-SC* fails.

In case we take $z = (0, 0, \lambda_1, \lambda_1 - 1),\ \lambda_1 > 1$, although MPCC-SC holds, since $J\Lambda_0(z) = \varnothing$, the condition MPCC-SOC will fail. Note that $T_z M_{\text{KKTBL}}$ is generated by the vector $(0, 0, 1, 1)$ while $\nabla^2_{(x,y,\lambda)} L(0, 0, \lambda_1, 1 - \lambda_1, 1, 2, 0, 0) = 0$, so, $\nabla^2_{(x,y,\lambda)} L|_{T_z M_{\text{KKTBL}}}(0, 0, \lambda_1, 1 - \lambda_1, 1, 2, 0, 0)$ is a singular matrix.

As can be seen in Figure 5.1, if we slightly perturb the functions describing the constraints as $-x + y + \epsilon_3(x, y)$ and $-y + \epsilon_4(x, y)$, there will be a point (x^*, y^*) satisfying

$$\begin{aligned} -x + y + \epsilon_3(x, y) &= 0, \\ -y + \epsilon_4(x, y) &= 0. \end{aligned}$$

Now, if we perturb the objective function of the lower level problem as $y + \epsilon_2(x, y)$, the lower level multipliers λ_1, λ_2 should fulfill the KKT condition for problem $Q(x^*)$:

$$1 + \frac{\partial \epsilon_2}{\partial y}(x^*, y^*) - \lambda_1(1 + \frac{\partial \epsilon_3}{\partial y}(x^*, y^*)) + \lambda_2(1 - \frac{\partial \epsilon_4}{\partial y}(x^*, y^*)) = 0.$$

So in particular $(\lambda_1^*, \lambda_2^*) = (\frac{1+\frac{\partial \epsilon_2}{\partial y}(x,y)}{1+\frac{\partial \epsilon_3}{\partial y}(x^*,y^*)}, 0)$ solves the previous equation. This means that the point $(x^*, y^*, \lambda_1^*, 0)$ is a feasible point of M_{KKTBL} for the perturbed functions. Now we turn to the objective function of the upper level problem. Assume it is perturbed as $-x - y + \epsilon_1(x, y)$ with $|\epsilon_1(x, y)| \ll 1$, $|\nabla\epsilon_1(x, y)| \ll 1$ and $|\nabla^2\epsilon_1(x, y)| \ll 1$, for all $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$. Then, the critical point system for $(x, y, \lambda) = (x^*, y^*, \lambda_1^*, 0),\ \alpha = \gamma = 0$ reads:

$$\begin{pmatrix} -1 + \frac{\partial \epsilon_1}{\partial x}(x^*, y^*) \\ -1 + \frac{\partial \epsilon_1}{\partial y}(x^*, y^*) \end{pmatrix} = \begin{pmatrix} -1 + \frac{\partial \epsilon_3}{\partial x}(x^*, y^*) \\ 1 + \frac{\partial \epsilon_3}{\partial y}(x^*, y^*) \end{pmatrix} \beta_1 + \begin{pmatrix} \frac{\partial \epsilon_4}{\partial x}(x^*, y^*) \\ -1 + \frac{\partial \epsilon_4}{\partial y}(x^*, y^*) \end{pmatrix} \beta_2.$$

As $|\nabla\epsilon_i| \ll 1,\ i = 3, 4$, the matrix $\begin{pmatrix} -1 + \frac{\partial \epsilon_3}{\partial x}(x^*, y^*) & \frac{\partial \epsilon_4}{\partial x}(x^*, y^*) \\ 1 + \frac{\partial \epsilon_3}{\partial y}(x^*, y^*) & -1 + \frac{\partial \epsilon_4}{\partial y}(x^*, y^*) \end{pmatrix}$, is non-singular. From these facts, it follows that the point $(x^*, y^*, \lambda_1^*, 0)$ is a critical point of this perturbed problem with corresponding multipliers (α, β, γ) and $\alpha = \gamma = 0$.

As a consequence we have shown that the failure of MPCC-SC at a solution of P_{KKTBL}, may remain stable under small perturbations.

The next result describes the generic properties of the critical points:

Theorem 5.2.2 *Given $(\hat{f}, \hat{\phi}, \hat{v}_1, \ldots, \hat{v}_l, \hat{g}_1, \ldots, \hat{g}_q)$, let us consider the perturbed functions $f = \hat{f} + b^T(x, y)$, $\phi = \hat{\phi}(x, y) + c_\phi^T y$, $v_i = \hat{v}_i + c_{v_i}^T y + d_{v_i}$, $g_j = \hat{g}_j + d_{g_j}$. Then for almost every $(b, c_\phi, c_{v_1}, \ldots, c_{v_l}, d_v, d_g)$, at all solutions $(\overline{x}, \overline{y}, \overline{\lambda}, \alpha, \beta, \gamma, \mu)$ of the corresponding system (5.2.5), i.e. $(\overline{x}, \overline{y}, \overline{\lambda})$ is a critical point of problem (5.1.3), the following holds:*

BL-1: If $\alpha \neq 0$, then the MPCC-LICQ, MPCC-SC and MPCC-SOC conditions are fulfilled w.r.t. the corresponding complementarity constrained problem (5.1.3).

BL-2: If $\alpha = 0$, then $rank(\nabla_y v_{J_{0v}(\overline{x},\overline{y})}(\overline{x}, \overline{y})) = m$ and the multipliers (μ, β) associated with $g_j(\overline{x}, \overline{y}), v_i(\overline{x}, \overline{y}),\ j \in J_{0g}(\overline{x}, \overline{y}),\ i \in J_{0v}(\overline{x}, \overline{y})$, are not equal to zero. Moreover there is some λ^* such that $(\overline{x}, \overline{y}, \lambda^*)$ is a critical point of problem P_{KKTBL} and $rank(\nabla_y v_{J_0(\overline{x},\overline{y},\lambda^*)}(\overline{x}, \overline{y})) =| J_0(\overline{x}, \overline{y}, \lambda^*) |= m$. For any λ^* such that the previous condition holds, the MPCC-SOC is satisfied at $(\overline{x}, \overline{y}, \lambda^*)$.

Proof. At a critical point $(\overline{x}, \overline{y}, \overline{\lambda})$ of the problem P_{KKTBL}, the following system has a solution:

$$\begin{aligned}
\nabla_{(x,y)} f(x,y) - \sum_{i \in J_0 \cup J\Lambda_0} \beta_i \nabla_{(x,y)} v_i(x,y) - \\
- \sum_{j \in J_{0g}} \mu_j \nabla_{(x,y)} g_j(x,y) - \\
- \begin{bmatrix} \nabla_x \nabla_y \phi(x,y) - \sum_{i=1}^{l} \lambda_i \nabla_x \nabla_y v_i(x,y) \\ \nabla_y^2 \phi(x,y) - \sum_{i=1}^{l} \lambda_i \nabla_y^2 v_i(x,y) \end{bmatrix} \alpha &= 0, \\
\nabla_y^T v_i(x,y) \alpha &= 0, \quad i \in J_0, \\
\nabla_y^T v_i(x,y) \alpha - \gamma_i &= 0, \quad i \in J\Lambda_0 \cup \Lambda_0, \\
\nabla_y \phi(x,y) - \sum_{i=1}^{l} \lambda_i \nabla_y v_i(x,y) &= 0, \\
v_i(x,y) &= 0, \quad i \in J_0 \cup J\Lambda_0, \\
\lambda_i &= 0, \quad J\Lambda_0 \cup \Lambda_0, \\
\\
g_j(x,y) &= 0, \quad j \in J_{0g}, \\
\beta_i &= 0, \quad i \in J\Lambda^*_{0\beta} \subset J\Lambda_0, \\
\gamma_i &= 0, \quad i \in J\Lambda^*_{0\gamma} \subset J\Lambda_0, \\
\mu_j &= 0, \quad j \in J^*_{0g} \subset J_{0g},
\end{aligned} \tag{5.2.8}$$

where some of the multipliers $\beta_i, \gamma_i, \mu_j,\ i \in J\Lambda_0,\ j \in J_{0g}$ are equal to zero. The corresponding index sets are denoted by $J\Lambda^*_{0\beta},\ J\Lambda^*_{0\gamma},\ J\Lambda^*_{0g}$. With this setting, $J^*_{0g} = J\Lambda^*_{0\beta} = J\Lambda^*_{0\gamma} = \varnothing$ means that MPCC-SC holds. For simplicity, we skipped the arguments (x, y, λ) in the active index sets.

Now, we will consider solutions $(x, y, \lambda, \alpha, \beta, \gamma, \mu)$ of the previous system for perturbed functions $f = \hat{f}(x,y) + b^T(x,y), \phi = \hat{\phi}(x,y) + c_\phi^T y, v_i = \hat{v}_i(x,y) + c_{v_i}^T y + d_{v_i},\ g_j = \hat{g}_j(x,y) + d_{g_j}$. We have to distinguish between two cases corresponding to $\alpha = 0$ and $\alpha \neq 0$.

Case $\alpha \neq 0$: We assume $\|\alpha\| > \frac{1}{N}$, for some fixed $N \in \mathbb{N}$. For simplicity we will consider $J_0 = \{1, 2, \ldots, l_1\},\ J\Lambda_0 = \{l_1 + 1, \ldots, l_2\},\ \Lambda_0 = \{l_2 + 1, \ldots, l\}$, $J_{0g} = \{1, 2, \ldots, q_1\},\ J\Lambda^*_{0\beta} = \{l_1 + 1 \ldots, l_3\},\ J\Lambda^*_{0\gamma} = \{l_4, \ldots, l_5\}$ and $J^*_{0g} = \{1, \ldots, q_2\}$, for some $0 \leq l_1 \leq l_3 \leq l_2 \leq l$ and $l_1 \leq l_4 \leq l_5 \leq l_2,\ q_2 \leq q_1$. Taking the Jacobian with respect to the variables and parameters, we find

$$
\begin{array}{ccccccccccc}
\partial x & \partial y & \partial \lambda & \partial \alpha & \partial \beta & \partial \gamma & \partial \mu & \partial b & \partial c_\phi^T & \partial c_v^T & \partial d_{v_{J_{0v}}} d_{g_{J_{0g}}} \\
\hline
\otimes & \otimes & \otimes & \otimes & \otimes & 0 & \otimes & I_{n+m} & 0 & \otimes & 0 \\
\otimes & \otimes & 0 & \begin{array}{c} [\nabla_y v_1]^T \\ \vdots \\ [\nabla_y v_l]^T \end{array} & 0 & \begin{array}{c} 0 \\ I_{l-l_1} \end{array} & 0 & 0 & 0 & \Omega & 0 \\
\otimes & \otimes & \otimes & 0 & 0 & 0 & 0 & 0 & I_m & \otimes & 0 \\
\otimes & \otimes & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \otimes & I_{l_2}|0|0|0 \\
0 & 0 & 0|I_{l-l_1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\otimes & \otimes & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0|0|I_{q_1}|0 \\
0 & 0 & 0 & 0 & 0|I_{\beta^*}|0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0|I_{\gamma^*}|0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & I_{\mu^*}|0 & 0 & 0 & 0 & 0
\end{array}
\tag{5.2.9}
$$

where $\Omega = \begin{pmatrix} \alpha^T & 0 & 0 \\ & \ddots & \\ 0 & 0 & \alpha^T \end{pmatrix}$ has l rows and $\beta^* = l_3 - l_1, \mu^* = q_2$, $\gamma^* = l_2 - l_4$.

Evidently, $\|\alpha\| > \frac{1}{N}$ implies $\alpha \neq 0$ and Ω has rank l. So, it follows directly that the rows of the matrix (5.2.9) are linearly independent. Then, by the Parameterized Sard Lemma, for almost every b, c_ϕ, d_v, d_g, c_v, the Jacobian matrix of the system (5.2.8), with respect to variables (x, y, λ) and multipliers $(\alpha, \beta, \gamma, \mu)$, has rank $n+m+l+m+|J_{0v}|+|J\Lambda_0|+|\Lambda_0|+|J_{0g}|+|\beta^*|+|\mu^*|+|\gamma^*|$, equal to the number of rows, at all solutions of the system with $\|\alpha\| > \frac{1}{N}$. But this rank cannot be greater than the number $n+m+l+m+|J_0|+|J\Lambda_0|+|J\Lambda_0|+|\Lambda_0|+|J_{0g}|+|\beta^*|+|\mu^*|+|\gamma^*|$ of involved variables. So, in view of $J_0 \cup J\Lambda_0 = J_{0v}$:

$$|\beta^*| + |\mu^*| + |\gamma^*| = 0.$$

i.e. MPCC-SC holds.
Now, we want to prove the regularity of $\mathcal{C} = \nabla^2_{(x,y,\lambda)} L|_{T_{(x,y,\lambda)} M_{\mathrm{KKTBL}}}$, the Hessian matrix at the solutions of the perturbed system (5.2.8) with $\|\alpha\| > \frac{1}{N}$. In view of Proposition 2.1.1 it is equivalent to prove the regularity of:

$$\mathcal{B} = \begin{pmatrix} \mathcal{A} & B \\ B^T & 0 \end{pmatrix}, \tag{5.2.10}$$

where $\mathcal{A} = \nabla^2_{(x,y,\lambda)} L(x, y, \lambda, \alpha, \beta, \gamma, \mu)$ and B is the matrix with the gradients of the active constraints as columns.

Note that $\mathcal{B}$ is the upper left part of the matrix (5.2.9), formed by the rows 1 to 6, and the columns corresponding to the derivatives $\partial_{(x,y,\lambda,\alpha,\beta,\gamma,\mu)}$. So it is non-singular. It is well known that the regularity of $\mathcal{B}$ is equivalent to the fulfillment of *MPCC-LICQ* and *MPCC-SOC*.

Taking all finitely many possible combinations of active index sets, it follows that, for almost every linear perturbation of (f, ϕ, v, g), the solutions of the system (5.2.5) with $\|\alpha\| > \frac{1}{N}$ are non-degenerate critical points.

By considering the intersection $\cap_{N\in\mathbb{N}}$ of all these sets of perturbations, we obtain for almost every linear perturbation the non-degeneracy condition at all critical points with $\alpha \neq 0$.

<u>Case $\alpha = 0$</u>: Now we will consider the case of solutions with $\alpha = 0$. This implies $\gamma_i = 0,\ i \in J\Lambda_0 \cup \Lambda_0$, see system (5.2.8). Then for some $(\beta, \mu) \in \mathbb{R}^{|J_{0v}|+|J_{0g}|}$, the following system has a solution $(\overline{x}, \overline{y}, \overline{\lambda})$:

$$\begin{array}{rcll} \nabla_{(x,y)} f(x,y) - \sum\limits_{i\in J_{0v}} \beta_i \nabla_{(x,y)} v_i(x,y) - \sum\limits_{j\in J_{0g}} \mu_j \nabla_{(x,y)} g_j(x,y) & = & 0, & \\ \nabla_y \phi(x,y) - \sum\limits_{i=1}^{l} \lambda_i \nabla_y v_i(x,y) & = & 0, & \\ v_i(x,y) & = & 0, & i \in J_0 \cup J\Lambda_0, \\ \lambda_i & = & 0, & i \in J\Lambda_0 \cup \Lambda_0, \\ g_j(x,y) & = & 0, & j \in J_{0g}. \end{array} \tag{5.2.11}$$

As the set $J_{0v}(\overline{x}, \overline{y}) = J_0(\overline{x}, \overline{y}, \overline{\lambda}) \cup J\Lambda_0(\overline{x}, \overline{y}, \overline{\lambda})$ does not depend on the particular choice of λ, this system can be decomposed into two parts. Firstly the system in (x, y, β, μ):

$$\begin{array}{rcll} \nabla_{(x,y)} f(x,y) - \sum\limits_{i\in J_{0v}} \beta_i \nabla_{(x,y)} v_i(x,y) - \sum\limits_{j\in J_{0g}} \mu_j \nabla_{(x,y)} g_j(x,y) & = & 0, & \\ v_i(x,y) & = & 0, & i \in J_0 \cup J\Lambda_0, \\ g_j(x,y) & = & 0, & j \in J_{0g}. \end{array} \tag{5.2.12}$$

In the second part, for a given solution $(\overline{x}, \overline{y})$ of system (5.2.12), the vector λ has to solve the following linear system of equalities and inequalities.

$$\begin{array}{rcl} \phi_y(\overline{x}, \overline{y}) - \sum\limits_{i\in J_{0v}(\overline{x},\overline{y})} \lambda_i \nabla_y v_i(\overline{x}, \overline{y}) & = & 0, \\ \lambda & \geq & 0. \end{array} \tag{5.2.13}$$

Note that, for any solution $(\overline{x}, \overline{y}, \beta, \mu)$ of the system (5.2.12), the vector $(\overline{x}, \overline{y})$ is a critical point of the standard problem:

$$\begin{array}{rrcl} & \min f(x,y) & & \\ \text{s.t.} & v_i(x,y) & \geq & 0, \quad i = 1, \ldots, l, \\ & g_j(x,y) & \geq & 0, \quad j = 1, \ldots, q. \end{array} \tag{5.2.14}$$

with corresponding multipliers β and μ.

By Caratheodory's theorem, if the system (5.2.13) is solvable for $(\overline{x}, \overline{y})$, we can choose a solution λ^*, such that

$$\text{the vectors } \nabla_y v_{J_0(\overline{x},\overline{y},\lambda^*)}(\overline{x},\overline{y}) \text{ are linearly independent.} \tag{5.2.15}$$

So, in the system (5.2.11) we can assume, w.l.o.g., that the vectors $\nabla_y v_i$, $i \in J_0(\overline{x},\overline{y},\overline{\lambda})$ are linearly independent. Note that $J_0(\overline{x},\overline{y},\overline{\lambda})$ may be the empty set.

We turn back to the system (5.2.11), where f, ϕ, g, v are linearly perturbed as in the first part of the proof. We will assume (as in system (5.2.8)) that some multipliers are zero, *e.g.*, $\beta_i = 0, \mu_j = 0,\ i \in J\Lambda^*_{0\beta},\ j \in J^*_{0g}$, with $|J\Lambda^*_{0\beta}| = \beta^*$ and $|J^*_{0g}| = \mu^*$. So, the critical points will be described by means of the system (5.2.11) and the previous equations for the multipliers (β_i, μ_j). The Jacobian of the (augmented) system with respect to variables, multipliers and perturbation parameters has now the form (see also (5.2.9) for the case $\alpha \neq 0$):

$$\begin{array}{ccccccccc}
\partial x & \partial y & \partial\lambda & \partial\beta & \partial\mu & \partial b & \partial c_\phi^T & \partial d_{v_i} & \partial d_{g_i} \\
\hline
\otimes & \otimes & 0 & \otimes & \otimes & I_{n+m} & 0 & 0 & 0 \\
\otimes & \otimes & \otimes & 0 & 0 & 0 & I_m & 0 & 0 \\
\otimes & \otimes & 0 & 0 & 0 & 0 & 0 & I_{|J_0 \cup J\Lambda_0|}|0 & 0 \\
0 & 0 & 0|I_{|J\Lambda_0 \cup \Lambda_0|} & 0 & 0 & 0 & 0 & 0 & 0 \\
\otimes & \otimes & 0 & 0 & 0 & 0 & 0 & 0 & I_{|J_{0g}|}|0 \\
0 & 0 & 0 & \begin{array}{cccc} I_{l_4} & 0 & 0 & 0 \\ 0 & 0 & I_{l_3-l_1} & 0 \end{array} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & I_{\mu^*} & 0 & 0 & 0 & 0
\end{array} \tag{5.2.16}$$

Again we assumed $J_0 = \{1, 2, \ldots, l_1\},\ J\Lambda_0 = \{l_1+1, \ldots, l_2\},\ \Lambda_0 = \{l_2+1, \ldots, l\}$, $J_{0g} = \{1, 2, \ldots, q_1\},\ J\Lambda^*_{0\beta} = \{1, \ldots, l_4\} \cup \{l_1+1 \ldots, l_3\}$, and $J^*_{0g} = \{l_1+1, \ldots, q_2\}$, for some $l_4 \le l_1 \le l_3 \le l_2 \le l,\ q_2 \le q_1$. Using our standard argument, we conclude that, for almost every perturbation (b, c_ϕ, d_v, d_g), the Jacobian matrix with respect to the variables and multipliers $(x, y, \lambda, \beta, \mu)$ has full row rank. So the number of variables is greater than or equal to the number of equations, *i.e.*:

$$n+m+l+|J_0 \cup J\Lambda_0|+|J_{0g}| \ge n+m+m+|J_0 \cup J\Lambda_0|+|J\Lambda_0 \cup \Lambda_0|+|J_{0g}|+|\beta^*|+|\mu^*|$$

or, equivalently,

$$|J_0| \ge m + |\beta^*| + |\mu^*|.$$

In particular $|J_0| \ge m$, but our assumption that $\nabla_y v_i,\ i \in J_0$, are linear independent, see (5.2.15), yields $|J_0| \le m$. So,

$$|J_0| = m \text{ and } |\beta^*| = |\mu^*| = 0. \tag{5.2.17}$$

To prove the statement concerning *MPCC-SOC* for the case $\alpha = 0$, see *BL*-2 in Theorem 5.2.2, we reconsider the original MPCC-problem (5.1.3) and show that the matrix $\mathcal{B} = \begin{pmatrix} \mathcal{A} & B \\ B^T & 0 \end{pmatrix}$ is regular, where $\mathcal{A} = \nabla^2_{(x,y,\lambda)} L(x, y, \lambda, \alpha, \beta, \gamma, \mu)$ and the columns of B are given by the gradients of the active constraints. From the rank condition for matrix (5.2.16) it follows that the square matrix, composed by the row blocks $1, \ldots, 5$, and the columns corresponding to $\partial_{(x,y,\lambda,\beta,\mu)}$, is non-singular. We will denote this matrix by:

$$\mathcal{A}' = \begin{pmatrix} \otimes & \otimes & 0 & \otimes & \otimes \\ \otimes & \otimes & \otimes & 0 & 0 \\ \otimes & \otimes & 0 & 0 & 0 \\ 0 & 0 & 0|I_{|J\Lambda_0 \cup \Lambda_0|} & 0 & 0 \\ \otimes & \otimes & 0 & 0 & 0 \end{pmatrix}.$$

Note that $\mathcal{B}$ can easily be constructed from $\mathcal{A}'$ as follows: we add the partial derivatives corresponding to the equations $\nabla_y v_{J_0} \alpha = 0$, $\nabla_y v_{J\Lambda_0 \cup \Lambda_0} \alpha = \gamma_{J\Lambda_0 \cup \Lambda_0}$, as new rows and the derivatives w.r.t. α and γ as new columns. Exchanging some rows and columns we have:

$$\mathcal{B} = \mathcal{I}_1 \begin{pmatrix} \mathcal{A}' & \otimes & 0 \\ 0 & \nabla_y v_{J_0} & 0 \\ 0 & \nabla_y v_{J\Lambda_0 \cup \Lambda_0} & I_{|J\Lambda_0 \cup \Lambda_0|} \end{pmatrix} \mathcal{I}_2,$$

where $\mathcal{I}_1$ and $\mathcal{I}_2$ are appropriate (non-singular) permutations matrices. As $\nabla_y v_{J_0}$ is a non-singular matrix, see (5.2.15) and (5.2.17), the regularity of $\mathcal{A}'$ and $\mathcal{B}$ are equivalent. This means that, the *MPCC-SOC* condition holds.

□

From the proof of Theorem 5.2.1 and Theorem 5.2.2 we also obtain the following fact.

Corollary 5.2.1 *For almost all perturbations of $(f, \phi, v_1, \ldots, v_l, g_1, \ldots, g_q)$, linear in $(f, v_1, \ldots, v_l, g_1, \ldots, g_q)$ and quadratic in ϕ, any critical point $(\overline{x}, \overline{y}, \overline{\lambda})$ of the corresponding problem, P_{KKTBL} with associated multiplier $\alpha \neq 0$ (see system (5.2.5)) is isolated and non-degenerate in the MPCC sense, see Chapter 4.*

We combine the generic properties of the previous results in a definition.

Definition 5.2.1 A bilevel problem P_{BL} is called *KKT-regular* if its corresponding KKT relaxation P_{KKTBL} has the regularity properties of the generic class in Theorem 5.2.1 and Theorem 5.2.2.

As a direct consequence of this definition we obtain the following result.

Corollary 5.2.2 *For almost all perturbations of $(f, \phi, v_1, \ldots, v_l, g_1, \ldots, g_q)$, linear in $(f, v_1, \ldots, v_l, g_1, \ldots, g_q)$ and quadratic in ϕ, the problems P_{BL} are KKT-regular.*

We end up with some additional information on the relation between the original problem P_{BL} and the corresponding relaxation P_{KKTBL} in the generic case. In all cases we assume that P_{BL} is KKT-regular. For local minimizers with $\alpha \neq 0$ the following holds.

Corollary 5.2.3 *Let P_{BL} be a KKT-regular problem and let $(\overline{x}, \overline{y}, \overline{\lambda})$ be a non-degenerate local minimizer of the corresponding program P_{KKTBL} in (5.1.3) with multiplier $\alpha \neq 0$. Then the strong-MFCQ condition holds at $\overline{y}$ w.r.t. the lower level problem $Q(\overline{x})$.*

Under these conditions, if $(\overline{x}, \overline{y})$ is a feasible point of problem P_{BL}, then it is a local minimizer of P_{BL} of order 1 or 2.

Proof. Recall from Corollary 5.2.1 that a non-degenerate critical point of (5.1.3) with $\alpha \neq 0$ is isolated. Assume now that two lower level multipliers, $\overline{\lambda} \neq \lambda_1$ exist w.r.t. $Q(\overline{x})$. Then for $\delta \in [0, 1]$ the points $(\overline{x}, \overline{y}, (1-\delta)\overline{\lambda}+\delta\lambda_1)$ will also be feasible points of problem (5.1.3) with the same minimal value of the objective function $f(\overline{x}, \overline{y})$. So, for δ small enough $(\overline{x}, \overline{y}, (1-\delta)\overline{\lambda}+\delta\lambda_1)$ will be a local minimizer of problem P_{KKTBL}, and this contradicts the fact that $(\overline{x}, \overline{y}, \overline{\lambda})$ is an isolated critical point of P_{KKTBL}. So strong-MFCQ holds.

Now we will show that the local minimizer condition is satisfied. Note that in this case $(\alpha \neq 0)$ $(\overline{x}, \overline{y}, \overline{\lambda})$ is a non-degenerate critical point in the MPCC-sense. Then, by Theorems 4.4.3 and 4.4.4, it is a (locally unique) minimizer of order 1 or 2 for P_{KKTBL}. Moreover, strong-MFCQ holds in the lower level problem $Q(\overline{x})$. As $(\overline{x}, \overline{y})$ is a feasible point of P_{BL}, *i.e.*, $\overline{y}$ solves $Q(\overline{x})$, and MFCQ holds at $\overline{y} \in Y(\overline{x})$, locally $M_{BL} \subset M_{\text{KKTBL}}|_{\mathbb{R}^n \times \mathbb{R}^m}$, where $M_{\text{KKTBL}}|_{\mathbb{R}^n \times \mathbb{R}^m}$ is the projection of the set M_{KKTBL} into the (x, y)-space. In view of the fulfillment of the strong-MFCQ condition, the multipliers associated to $(x, y) \in M_{BL}$, near to $(\overline{x}, \overline{y})$ are close to $\overline{\lambda}$. Altogether, this implies that the point $(\overline{x}, \overline{y})$ is a (locally unique) minimizer of order 1 or 2 for P_{BL}.

□

Now let $(\overline{x}, \overline{y}, \overline{\lambda})$ be a local minimizer of (5.1.3) (and thus a critical point of P_{KKTBL}) with associated multiplier $\alpha = 0$. Recall that we assume that P_{BL} is KKT-regular. Therefore, it satisfies the conditions given in BL-2, see Theorem 5.2.2. In particular $|J_{0v}(\overline{x}, \overline{y})| \geq m$.

First we consider the case $|J_{0v}(\overline{x}, \overline{y})| = m$. As $rank(\nabla_y v_{J_{0v}(\overline{x},\overline{y})}(\overline{x}, \overline{y})) = m = |J_{0v}(\overline{x}, \overline{y}, \overline{\lambda})|$, LICQ is satisfied for $Q(\overline{x})$ at $\overline{y}$ and $\overline{\lambda}$ is the unique solution of system (5.2.13). This implies that MPCC-SC and MPCC-SOC hold at $(\overline{x}, \overline{y}, \overline{\lambda})$, see condition BL-2 in Theorem 5.2.2. Consequently, $(\overline{x}, \overline{y}, \overline{\lambda})$ is an isolated critical point of problem (5.1.3). Using this fact and the same ideas as in the proof of Corollary 5.2.3, we obtain the following result.

Corollary 5.2.4 *Let P_{BL} be a KKT-regular problem and let $(\overline{x}, \overline{y}, \overline{\lambda})$ be a non-degenerate local minimizer of the corresponding program $P_{\scriptscriptstyle KKTBL}$ in (5.1.3) with multiplier $\alpha = 0$ and $rank(\nabla_y v_{J_{0v}(\overline{x},\overline{y})}(\overline{x}, \overline{y})) = |J_{0v}(\overline{x}, \overline{y})| = m$. Then LICQ holds at $\overline{y}$ w.r.t. the lower level problem $Q(\overline{x})$.*

Under these hypotheses, if $(\overline{x}, \overline{y})$ is a feasible point of problem P_{BL}, then it is a local minimizer of P_{BL} of order 1 or 2.

Now we consider the case that $(\overline{x}, \overline{y}, \overline{\lambda})$ is a local minimizer of $P_{\scriptscriptstyle\mathrm{KKTBL}}$, with $\alpha = 0$, where strong-MFCQ fails at $Q(\overline{x})$ w.r.t. $\overline{y}$, hence $|J_{0v}(\overline{x}, \overline{y})| > m$. Then the set of solutions of the system (5.2.13) is a polyhedron of dimension d, $d \leq |J_{0v}(\overline{x}, \overline{y})| - m$. We will denote this polyhedron by $R(\overline{x}, \overline{y})$. Note that the vertices of $R(\overline{x}, \overline{y})$ are given by those solutions λ^* such that $rank(\nabla_y v_{J_0(\overline{x},\overline{y},\lambda^*)}(\overline{x}, \overline{y})) = |J_0(\overline{x}, \overline{y}, \lambda^*)| = m$.

For each $\lambda^* \in R(\overline{x}, \overline{y})$, the point $(\overline{x}, \overline{y}, \lambda^*)$ is a critical point of $P_{\scriptscriptstyle\mathrm{KKTBL}}$. In this situation the following bad behavior may occur: the points $(\overline{x}, \overline{y}, \lambda)$ with $\lambda \in R(\overline{x}, \overline{y})$ and $J\Lambda_0(\overline{x}, \overline{y}, \lambda) = \varnothing$, *i.e.*, λ is in the relative interior of $R(\overline{x}, \overline{y})$, are local minimizers of $P_{\scriptscriptstyle\mathrm{KKTBL}}$, but if $\overline{\lambda}$ is a vertex of $R(\overline{x}, \overline{y})$, the point $(\overline{x}, \overline{y}, \overline{\lambda})$ is no longer a local minimizer. This means, in particular, that the set of local minimizers may not be closed. We give an example:

Example 5.2.1

$$
\begin{array}{rl}
 & \min -x + y \\
\text{s.t.} \quad y \text{ solves } Q(x): & \qquad \min y \\
 & \qquad \text{s.t.} \quad \begin{array}{rcl} y & \geq & 0, \\ x & \geq & 0. \end{array}
\end{array}
$$

Note that the corresponding KKT relaxation is

$$
\begin{array}{rrcl}
 & \min -x + y & & \\
\text{s.t.} & 1 - \lambda_1 & = & 0, \\
 & y & \geq & 0, \\
 & x & \geq & 0, \\
 & \lambda_1, \lambda_2 & \geq & 0, \\
 & y\lambda_1 & = & 0, \\
 & x\lambda_2 & = & 0.
\end{array}
$$

Here, at the minimizer $(\overline{x}, \overline{y}) = (0, 0)$, we have $J_{0v}(\overline{x}, \overline{y}) = \{1, 2\}$. So $|J_{0v}(\overline{x}, \overline{y})| = 2 > 1 = m$. The points $(x, y, \lambda_1, \lambda_2) = (0, 0, 1, \lambda_2)$, with $\lambda_2 > 0$, are clearly local minimizers of the KKT relaxation, with the same value of the objective function, $f(0, 0) = 0$. However, if we take $(\lambda_1, \lambda_2) = (1, 0)$, the vertex solution of the corresponding system (5.2.13) (for $(\overline{x}, \overline{y}) = (0, 0)$) the point $(0, 0, 1, 0)$ is no longer a local minimizer. Indeed, when letting $x > 0$, the value of the objective function will be smaller. Note that $(0,0)$ is not a local minimizer of the original bilevel problem.

In the case where $(\overline{x},\overline{y},\overline{\lambda})$ is a local minimizer of P_{KKTBL}, with $\alpha = 0$, and such that strong-MFCQ fails for $\overline{y}$ at $Q(\overline{x})$, in comparison with Corollaries 5.2.3 and 5.2.4, we only have the following weaker result.

Corollary 5.2.5 *Let P_{BL} be a KKT-regular problem and let $(\overline{x},\overline{y},\overline{\lambda})$ be a local minimizer of P_{KKTBL}, with $\alpha = 0$, such that $\operatorname{rank}[\nabla v_{J_0(\overline{x},\overline{y},\overline{\lambda})}(\overline{x},\overline{y})] = |J_0(\overline{x},\overline{y},\overline{\lambda})| = m$ (i.e. $\overline{\lambda}$ is a vertex solution of system (5.2.13)) and strong-MFCQ fails for $\overline{y}$ at $Q(\overline{x})$. Then $(\overline{x},\overline{y},\overline{\lambda})$ is an isolated local minimizer of P_{KKTBL} (5.1.3) of order 2, in the following sense: There exists κ, $\kappa > 0$, such that for all (x,y) near $(\overline{x},\overline{y})$, with $(x,y,\lambda) \in M_{KKTBL}$ for some λ satisfying $\lambda_{J_0(\overline{x},\overline{y},\overline{\lambda})} > 0$, we obtain $f(x,y) - f(\overline{x},\overline{y}) \geq \kappa \|(x-\overline{x}, y-\overline{y})\|^2$.*

Proof. At a local minimizer $(\overline{x},\overline{y},\overline{\lambda})$ the multipliers $\beta_{J\Lambda_0}$, $\gamma_{J\Lambda_0}, \mu_{J_{0g}}$ are nonnegative and the Hessian of the Lagrangean of problem (5.1.3), should be positive semi-definite on $T_{(x,y,\lambda)}M_{\text{KKTBL}}$, *i.e.*, $\nabla^2_{(x,y,\lambda)} L(\overline{x},\overline{y},\overline{\lambda},\alpha,\beta,\mu,\gamma)|_{T_{(x,y,\lambda)}M_{\text{KKTBL}}} \succeq 0$, see the end of Section 4.3. In this case in view of $\alpha = \gamma = 0$ we have:

$$\nabla^2_{(x,y,\lambda)} L(\overline{x},\overline{y},\overline{\lambda},\alpha,\beta,\mu,\gamma) = \begin{pmatrix} \nabla^2_{(x,y)} \hat{L}(\overline{x},\overline{y},\beta,\mu) & 0 \\ 0 & 0 \end{pmatrix}$$

where

$$\hat{L}(\overline{x},\overline{y},\beta,\mu) = f(\overline{x},\overline{y}) - \sum_{i \in J_0(\overline{x},\overline{y},\overline{\lambda}) \cup J\Lambda_0(\overline{x},\overline{y},\overline{\lambda})} \beta_i v_i(\overline{x},\overline{y}) - \sum_{j \in J_{0g}(\overline{x},\overline{y})} \mu_j g_j(\overline{x},\overline{y}).$$

Now we will construct the tangent subspace $T_{(\overline{x},\overline{y},\overline{\lambda})}M_{\text{KKTBL}}$. Recall that, by definition, it is the tangent subspace of the manifold described by the system (5.2.2). As can be easily seen, this sub-space is generated by the $k = n - |J_{0g}| - |J\Lambda_0|$ columns of the matrix $\begin{pmatrix} V \\ \odot \\ 0_{l-m} \end{pmatrix}$ where the k columns of V form a basis of the subspace tangent to $\nabla_{(x,y)} v_{J_{0v}(\overline{x},\overline{y})}(\overline{x},\overline{y}), \nabla_{(x,y)} g_{J_{0g}}(\overline{x},\overline{y})$ and $\odot$ is the $m \times k$-matrix solving

$$\nabla_y v_{J_0(\overline{x},\overline{y},\overline{\lambda})}(\overline{x},\overline{y})^T \odot = -\nabla_{(x,y)} [\nabla_y \phi - \sum_{i=1}^{l} \lambda_i \nabla_y v_i(\overline{x},\overline{y})]^T V.$$

As a consequence

$$\nabla^2_{(x,y,\lambda)} L(\overline{x},\overline{y},\overline{\lambda},\alpha,\beta,\mu,\gamma)|_{T_{(\overline{x},\overline{y},\overline{\lambda})}M_{\text{KKTBL}}} = V^T \nabla^2_{(x,y)} \hat{L}(\overline{x},\overline{y},\beta,\mu) V \succeq 0.$$

But $\overline{\lambda}$ is such that $rank[\nabla v_{J_0(\overline{x},\overline{y},\overline{\lambda})}(\overline{x},\overline{y})] = |J_0(\overline{x},\overline{y},\overline{\lambda})| = m$ and for these critical points *MPCC-SOC* holds, see Condition *BL*-2 in Theorem 5.2.2. So

$$\nabla^2_{(x,y,\lambda)} L(\overline{x},\overline{y},\overline{\lambda},\alpha,\beta,\mu,\gamma)|_{T_{(\overline{x},\overline{y},\overline{\lambda})}M_{\mathrm{KKTBL}}} = V^T \nabla^2_{(x,y)} \hat{L}(\overline{x},\overline{y},\beta,\mu) V \text{ is positive definite.} \tag{5.2.18}$$

Recall that in case $\alpha = 0$ the regularity means $\mu_j, \beta_i \neq 0$, for all $j \in J_{0g}(\overline{x},\overline{y})$, $i \in J_0(\overline{x},\overline{y},\overline{\lambda}) \cup J\Lambda_0(\overline{x},\overline{y},\overline{\lambda})$, see Condition *BL*-2 in Theorem 5.2.2. As $(\overline{x},\overline{y},\overline{\lambda})$ is a minimizer of P_{KKTBL}, we must have:

$$\beta_i, \mu_j > 0,\ \forall i \in J\Lambda_0(\overline{x},\overline{y},\overline{\lambda}),\ j \in J_{0g}(\overline{x},\overline{y}). \tag{5.2.19}$$

Let us consider the problem:

$$\begin{gathered} \min f(x,y) \\ \text{s.t. } (x,y) \in \hat{\mathcal{M}}_{J_0(\overline{x},\overline{y},\overline{\lambda})} \end{gathered}$$

$$\hat{\mathcal{M}}_{J_0(\overline{x},\overline{y},\overline{\lambda})} = \left\{ (x,y) \in \mathbb{R}^{n+m} \;\middle|\; \begin{array}{rcll} v_i(x,y) & = & 0, & i \in J_0(\overline{x},\overline{y},\overline{\lambda}), \\ v_i(x,y) & \geq & 0, & i \in J\Lambda_0(\overline{x},\overline{y},\overline{\lambda}), \\ g_j(x,y) & \geq & 0, & j = 1,\ldots,q. \end{array} \right\} \tag{5.2.20}$$

It is easy to see that $(\overline{x},\overline{y})$ is a critical point of the previous problem with associated multipliers (β,μ). Since the Hessian of the Lagrangean of this problem coincides with the Hessian of $\hat{L}(\overline{x},\overline{y},\beta,\mu)$ and $T_{(\overline{x},\overline{y})}\hat{\mathcal{M}}_{J_0(\overline{x},\overline{y},\overline{\lambda})}$ is generated by the columns of V, conditions (5.2.19) and (5.2.18) mean that $(\overline{x},\overline{y})$ is a local minimizer of problem (5.2.20) of order 2, *i.e.*, there exists $\kappa,\ \kappa > 0$, such that, locally,

$$f(x,y) - f(\overline{x},\overline{y}) \geq \kappa \|(x-\overline{x}, y-\overline{y})\|^2, \forall (x,y) \in \hat{\mathcal{M}}_{J_0}(\overline{x},\overline{y},\overline{\lambda}).$$

Now, we take (x,y) near $(\overline{x},\overline{y})$ satisfying $(x,y,\lambda) \in M_{\mathrm{KKTBL}}$, for some λ, $\lambda_{J_0(\overline{x},\overline{y},\overline{\lambda})} > 0$. By the feasibility of (x,y,λ) we have $v_{J_0(\overline{x},\overline{y},\overline{\lambda})}(x,y) = 0$. Then $(x,y) \in \hat{\mathcal{M}}_{J_0(\overline{x},\overline{y},\overline{\lambda})}$ and the inequality $f(x,y) - f(\overline{x},\overline{y}) \geq \kappa\|(x-\overline{x}, y-\overline{y})\|^2$ holds. □

Remark 5.2.1 *Note that the condition obtained in Corollary 5.2.5 does not mean that if $(\overline{x},\overline{y}) \in M_{BL}$, then $(\overline{x},\overline{y})$ is a local minimizer of P_{BL}, as in the other cases.*

We now summarize the genericity results of this section. Generically, MPCC-LICQ holds for all feasible points of M_{KKTBL}. At the local minimizers $(\overline{x},\overline{y},\overline{\lambda})$ of P_{KKTBL}, with corresponding multipliers $(\alpha,\beta,\gamma,\mu)$, the following cases may appear generically:

- $\alpha \neq 0$: in this case for the lower level problem $Q(\overline{x})$, strong-MFCQ holds at $\overline{y}$ and $(\overline{x}, \overline{y}, \overline{\lambda})$ is non-degenerate, as a critical point, in the *MPCC-sense*.

- $\alpha = 0$: $(\overline{x}, \overline{y})$ is a non-degenerate critical point of problem (5.2.14) and $rank(\nabla_y v_{J_0(\overline{x},\overline{y},\overline{\lambda}) \cup J\Lambda_0(\overline{x},\overline{y},\overline{\lambda})}) = m$. Moreover there is some vertex solution λ^* of the system (5.2.13), such that $(\overline{x}, \overline{y}, \lambda^*)$ is a critical point of P_{KKTBL}, $|J_0(\overline{x}, \overline{y}, \lambda^*)| = m$ and $rank(\nabla_y v_{J_0(\overline{x},\overline{y},\lambda^*)}(\overline{x}, \overline{y})) = m$. More precisely, recall $J_{0v}(\overline{x}, \overline{y}) = J_0(\overline{x}, \overline{y}, \lambda^*) \cup J\Lambda_0(\overline{x}, \overline{y}, \lambda^*)$, we have :

    · in case $|J_{0v}(\overline{x}, \overline{y})| = m$:
    $\overline{\lambda}$ is the unique solution of system (5.2.13). LICQ holds at $\overline{y}$ for $Q(\overline{x})$ and $(\overline{x}, \overline{y}, \overline{\lambda})$ is a non-degenerate critical point of P_{KKTBL} in the MPCC-sense, *i.e., MPCC-SOC* and *MPCC-SC* are fulfilled.

    · in case $|J_{0v}(\overline{x}, \overline{y})| > m$:
    For the vertex solutions λ^* of the system (5.2.13) it follows $J\Lambda_0(\overline{x}, \overline{y}, \lambda^*) \neq \varnothing$. Moreover *MPCC-SC* fails for $(\overline{x}, \overline{y}, \lambda^*)$ since $\gamma_i = 0$, $\forall i \in J\Lambda_0(\overline{x}, \overline{y}, \lambda^*)$, but *MPCC-SOC* holds.

5.3 A numerical approach for solving BL

In this part we will present a procedure for computing local minimizers of P_{KKTBL}. As already sketched at the end of Section 5.2, there may appear two different cases: local minimizers corresponding to critical points where $\alpha = 0$ or where $\alpha \neq 0$. First let us reconsider the reduced problem, see (5.2.14):

$$\begin{array}{rrcll} & \multicolumn{4}{l}{\min f(x,y)} \\ \text{s.t.} & v_i(x,y) & \geq & 0, & i = 1, \dots, l, \\ & g_j(x,y) & \geq & 0, & j = 1, \dots, q. \end{array} \tag{5.3.1}$$

With the aid of this program we have the following possible method.

Conceptual method:

1. *Possible solutions with* $\alpha = 0$: Try to compute a critical point $(\overline{x}, \overline{y})$ of the reduced nonlinear program (5.3.1). If there is a corresponding solution λ of system (5.2.13), then $(\overline{x}, \overline{y}, \lambda)$ is a critical point of problem P_{KKTBL}, with $\alpha = 0$. Now we calculate a vertex solution $\overline{\lambda}$ of system (5.2.13). Then, generically, $rank(\nabla_y v_{J_0(\overline{x},\overline{y},\overline{\lambda})}(\overline{x}, \overline{y})) = |J_0(\overline{x}, \overline{y}, \overline{\lambda})| = m$. If conditions (5.2.19) and (5.2.18) hold, $(\overline{x}, \overline{y}, \overline{\lambda})$ is a local minimizer of problem (5.1.3).

2. *Possible solutions with* $\alpha \neq 0$: Try to find a solution $(\overline{x}, \overline{y}, \overline{\lambda})$ of the full system (5.2.5), by applying the parametric smoothing approach of Chapter 4. If the procedure converges to a point $(\overline{x}, \overline{y}, \overline{\lambda})$ with $\alpha = 0$, try to identify the active index sets $J_{0v}(\overline{x}, \overline{y})$, $J_{0g}(\overline{x}, \overline{y})$ and switch to step 1.

The first step can be easily done by applying standard nonlinear optimization techniques. Note that the program (5.3.1) will generically satisfy the non degeneracy conditions of nonlinear programming problems.

For the smoothing approach in Step 2, we solve the perturbed problem:

$$
\begin{array}{rrcll}
 & \min f(x,y) & & & \\
\text{s.t.} \quad \nabla_y \phi(x,y) - \sum\limits_{i=1}^{l} \lambda_i \nabla_y v_i(x,y) & & = & 0, & \\
 & v_i(x,y) & \geq & 0, & i = 1,\ldots,l, \\
 & \lambda_i & \geq & 0, & i = 1,\ldots,l, \\
 & \lambda_i v_i(x,y) & = & \tau, & i = 1,\ldots,l, \\
 & g_j(x,y) & \geq & 0, & j = 1,\ldots,q
\end{array}
\tag{5.3.2}
$$

and let $\tau \to 0^+$. So suppose that $(\overline{x}, \overline{y}, \overline{\lambda})$ is a critical point of P_{KKTBL} with $\alpha \neq 0$, where α is the vector of multipliers associated with the equality constraints $\nabla_y \phi(x,y) - \sum_{i=1}^{l} \lambda_i v_i(x,y)$. The existence of a sequence of local minimizers $(\overline{x}_\tau, \overline{y}_\tau, \overline{\lambda}_\tau)$ of problem (5.3.2), and its rate of convergence follows from the analysis in Section 4.6.

Proposition 5.3.1 *Let $(\overline{x}, \overline{y}, \overline{\lambda})$ be a strongly stationary point of the MPCC problem (5.1.3) such that* MPCC-LICQ, MPCC-SC MPCC-SOC *hold. Then there is a sequence of stationary points $(x_\tau, y_\tau, \lambda_\tau)$ of problem (5.3.2) converging to $(\overline{x}, \overline{y}, \overline{\lambda})$ with rate $O(\sqrt{\tau})$.*

Proof. It is a direct corollary of Theorem 4.5.1. □

In the generic case, the hypotheses in Proposition 5.3.1 hold if strong-MFCQ is fulfilled for $\overline{y}$ in $Q(\overline{x})$. Sufficient conditions under KKT-regularity are that the multiplier α corresponding to $(\overline{x}, \overline{y}, \overline{\lambda})$ is not equal to 0 or that $|J_{0v}(\overline{x}, \overline{y})| = m$. Unfortunately, around a stationary point $(\overline{x}, \overline{y}, \overline{\lambda})$ with $\alpha = 0$ and $J\Lambda_0(\overline{x}, \overline{y}, \overline{\lambda}) \neq \varnothing$, we cannot guarantee the existence of a sequence of stationary points $(x_\tau, y_\tau, \lambda_\tau)$ of problem (5.3.2), converging to $(\overline{x}, \overline{y}, \overline{\lambda})$, when $\tau \to 0$, as in Propostion 5.3.1. In the following example we present a generic counterexample.

Example 5.3.1

$$
\begin{array}{lcrcl}
 & \min x + y & & & \\
\text{s.t. } y \text{ solves } Q(x): & & \min y & & \\
 & s.t. & y & \geq & 0, \\
 & & 1 - y & \geq & 0, \\
 & & x & \geq & 0.
\end{array}
$$

The KKT approach leads us to the following problem:

$$\begin{array}{rcll} & \min x+y & & \\ \text{s.t.} \quad 1-\lambda_1+\lambda_2 & = & 0, & \\ y & \geq & 0, & \\ 1-y & \geq & 0, & \\ x & \geq & 0, & \\ \lambda_i & \geq & 0, & i=1,2,3, \\ \lambda_1 y & = & 0, & \\ \lambda_2(1-y) & = & 0, & \\ \lambda_3 x & = & 0. & \end{array}$$

This problem has the solution $(\overline{x}, \overline{y}, \overline{\lambda}) = (0,0,1,0,0)$ with multipliers $\beta_1 = \beta_2 = 1$, $\alpha = \gamma_1 = \gamma_2 = \gamma_3 = 0$, and $|J\Lambda_0(\overline{x}, \overline{y}, \overline{\lambda})| = 1$. Now, we consider the smoothing parameterization of Section 4.6 and try to solve:

$$P_\tau: \quad \begin{array}{rcl} & \min x+y & \\ \text{s.t.} \quad 1-\lambda_1+\lambda_2 & = & 0, \\ \lambda_1 & \geq & 0, \\ y & \geq & 0, \\ \lambda_2 & \geq & 0, \\ 1-y & \geq & 0, \\ \lambda_3 & \geq & 0, \\ x & \geq & 0, \\ y\lambda_1 & = & \tau, \\ (1-y)\lambda_2 & = & \tau, \\ x\lambda_3 & = & \tau, \end{array}$$

for $\tau \to 0^+$. Note that, for $\tau > 0$ small enough, LICQ holds at any feasible point. In fact, as the active index set is empty, the linear dependence implies that the gradients of equations number 1, 8, 9 and 10 are linearly dependent, *i.e.*, the following system has a non-trivial solution (a,b,c,d):

$$\begin{pmatrix} 0 & 0 & 0 & \lambda_3 \\ 0 & \lambda_1 & -\lambda_2 & 0 \\ -1 & y & 0 & 0 \\ 1 & 0 & 1-y & 0 \\ 0 & 0 & 0 & x \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = 0.$$

$x > 0$ implies $d = 0$. By multiplying the third and fourth equations by λ_1^2 and $-\lambda_2^2$, respectively, and using the feasibility condition, we obtain:

$$\begin{array}{rcl} \lambda_1 b\tau & = & \lambda_1^2 a, \\ -\lambda_2 c\tau & = & \lambda_2^2 a. \end{array}$$

Summing up and using $b\lambda_1 - c\lambda_2 = 0$, we obtain $a(\lambda_1^2 + \lambda_2^2) = 0$. But $\lambda_1, \lambda_2 > 0$, so $a = 0$ and this implies $b = c = 0$, a contradiction to $(a, b, c, d) \neq 0$.

Now we will prove by contradiction that there cannot exist a sequence of critical points of P_τ converging to (0,0,1,0,0) when $\tau \to 0^+$. The critical point condition for problem P_τ reads:

$$\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & \lambda_3 \\ 0 & \lambda_1 & -\lambda_2 & 0 \\ -1 & y & 0 & 0 \\ 1 & 0 & 1-y & 0 \\ 0 & 0 & 0 & x \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}$$

The condition $x > 0$ yields $d = 0$, so in the first equation we have $1 = 0$ contradicting the existence of multipliers (a, b, c, d) solving the system.

We want to recall that there may exist feasible points $(x, y) \in M_{BL}$, or even minimizers $(\overline{x}, \overline{y})$ of problem P_{BL} in (5.1.1), where the MFCQ condition is violated in the lower level problem, see problem (5.2.6). So $\overline{y}$ may not satisfy the KKT condition for $Q(\overline{x})$. Clearly minimizers $(\overline{x}, \overline{y})$, such that the KKT condition fails at $\overline{y}$ for $Q(\overline{x})$, cannot be found by the KKT approach. Therefore from a practical viewpoint it is advisable to use the FJ approach in which the KKT condition (5.1.2) is replaced by the FJ optimality condition:

$$\begin{aligned} \lambda_0 \nabla_y \phi(x, y) - \sum_{i=1}^{l} \lambda_i \nabla_y v_i(x, y) &= 0, \\ \lambda_0 &\geq 0, \\ \lambda_i &\geq 0, \quad i = 1, \ldots, l, \\ v_i(x, y) &\geq 0, \quad i = 1, \ldots, l, \\ \lambda_i v_i(x, y) &= 0, \quad i = 1, \ldots, l, \\ \lambda_0 + \sum_{i=1}^{l} \lambda_i &= 1. \end{aligned}$$

So we are lead to the problem:

$$\begin{aligned} &\min f(x, y) \\ \text{s.t.} \quad \lambda_0 \nabla_y \phi(x, y) - \sum_{i=1}^{l} \lambda_i \nabla_y v_i(x, y) &= 0, \\ \lambda_0 &\geq 0, \\ \lambda_i &\geq 0, \quad i = 1, \ldots, l, \\ v_i(x, y) &\geq 0, \quad i = 1, \ldots, l, \\ \lambda_i v_i(x, y) &= 0, \quad i = 1, \ldots, l, \\ \lambda_0 + \sum_{i=1}^{l} \lambda_i &= 1, \\ g_j(x, y) &\geq 0, \quad j = 1, \ldots, q. \end{aligned} \tag{5.3.3}$$

The FJ approach leads, up to some technical modification, to a problem with a similar structure as in the KKT approach. We, therefore, expect similar genericity results.

5.4 Numerical examples

In this section we will consider four bilevel problems to illustrate the solution method given before. Three of them appear on the web page http://www-unix.mcs.anl.gov/~leyffer/MacMPEC. The fourth is the problem in Example 5.3.1. We will present the numerical results obtained when solving the problem P_{KKTBL}, corresponding to the KKT approach applied to problem P_{BL}. The solutions are computed by means of the smoothing techniques described in Chapter 4. We use the solver of nonlinear optimization problems of MATLAB, *fmincon*, to find a critical point of the corresponding problem P_τ, $\tau \to 0^+$. All programs run under MATLAB 7.0.1.

We want to remark that the numerical results will be displayed with two decimal places, even if the corresponding number is exact. For instance, the values 12 and 12.00009 will be both written as 12.00.

The first problem originally appeared in Bard [2].

$$
\begin{array}{lrl}
 & \min(x_1-5)^2+(2y+1)^2 & \\
\text{s.t.} & x \geq 0, & \\
 & y \geq 0, & \\
y \text{ solves } Q(x): & \min(y-1)^2 - 1.5xy & \\
 & \text{s.t.} \quad 3x - y - 3 & \geq 0, \\
 & -x + 0.5y + 4 & \geq 0, \\
 & -x - y + 7 & \geq 0.
\end{array}
$$

The reported local minimizer is $(\overline{x}, \overline{y}, \overline{\lambda}) = (1, 0, 3.5, 0, 0)$. We obtained:

Method	Smoothing method
Starting Solution	(0.00, 0.00)
Start lower level multipliers	(1.00, 1.00, 1.00)
Solution	(1.00, 0.00)
Lower level multipliers	(3.50, 0.00, 0.00)
Value of the obj. function	17.00
CPU time	.20

The smoothing method converged to the minimizer. The error between the approximate and the real minimizer for $\tau = 9.54$ *e*-07 is 0.39 *e*-06.

We solved the same problem using 50 random starting points on the intervals $[-1, 5]^2$ and $[-20, 20]^2$. It performed as follows:

Behavior\Interval	$[-1,5]^2$	$[-20,20]^2$
Smoothing approach succeeded in	33 cases	16 cases

The second example, also appeared in [2], and reads:

$$\min -x_1^2 - 3x_2 - 4y_1 + y_2^2$$

s.t.

$$\begin{array}{rrcl}
& x & \geq & 0, \\
& y & \geq & 0, \\
& -x_1^2 - 2x_2 + 4 & \geq & 0, \\
y \text{ solves } Q(x): & \min y_1^2 - 5y_2 & & \\
& \text{s.t.} \quad x_1^2 - 2x_1 + x_2^2 - 2y_1 + y_2 + 3 & \geq & 0, \\
& x_2 + 3y_1 - 4y_2 - 4 & \geq & 0.
\end{array}$$

The numerical results are given in the next table:

Method	Inner point
Starting solution	(0.00, 0.00, 0.00, 0.00)
Start lower level multipliers	(1.00, 1.00)
Solution	(0.00, 2.00, 1.88, 0.91)
Lower level multipliers	(0.00, 1.25)
Value of the objective function	-12.67
CPU time	33.31

Again the smoothing approach converges to the reported minimizer, $(\overline{x}, \overline{y}, \overline{\lambda}) = (0, 2, 1.875, 0.9062, 0, 1.25)$. The error between the obtained point and the real minimizer is 4.97 e-05 for $\tau = 9.54$ e-07.

Now, by considering 50 starting points randomly generated in the intervals $[-1,5]^4$ and $[-20,20]^4$ we obtained:

Behavior\Interval	$[-1,5]^4$	$[-20,20]^4$
Smoothing approach succeeded in	44 cases	45 cases

Note that in the previous examples, at the solutions we have $J\Lambda_0(\overline{x}, \overline{y}, \overline{\lambda}) = \varnothing$ *i.e.*, strong SC holds (see Section 4.3). However, this condition may fail in generic examples as in Example 5.3.1, where at the solutions $(\overline{x}, \overline{y}, \overline{\lambda}) = (0, 0, 1, 0, \lambda_3)$, $\lambda_3 \geq 0$ the corresponding multiplier satisfies $\alpha = 0$. Recall the problem is

$$\begin{array}{lrcl}
& \min x + y & & \\
\text{s.t. } y \text{ solves } Q(x): & \min y & & \\
\text{s.t.} & y & \geq & 0, \\
& 1 - y & \geq & 0, \\
& x & \geq & 0.
\end{array}$$

With the starting point (0,0,1,1,1), the smoothing approach obtained the solution (2.17 e-09, 9.54 e-07, 1.00, 9.547 e-07, 438.59). As expected, at the solution, the lower level multipliers became large. When τ is small, we observed that the value of λ_3 remained almost unchanged. In fact, for τ small, $x = \frac{\tau}{438.593}$ is almost equal to 0. So, the point $(\frac{\tau}{438.593}, 0, 1, 0, \lambda_3)$ solves the critical points system within an acceptable error.

We end with a non-generic example, where *MPCC-LICQ* fails at the solution point. This example can be found in [13]. The problem is

$$
\begin{array}{lrcl}
 & \min -x_1^2 - 2x_1 + x_2^2 - 2x_2 + y_1^2 + y_2^2 & & \\
\text{s.t.} & & & \\
 & x & \geq & 0, \\
 & y & \geq & 0, \\
 & -x_1 + 2 & \geq & 0, \\
y \text{ solves } Q(x): & \min y_1^2 - 2x_1y_1 + y_2^2 - 2x_2y_2 & & \\
 & \text{s.t.} \quad .25 - (y_1 - 1)^2 & \geq & 0, \\
 & .25 - (y_2 - 1)^2 & \geq & 0.
\end{array}
$$

In this case the minimizer is $(.5, .5, .5, .5)$ and the lower level multipliers, $\lambda = (0, 0)$. At that point, *MPCC-LICQ* is not satisfied. Our approach behaved surprisingly good with error of order $\sqrt{\tau}$ for $\tau \geq 8.6736$ e-19.

Chapter 6

Equilibrium constrained problems

6.1 Introduction

In this part we will consider the general case of optimization problems with equilibrium constraints (see Section 1.1)

$$P_{EC}: \qquad \begin{array}{c} \min f(x,y) \\ \text{s.t. } (x,y) \in M_{EC} \end{array} \tag{6.1.1}$$

$$M_{EC} = \left\{ (x,y) \in \mathbb{R}^n \times \mathbb{R}^m \;\middle|\; \begin{array}{rcll} g_j(x,y) & \geq & 0, & j = 1, \ldots, q, \\ y & \in & Y(x), & \\ \phi(x,y,z) & \geq & 0, & \forall z \in Y(x) \end{array} \right\}$$

where $Y(x) = \{y \in \mathbb{R}^m \,|\, v_i(x,y) \geq 0,\ i = 1, \ldots, l\}$, $f \in [C^\infty]^1_{n+m}$, $\phi \in [C^\infty]^1_{n+m+m}$, $(g_1, \ldots, g_q) \in [C^\infty]^q_{n+m}$ and $(v_1, \ldots, v_l) \in [C^\infty]^l_{n+m}$. For simplicity we omit equality constraints.

The following problems can be seen as special instances of P_{EC}:

- Optimization problems with a VI constraint, see (1.1.8), where $\phi(x,y,z) = \Phi(x,y)^T(z-y)$.

- Stackelberg Games $\min f(x,y)$ s.t. $y \in S(x)$ where $S(x)$ is the set of Nash strategies.

- Bilevel problems, see Chapter 5.

- Generalized semi-infinite problem (GSIP), see [58].

As for the case of BL programs, the structure of the feasible set of P_{EC} is not suitable for classical optimization approaches. Instead, we again propose and analyze the KKT approach based on a transformation of the problems into a program with complementarity constraints.

The chapter is organized as follows. Firstly we are interested in the topological structure of the feasible set of P_{EC} and we give some necessary condition for convexity and closedness. In Section 6.3 we consider the MPCC problem obtained by the KKT approach for solving P_{EC}. We will show that one cannot expect *MPCC-LICQ* to be fulfilled generically at all feasible points. Section 6.4 deals with the special case where LICQ holds in $Y(x)$, and $\phi,\ v_i,\ g_j$, $i = 1, \ldots, l,\ j = 1, \ldots, q$, satisfy certain convexity conditions. For the MPCC form of P_{EC} we prove that for almost all perturbations *MPCC-LICQ* is satisfied at all feasible points and that the critical points are non-degenerate in the MPCC-sense. In the last section we discuss the linear case, where the functions $f,\ v_i,\ g_j$, $i = 1, \ldots, l,\ j = 1, \ldots, q$ are linear and $\phi(x, y, z)$ is of the form $[C(x, y)+d](z-y)$. We derive the generic properties for these problems and present a numerical algorithm.

6.2 Structure of the feasible set

In this section we are interested in conditions which guarantee that the feasible set M_{EC} of P_{EC} is closed and/or convex. Note that in general, as in the case of GSIP (see [58]), M_{EC} need not to be closed. We give the illustrative example:

Example 6.2.1 Choose $\phi(x, y, z) = z - \frac{3}{2},\ x \geq 0,\ x \leq \frac{\pi}{2},\ y \in Y(x)$ and $Y(x) = \{z \in \mathbb{R} \mid (z - 2)(z - \sin(x))^2 \geq 0,\ z \geq 1\}$.

The feasible set then reads:

$$M_{EC} = \left\{ (x, y) \in \mathbb{R} \times \mathbb{R} \;\middle|\; \begin{array}{rcl} x & \in & [0, \frac{\pi}{2}] \\ y & \in & Y(x), \\ z - \frac{3}{2} & \geq & 0,\ \forall z \in Y(x) \end{array} \right\}.$$

We find

$$Y(x) = \begin{cases} [2, \infty), & \text{for } 0 \leq x < \frac{\pi}{2} \\ \{1\} \cup [2, \infty), & \text{for } x = \frac{\pi}{2}, \end{cases}$$

leading to the feasible set $M_{EC} = [0, \frac{\pi}{2}) \times [2, \infty)$, which is neither closed nor open. Note that this is due to the failure of MFCQ at $z = 1$ with respect to $Y(\frac{\pi}{2})$. As a result, the mapping $Y(x)$ is not lower semi-continuous at $z = 1$, since there is no sequence $(x_k, z_k) \to (\frac{\pi}{2}, 1)$, with $z_k \in Y(x_k)$. Recall that a set valued map $Y : \mathbb{R}^n \to 2^{\mathbb{R}^m}$ is called lower semi-continuous at $\overline{x} \in \mathbb{R}^n$ if for all open sets $V \subset \mathbb{R}^m$ such that $V \cap Y(\overline{x}) \neq \varnothing$ there is a neighborhood U of $\overline{x}$ such that for all $x \in U,\ Y(x) \cap V \neq \varnothing$. For more details we refer to Klatte and Kummer [36].

The next result provides sufficient conditions for closedness/convexity of M_{EC}, see [7] and also [58] for GSIP.

Proposition 6.2.1 (*cf.* [**7**]) *Suppose that $Y(x) \subset K$ holds for all $x \in \mathbb{R}^n$ with some compact set K, $K \subset \mathbb{R}^m$. Assume that for all $x \in \mathbb{R}^n$ the condition MFCQ holds at all $z \in Y(x)$. Then the set valued mapping $Y(x)$ is lower semi-continuous and M_{EC} is closed.*

Suppose that for any x, the set $Y(x)$ is convex. Let the functions $\phi(x, y, z)$ and $g_j(x, y)$, $j = 1, \dots, q$ be concave and let $Y(\alpha x_1 + (1-\alpha)x_2) \subset \alpha Y(x_1) + (1-\alpha)Y(x_2)$, $\alpha \in [0, 1]$ hold for all x_1, x_2. Then M_{EC} is convex.

Proof. It is well known that the MFCQ condition implies that $Y(x)$ is lower semi-continuous, see [36]. To prove that the feasible set M_{EC} is closed, we have to show:

$$\text{for all } (x_k, y_k) \in M_{EC},\ (x_k, y_k) \to (\overline{x}, \overline{y}), \quad \Rightarrow \quad (\overline{x}, \overline{y}) \in M_{EC}.$$

By continuity, the assumptions entail $\overline{y} \in Y(\overline{x})$ and $g_j(\overline{x}, \overline{y}) \geq 0$. As $Y(\overline{x})$ is compact, we can consider $z^* \in Y(\overline{x})$ such that

$$z^* \text{ solves} \quad \begin{array}{l} \min\limits_z \phi(\overline{x}, \overline{y}, z) \\ \text{s.t. } z \in Y(\overline{x}). \end{array}$$

But $Y(\overline{x})$ is lower semi-continuous, so there is a sequence of points of $z_k \in Y(x_k)$, such that $z_k \to z^*$. Due to the feasibility assumption $\phi(x_k, y_k, z_k) \geq 0$ and taking limits $k \to \infty$, yields $0 \leq \phi(\overline{x}, \overline{y}, z^*) \leq \phi(\overline{x}, \overline{y}, z)$, $\forall z \in Y(x)$. So, $(\overline{x}, \overline{y}) \in M_{EC}$.

For the second part, if (x_1, y_1) and (x_2, y_2), are two feasible points, by the convexity assumptions, it follows (using standard arguments) that for any $\alpha \in [0, 1]$, the point $(\alpha x_1 + (1-\alpha)x_2, \alpha y_1 + (1-\alpha)y_2)$, is also feasible.

$\square$

Let us write the program P_{EC} equivalently in the bilevel form:

$$\begin{array}{lrcll} & \min\limits_{x,y,z} f(x, y) & & & \\ \text{s.t.} & g_j(x, y) & \geq & 0, & j = 1, \dots, q, \\ & y & \in & Y(x), & \\ & \phi(x, y, z) & \geq & 0, & \\ z \text{ solves } Q(x, y): & \min\limits_u \phi(x, y, u) & & & \\ & \text{s.t. } u \in Y(x). & & & \end{array} \tag{6.2.1}$$

Unfortunately, even under the strong assumptions of Proposition 6.2.1 we cannot guarantee the convexity of the corresponding feasible set. The reason is that if z_i, $i = 1, 2$ are minimizers of the problems

$$\begin{array}{l} \min\limits_u \phi(x_i, y_i, u) \\ \text{s.t. } u \in Y(x_i) \end{array}$$

the point $\alpha z_1 + (1-\alpha)z_2, \alpha \in [0,1]$ is not necessarily a minimizer of

$$\min_u \phi(\alpha x_1 + (1-\alpha)x_2, \alpha y_1 + (1-\alpha)y_2, u)$$

$$\text{s.t. } u \in Y(\alpha x_1 + (1-\alpha)x_2).$$

For instance, by taking $\phi(x,y,z) = xz$ and $Y(x) = [0,1]$, the feasible set M_{EC} is

$$M_{EC} = \left\{ (x,y) \in \mathbb{R} \times \mathbb{R} \;\middle|\; \begin{array}{c} y \in [0,1], \\ xz \geq 0,\ \forall z \in [0,1] \end{array} \right\} = \left\{ (x,y) \;\middle|\; \begin{array}{ccc} y & \in & [0,1], \\ x & \geq & 0 \end{array} \right\}.$$

The feasible set of the corresponding BL problem (6.2.1) reads:

$$M_{BL} = \left\{ (x,y,z) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R} \;\middle|\; \begin{array}{rcl} y & \in & [0,1], \\ xz & \geq & 0, \\ z \text{ solves } Q(x,y): & & \min\limits_u xu \\ & & \text{s.t. } u \in [0,1]. \end{array} \right\}$$

At $x_1 = y_1 = 0$ the set of minimizers of the lower level problem $Q(0,0)$ is [0,1]. Let us take $z_1 = 1$. For $x_2 = 1,\ y_2 = 0$ the only solution of $Q(1,0)$ is $z_2 = 0$. However the convex combination $(1-\alpha)(x_1,y_1,z_1) + \alpha(x_2,y_2,z_2) = (\alpha, 0, 1-\alpha)$ is not feasible for $0 < \alpha < 1$ since $z = (1-\alpha)$ does not solve the lower level problem $\min\limits_u \alpha u$ s.t. $u \in [0,1]$.

6.3 Genericity analysis of the KKT approach to EC

Now we will consider the KKT-relaxation for equilibrium problems P_{EC} in (6.1.1). Firstly the equilibrium problem is written in the bilevel form (6.2.1). Then the lower level minimization problem is replaced by the KKT condition for the minimizer z of $Q(x,y)$. We obtain the MPCC problem:

$$P_{\text{KKTEC}}: \quad \begin{array}{c} \min\limits_{x,y,z,\lambda} f(x,y) \\ \text{s.t. } (x,y,z,\lambda) \in M_{\text{KKTEC}} \end{array} \tag{6.3.1}$$

$$M_{\text{KKTEC}} = \left\{ \begin{array}{c} (x,y,z,\lambda) \in \\ \mathbb{R}^{n+m+m+l} \end{array} \;\middle|\; \begin{array}{rcll} g_j(x,y) & \geq & 0, & j = 1,\dots,q, \\ v_i(x,y) & \geq & 0, & i = 1,\dots,l, \\ \nabla_z \phi(x,y,z) - \sum\limits_{i=1}^{l} \lambda_i \nabla_z v_i(x,z) & = & 0, & \\ \lambda_i & \geq & 0, & i = 1,\dots,l, \\ v_i(x,z) & \geq & 0, & i = 1,\dots,l, \\ \lambda_i v_i(x,z) & = & 0, & i = 1,\dots,l, \\ \phi(x,y,z) & \geq & 0. & \end{array} \right\}$$

Recall that this problem is a relaxation of P_{EC} in the following sense. If MFCQ holds for the local minimizer $\overline{z}$ of $Q(\overline{x},\overline{y})$, then $(\overline{x},\overline{y},\overline{z},\overline{\lambda}) \in M_{\text{KKTEC}}$ for some $\overline{\lambda} \in \mathbb{R}^l_+$. Using a notation similar to the previous chapter, for points $(\overline{x},\overline{y},\overline{z},\overline{\lambda}) \in M_{\text{KKTEC}}$, we introduce the active index sets:

$$\begin{array}{rcl} J_{0v}(\overline{x},\overline{y}) & = & \{i \mid v_i(\overline{x},\overline{y}) = 0\} \\ J_0(\overline{x},\overline{z},\overline{\lambda}) & = & \left\{i \mid v_i(\overline{x},\overline{z}) = 0,\ \overline{\lambda}_i > 0\right\} \\ J\Lambda_0(\overline{x},\overline{z},\overline{\lambda}) & = & \left\{i \mid v_i(\overline{x},\overline{z}) = \overline{\lambda}_i = 0\right\} \\ \Lambda_0(\overline{x},\overline{z},\overline{\lambda}) & = & \left\{i \mid v_i(\overline{x},\overline{z}) > 0,\ \overline{\lambda}_i = 0\right\} \\ J_{0g}(\overline{x},\overline{y}) & = & \{j \mid g_j(\overline{x},\overline{y}) = 0\}\,. \end{array} \tag{6.3.2}$$

We begin with a negative example which shows that, in contrast with the BL case, MPCC-LICQ is no more generically fulfilled on the whole feasible set M_{KKTEC}.

Example 6.3.1 Consider the problem P_{EC} with feasible set M_{EC} defined by $\phi(x,y,z) = z+1,\ v_1(x,y) = y,\ v_2(x,y) = x-y,\ x,y,z \in \mathbb{R}$ and with $q = 0$, *i.e.*, no constraints $g_j(x,y) \geq 0$.

So

$$M_{EC} = \left\{ (x,y) \in \mathbb{R}\times\mathbb{R} \;\middle|\; \begin{array}{rcl} y & \geq & 0, \\ x-y & \geq & 0, \\ z+1 & \geq & 0,\ \forall z \in [0,x] \end{array} \right\}$$

The set M_{KKTEC} corresponding to the KKT relaxation P_{KKTEC} is described by:

$$\begin{array}{rcl} y & \geq & 0, \\ x-y & \geq & 0, \\ 1-\lambda_1+\lambda_2 & = & 0, \\ \lambda_1,\lambda_2 & \geq & 0, \\ z & \geq & 0, \\ x-z & \geq & 0, \\ \lambda_1 z & = & 0, \\ \lambda_2(x-z) & = & 0, \\ z+1 & \geq & 0. \end{array} \tag{6.3.3}$$

Note that $(x,y,z,\lambda_1,\lambda_2) = (0,0,0,1,0)$ is feasible. At this point the condition *MPCC-LICQ* fails, because the gradients of the active constraints $v_1(x,y), v_2(x,y)$, $v_1(x,z), v_2(x,z)$, *i.e.* the columns of the matrix

$$\begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

are linearly dependent. Moreover this situation is stable under small perturbations of the involved functions.

To see this, let us consider the set $\overline{M}_{\text{KKTEC}}$ defined by the functions $v_1(x,y)+\epsilon_1(x,y)$, $v_2(x,y)+\epsilon_2(x,y)$, $\phi(x,y,z)+\epsilon_3(x,y,z)$ where $|\epsilon_i(x,y)| \ll 1$, $|\nabla\epsilon_i(x,y)| \ll 1$, $i=1,2$, $|\epsilon_3(x,y,z)| \ll 1$ and $|\nabla\epsilon_3(x,y,z)| \ll 1$. As in problem (5.2.6), it can be seen that for fixed functions $\epsilon_1(x,y), \epsilon_2(x,y)$ small enough there exists (x^*,y^*) solving:

$$\begin{aligned} y+\epsilon_1(x,y) &= 0, \\ x-y+\epsilon_2(x,y) &= 0. \end{aligned}$$

The KKT condition of the perturbed problem reads:

$$1+\frac{\partial\epsilon_3}{\partial z}(x,y,z)-\lambda_1(1+\frac{\partial\epsilon_1}{\partial z}(x,z))+\lambda_2(1-\frac{\partial\epsilon_2}{\partial z}(x,z))=0,\ \lambda_1,\lambda_2\geq 0.$$

Let $x=x^*,\ y=y^*$ and $z=y^*$. As $|\nabla\epsilon_i| \ll 1$, it follows $1+\frac{\partial\epsilon_1}{\partial z}(x^*,y^*)>0$ and $1+\frac{\partial\epsilon_3}{\partial z}(x^*,y^*)>0$. Then for $(\lambda_1^*,\lambda_2^*)=\left(\frac{1+\frac{\partial\epsilon_3}{\partial z}(x^*,y^*,y^*)}{1+\frac{\partial\epsilon_1}{\partial z}(x^*,y^*)},0\right)$ the point

$$(x,y,z,\lambda_1,\lambda_2)=\left(x^*,y^*,y^*,\frac{1+\frac{\partial\epsilon_3}{\partial z}(x^*,y^*,y^*)}{1+\frac{\partial\epsilon_1}{\partial z}(x^*,y^*)},0\right)$$

is an element of the perturbed set $\overline{M}_{\text{KKTEC}}$. The gradients of the active constraints $v_1(x,y), v_2(x,y), v_1(x,z), v_2(x,z)$ remain linearly dependent at the point $(x,y,z,\lambda)=(x^*,y^*,y^*,\lambda^*)$. So, *MPCC-LICQ* fails.

Remark 6.3.1 *The previous example shows that EC problems are more complicated than BL programs. While MPCC-LICQ holds generically for all feasible points of the KKT-relaxation for BL-problems, the extra condition $y \in Y(x)$ implies that the MPCC-LICQ condition is not generically fulfilled for equilibrium constrained problems.*

In fact the failure of MPCC-LICQ is a consequence of the condition $z,y \in Y(x)$ and the violation of LICQ in problem $Q(x,y)$. Note that if $\overline{y}=\overline{z}$ and $\nabla_y v_{J_{0v}(\overline{x},\overline{y})}(\overline{x},\overline{y})$ are linearly dependent (see (6.3.2)) the MPCC-LICQ condition does not hold at $(\overline{x},\overline{y},\overline{z},\lambda) \in M_{\text{KKTEC}}$. Precisely, as shown in the previous example, we cannot expect $\nabla_y v_{J_{0v}(\overline{x},\overline{y})}$ to have rank $|J_{0v}(\overline{x},\overline{y})|$ for all $y \in Y(x)$.

6.4 Convex case

In this section we will consider the particular case of problems P_{EC} such that the lower level problem $Q(x,y)$ is a convex problem. More precisely throughout this section we will always assume:

CC-1: For any (x, y), $\phi(x, y, z)$ is a strictly convex function of z.

CC-2: For any $x \in \mathbb{R}^n$, $i = 1, \ldots, l$, $v_i(x, z)$ is a concave function of z.

CC-3: For any $x \in \mathbb{R}^n$, LICQ holds for all points $y \in Y(x)$.

Under these assumptions, the problems P_{EC} and P_{KKTEC} are equivalent in the sense that $(\overline{x}, \overline{y}) \in M_{EC}$ if and only if there exists some $(\overline{z}, \overline{\lambda}) \in \mathbb{R}^m \times \mathbb{R}^l$ such that $(\overline{x}, \overline{y}, \overline{z}, \overline{\lambda}) \in M_{\text{KKTEC}}$, see problems (6.1.1) and (6.3.1).

We want to point out that the previous conditions exclude the bad behavior of Example 6.3.1. We will show that the LICQ assumption CC-3 for $Y(x)$ will assure that for almost every linear perturbation of $(g_1, \ldots, g_q)$ *MPCC-LICQ* is satisfied on M_{KKTEC}.

Proposition 6.4.1 *Let the functions* $(\hat{g}_1, \ldots, \hat{g}_q, \hat{v}_1, \ldots, \hat{v}_l) \in [C^\infty]^{q+l}_{n+m}$*, and* $\hat{\phi} \in [C^\infty]^1_{n+m+m}$ *be fixed such that the conditions CC-1, CC-2, CC-3 hold. Then for almost every* $(d_g, b_\phi, c_\phi) \in \mathbb{R}^q \times \mathbb{R}^m \times \mathbb{R}$*, the* MPCC-LICQ *condition holds for all points of the feasible set* M_{KKTEC} *corresponding to the functions* $\phi = \hat{\phi} + b_\phi z + c_\phi$*,* $(g_1, \ldots, g_q) = ((\hat{g}_1, \ldots, \hat{g}_q) + d_g)$ *and* $(v_1, \ldots, v_l) = (\hat{v}_1, \ldots, \hat{v}_l)$*.*

Proof. Let $(x, y, z, \lambda) \in M_{\text{KKTEC}}$ satisfy $\phi(x, y, z) = 0$. Then, (x, y, z, λ) solves the following system of equations:

$$\begin{array}{rcll} g_j(x, y) & = & 0, & j \in J_{0g}, \\ v_i(x, y) & = & 0, & i \in J_{0v}, \\ \nabla_z \phi(x, y, z) - \sum\limits_{i=1}^{l} \lambda_i \nabla_z v_i(x, z) & = & 0, & \\ v_i(x, z) & = & 0, & i \in J_0 \cup J\Lambda_0, \\ \lambda_i & = & 0, & i \in J\Lambda_0 \cup \Lambda_0, \\ \phi(x, y, z) & = & 0, & \end{array} \tag{6.4.1}$$

for active index sets $J_0 = J_0(x, z, \lambda)$, $J\Lambda_0 = J\Lambda_0(x, z, \lambda)$, $\Lambda_0 = \Lambda_0(x, z, \lambda)$, $J_{0v} = J_{0v}(x, y)$ and $J_{0g} = J_{0g}(x, y)$, see (6.3.2). W.l.o.g. we assume $J_0 = \{1, \ldots, l_1\}$, $J\Lambda_0 = \{l_1{+}1, \ldots, l_2\}$, $\Lambda_0 = \{l_2{+}1, \ldots, l\}$ and $J_{0g} = \{1, \ldots, q_1\}$. The Jacobian matrix of the system (6.4.1) takes the form

$$\begin{array}{ccccccc} \partial_x & \partial_y & \partial_z & \partial_\lambda & \partial_{d_g} & \partial_{b_\phi} & \partial_{c_\phi} \\ \otimes & \otimes & 0 & 0 & I_{q_1}|0 & 0 & 0 \\ \otimes & \nabla_y^T v_{J_{0v}}(x, y) & 0 & 0 & 0 & 0 & 0 \\ \otimes & \otimes & \otimes & \otimes & 0 & I_m & 0 \\ \otimes & 0 & \nabla_z^T v_{J_0 \cup J\Lambda_0} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \mid I_{l-l_1} & 0 & 0 & 0 \\ \otimes & \otimes & \otimes & 0 & 0 & \otimes & 1 \end{array} \tag{6.4.2}$$

Since by assumption CC-3, $[\nabla_y v_{J_{0v}}(x, y)]$ has rank $|J_{0v}|$ and $[\nabla_z v_{J_0 \cup J\Lambda_0}(x, z)]$ is of rank $|J_0 \cup J\Lambda_0|$, the matrix in (6.4.2) has full row rank. So, by the Parameterized

Sard Lemma, for almost every d_g, b_ϕ, c_ϕ, the matrix composed by the columns corresponding to $\partial_{x,y,z,\lambda}$ has full row rank at all solutions (x, y, z, λ) of the system (6.4.1). Consequently, the gradients of the active constraints (*i.e.* the rows corresponding to $\partial_{x,y,z,\lambda}$ are linearly independent.
For the case $\phi(x, y, z) > 0$, the same result follows after considering the system (6.4.1) without the condition $\phi(x, y, z) = 0$.
Finally by taking all possible combinations of the active index sets we recognize that, for almost every (d_g, b_ϕ, c_ϕ), the condition *MPCC-LICQ* holds at all feasible points of the corresponding perturbed problem.

$\square$

Remark 6.4.1 *The result of Proposition 6.4.1 can be obtained for functions $\hat{\phi} \in [C^3]^1_{n+m+m}$, $(\hat{g}_1, \ldots, \hat{g}_q,) \in [C^2]^q_{n+m}$, $(\hat{v}_1, \ldots, \hat{v}_l) \in [C^3]^l_{n+m}$. To do so, a larger number of parameters has to be taken into account, cf. Lemma 2.3.1.*

We are now interested in proving that *MPCC-SC* and *MPCC-SOC* are satisfied at all critical points of P_{KKTEC} for generic perturbations of the objective function.

Theorem 6.4.1 *Let the problem functions $(\hat{f}, \hat{g}, \hat{v}) \in [C^\infty]^{1+q+l}_{n+m}$, $\hat{\phi} \in [C^\infty]^1_{n+m+m}$ be fixed such that MPCC-LICQ holds at all feasible points of M_{KKTEC}. Then for almost every $b \in \mathbb{R}^{n+m}$, the conditions* MPCC-SC *and* MPCC-SOC *hold at all critical points (x, y, z, λ) of P_{KKTEC}, defined by the perturbed functions $f(x, y) = \hat{f}(x, y) + b^T(x, y)$, $(g, \phi, v) = (\hat{g}, \hat{\phi}, \hat{v})$.*

Proof. Let (x, y, z, λ) be a critical point of problem (6.3.1) with $\phi(x, y, z,) = 0$. Then it satisfies the feasibility conditions given in (6.4.1) and for some $(\alpha, \beta, \gamma, \Delta, \delta, \mu)$, it solves the system:

$$\begin{pmatrix} \nabla_x f \\ \nabla_y f \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \nabla_x \nabla_z \phi - \sum\limits_{i=1}^{l} \lambda_i \nabla_x \nabla_z v_i & \nabla_x v_{J_0 \cup J\Lambda_0} & 0 & \nabla_x \phi & \nabla_x v_{J_{0v}} & \nabla_x g_{J_{0g}} \\ \nabla_y \nabla_z \phi & 0 & 0 & \nabla_y \phi & \nabla_y v_{J_{0v}} & \nabla_y g_{J_{0g}} \\ \nabla^2_{zz} \phi - \sum\limits_{i=1}^{l} \lambda_i \nabla^2_{zz} v_i & \nabla_z v_{J_0 \cup J\Lambda_0} & 0 & \nabla_z \phi & 0 & 0 \\ \begin{matrix} -\nabla_z^T v_1 \\ \vdots \\ -\nabla_z^T v_l \end{matrix} & 0 & \begin{matrix} 0 \\ 0 | I_{J\Lambda_0 \cup \Lambda_0} \end{matrix} & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \Delta \\ \delta \\ \mu \end{pmatrix} \tag{6.4.3}$$

with active index sets $J_0, J\Lambda_0, \Lambda_0, J_{0v}, J_{0g}$, see (6.3.2). For the case $\phi(x, y, z) > 0$, we simply assume $\Delta = 0$.

We will only sketch the main ideas of the proof. Let us fix the active index sets $J_0, J\Lambda_0, \Lambda_0, J_{0v}, J_{0g}$. Firstly, by contradiction, we show that for almost every b, there is no critical point such that $\gamma_i = 0$ holds for some $i \in J\Lambda_0$. To do so, let us consider a critical point with $J\Lambda_0^0 := \{i \in J\Lambda_0 \mid \gamma_i = 0\}$. From system (6.4.3), we find $-\nabla_z v_i(x, z)\alpha = 0$, $i \in J_0 \cup J\Lambda_0^0$ and $-\nabla_z v_i(x, z)\alpha + \gamma_i = 0$,

$i \in [J\Lambda_0 \setminus J\Lambda_0^0] \cup \Lambda_0$. The variables γ_i can be eliminated and we can equivalently replace the system:

$$\begin{array}{rcl} -\nabla_z v_i(x,z)\alpha & = & 0, \;\; i \in J_0 \cup J\Lambda_0^0 \\ -\nabla_z v_i(x,z)\alpha + \gamma & = & 0, \;\; i \in [J\Lambda_0 \setminus J\Lambda_0^0] \cup \Lambda_0 \end{array}$$

in (6.4.3) by:

$$-\nabla_z v_i(x,z)\alpha = 0, \; i \in J_0 \cup J\Lambda_0^0.$$

Considering this reduced system (with the perturbed function $f = \hat{f} + b^T(x,y)$) and the feasibility condition (6.4.1), it can be proven that, under the hypotheses of the theorem, the Jacobian with respect to variables (x, y, z, λ), multipliers $(\alpha, \beta, \Delta, \delta, \mu)$ and the parameter b has full row rank. Then, by the Parameterized Sard Lemma, it follows that for almost every b the Jacobian matrix with respect to $\partial_{x,y,z,\lambda,\alpha,\beta,\Delta,\mu,\delta}$ has full row rank at all solutions of the (reduced) system. This means that the number of rows has to be smaller than or equal to the number of columns and this fact appears to be equivalent to $J\Lambda_0^0 = \varnothing$. So, for almost every b, the set $J\Lambda_0^0$ is empty as we wanted to prove.

Now we prove the fulfillment of the MPCC-SC. Let us assume that there are some multipliers $\beta_i = 0$, $\delta_k = 0$ and $\mu_j = 0$, for some $i \in J\Lambda_0$, $k \in J_{0v}$ and $j \in J_{0g}$ respectively. W.l.o.g. we take the active index sets as before and assume $\beta_i = 0$, $i \in J\Lambda_{0\beta}^0 \subset J\Lambda_0$, $\delta_k = 0$, $k \in J_{0v}^0 \subset J_{0v}$ and $\mu_j = 0$, $j \in J_{0g}^0 \subset J_{0g}$. If we consider the whole system composed by these conditions, the critical point equations and the feasibility conditions (see (6.4.3) and (6.4.1) respectively) it can be seen that its Jacobian with respect to variables, multipliers and the parameter b has also full row rank. The application of the Parameterized Sard Lemma means that for almost every b the Jacobian matrix with respect to $\partial_{x,y,z,\lambda,\alpha,\beta,\Delta,\mu,\delta}$ has full row rank at all solutions of the considered system. But the number of rows has to be smaller than or equal to the number of columns and by comparing their dimension it follows $\Delta \neq 0$ if $\phi(x,y,z) = 0$ and $J\Lambda_{0\beta}^0 = J_{0v}^0 = J_{0g}^0 = \varnothing$. So, the MPCC-SC condition must be valid. Moreover for almost all perturbations $b^T(x,y)$ of f, the sub-matrix of the Jacobian with columns corresponding to $\partial_{x,y,z,\lambda,\alpha,\beta,\gamma,\Delta,\mu,\delta}$ is non-singular. This fact together with Proposition 2.1.1, implies that the Hessian matrix of problem (6.3.1) restricted to the tangent subspace is regular. This means that *MPCC-SOC* is satisfied. Finally, by taking all possible combinations of active index sets into account, the result is proven.

□

6.5 Linear equilibrium constrained problems

In this section we will study a special case of P_{EC} in which the KKT approach leads to a MPCC-problem with only linear functions. We consider a P_{EC} program

(6.1.1) with linear functions $f, v_i, i = 1, \ldots, l$, $\phi(x, y, z) = [C(x, y) + d]^T(z - y)$ and $q = 0$,

$$P_{LEC}: \quad \begin{array}{c} \min c^T(x,y) \\ \text{s.t. } (x,y) \in M_{LEC} \end{array} \tag{6.5.1}$$

$$M_{LEC} = \left\{ (x,y) \in \mathbb{R}^n \times \mathbb{R}^m \;\middle|\; \begin{array}{rcll} [C(x,y)+d]^T(v-y) & \geq & 0, & \forall v \in Y(x), \\ y & \in & Y(x), & \end{array} \right\}$$

$$Y(x) = \{y \in \mathbb{R}^m \mid B(x,y) \geq b\}$$

where $c \in \mathbb{R}^n \times \mathbb{R}^m$, $B \in \mathbb{R}^{l\times(n+m)}$, $b \in \mathbb{R}^l$, $C \in \mathbb{R}^{m\times(n+m)}$ and $d \in \mathbb{R}^m$. This problem will be called linear equilibrium constrained problem, LEC for short. Taking into account that $\phi(x, y, y) = 0$ holds, the KKT approach leads to the problem (see also Section 1.2)

$$P_{\text{KKTLEC}}: \quad \begin{array}{c} \min c^T z \\ \text{s.t. } (z,\lambda) \in M_{\text{KKTLEC}} \end{array} \tag{6.5.2}$$

$$M_{\text{KKTLEC}} = \left\{ (z,\lambda) \in \mathbb{R}^{n+m+l} \;\middle|\; \begin{array}{rcl} Cz + d - BY^T\lambda & = & 0, \\ Bz & \geq & b, \\ \lambda & \geq & 0, \\ (Bz-b)^T\lambda & = & 0, \end{array} \right\}$$

Here we use the abbreviation $z = (x, y) \in \mathbb{R}^n \times \mathbb{R}^m$ and $BY \in \mathbb{R}^{l\times m}$ is the submatrix of B corresponding to the variable y.

Obviously, problem (6.5.2) can be seen as a MPCC program (4.1.2) where the functions (f, h, g, r, s) are linear. In this case, the lower level problem

$$Q(x,y): \quad \begin{array}{c} \min\limits_v (C(x,y)+d)^T(v-y) \\ \text{s.t. } v \in Y(x) \end{array}$$

is a linear problem in v and so

$$z \in M_{LEC} \Leftrightarrow (z,\lambda) \in M_{\text{KKTLEC}} \text{ for some } \lambda \in \mathbb{R}^l.$$

The problem P_{LEC} has also been studied earlier in [7]. In this paper the structure of the feasible set of (6.5.1) was investigated in the generic case. In the present section we analyze this problem from the viewpoint of the corresponding complementarity constraint program P_{KKTLEC} in (6.5.2).

In the next subsections we firstly present a genericity analysis of problem (6.5.2) and based on this analysis, we then describe a solution method.

6.5.1 Genericity results.

We will now prove that generically with respect to the problem data (c, B, b, C, d), the condition *MPCC-LICQ* holds at all feasible points of P_{KKTLEC}. We also give conditions, under which, generically the local minimizers of P_{KKTLEC} are of order 1 with respect to the z variable.

As before, the active index sets are defined as:

$$
\begin{array}{rcl}
J_0(\overline{z}, \overline{\lambda}) & = & \left\{ i \mid [B\overline{z}]_i = b_i,\ \overline{\lambda}_i > 0 \right\} \\
J\Lambda_0(\overline{z}, \overline{\lambda}) & = & \left\{ i \mid [B\overline{z}]_i = b_i,\ \overline{\lambda}_i = 0 \right\} \\
\Lambda_0(\overline{z}, \overline{\lambda}) & = & \left\{ i \mid [B\overline{z}]_i > b_i,\ \overline{\lambda}_i = 0 \right\}
\end{array}
$$

The matrices B_{J_0}, $B_{J\Lambda_0}$ and B_{Λ_0} denote submatrices of B composed by the rows of B with indices in J_0, $J\Lambda_0$ and Λ_0 respectively.

Proposition 6.5.1 *In problem (6.5.2) let (c, B, C) be fixed. Then for almost all $(b, d) \in \mathbb{R}^{l+m}$ the MPCC-LICQ condition is fulfilled at all feasible points.*

Proof. The proof is similar to that of Lemma 5 in [53].

Consider a feasible point $(\overline{z}, \overline{\lambda})$ of (6.5.2). Then with active index sets $J_0 = J_0(\overline{z}, \overline{\lambda})$, $J\Lambda_0 = J\Lambda_0(\overline{z}, \overline{\lambda})$, $\Lambda_0 = \Lambda_0(\overline{z}, \overline{\lambda})$ the following system is satisfied:

$$
\begin{array}{rcl}
C\overline{z} + d - BY^T\overline{\lambda} & = & 0 \\
B_{J_0}\overline{z} - b_{J_0} & = & 0 \\
B_{J\Lambda_0}\overline{z} - b_{J\Lambda_0} & = & 0 \\
\overline{\lambda}_{J\Lambda_0} & = & 0 \\
\overline{\lambda}_{\Lambda_0} & = & 0
\end{array}
\tag{6.5.3}
$$

The Jacobian of this system w.r.t. (z, λ) and the perturbation parameters (b, d) is of the form:

$$
\begin{array}{ccccc}
\partial_z & \partial_\lambda & \multicolumn{2}{c}{\partial_b} & \partial_d \\
C & -BY^T & \multicolumn{2}{c}{0} & I_m \\
B_{J_0} & 0 & -I_{J_0} & 0 & 0 \\
B_{J\Lambda_0} & & 0 & -I_{J\Lambda_0} & \\
0 & I_{J\Lambda_0 \cup \Lambda_0} & \multicolumn{2}{c}{0} & 0
\end{array}
$$

This matrix has full row rank. So, from the Parameterized Sard Lemma, Lemma 2.3.1, for almost all (b, d) the Jacobian of the (reduced) system with respect to the variables (z, λ) has full row rank

$$
m + |J_0| + \mid J\Lambda_0 \mid + \mid J\Lambda_0 \mid + \mid \Lambda_0 \mid = m + l + \mid J\Lambda_0 \mid
$$

at all solution points of system (6.5.3). But this Jacobian has the gradients of the active constraints as rows, *i.e., MPCC-LICQ* holds. Considering all finitely many possible combinations of active index sets $J_0, J\Lambda_0, \Lambda_0$, the result is proven. $\square$

Remark 6.5.1 *In Proposition 6.5.1 we have shown that MPCC-LICQ holds generically without the assumption of any extra constraint qualification in the set* $\{z \in \mathbb{R}^{n+m} \mid Bz \geq b\}$ *as needed in the general case (cf. condition CC-3 in Proposition 6.4.1). It appears that, MPCC-LICQ will also hold generically without this constraint qualification if the function* ϕ *satisfies the condition* $\phi(x, y, y) = 0$*, for all* $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$*.*

Next we are going to examine the generical properties of the critical points of problem (6.5.2). In view of the analysis above, we can generically assume that the *MPCC-LICQ* condition holds for the feasible set.

Let $(\overline{z}, \overline{\lambda})$ be a critical point of the problem (6.5.2). Then, for some $(\alpha, \beta, \gamma) \in \mathbb{R}^{m+|J_0 \cup J\Lambda_0|+|J\Lambda_0 \cup \Lambda_0|}$ it solves with active index sets $J_0, J\Lambda_0, \Lambda_0$ the following system:

$$\begin{array}{rcl} \begin{pmatrix} c \\ 0 \end{pmatrix} - \begin{pmatrix} C^T \\ -BY \end{pmatrix} \alpha - \begin{pmatrix} [B^T_{J_0 \cup J\Lambda_0}] \\ 0 \end{pmatrix} \beta - \begin{pmatrix} 0 \\ I_{J\Lambda_0 \cup \Lambda_0} \end{pmatrix} \gamma & = & 0 \\ C\overline{z} + d - BY^T\overline{\lambda} & = & 0 \\ B_{J_0}\overline{z} - b_{J_0} & = & 0 \\ B_{J\Lambda_0}\overline{z} - b_{J\Lambda_0} & = & 0 \\ \overline{\lambda}_{J\Lambda_0 \cup \Lambda_0} & = & 0. \end{array} \tag{6.5.4}$$

We are going to investigate whether generically at a critical point of P_{KKTLEC} the conditions MPCC-SC and MPCC-SOC hold. Since the problem functions in P_{KKTLEC} are linear, MPCC-SOC means that the active gradients span the whole $\mathbb{R}^{n+m+l}$.

Note that if $[C_{n+1}, \ldots, C_{n+m}]$ is a symmetric matrix, the problem P_{KKTLEC} is equivalent to the KKT-formulation of the bilevel problem:

$$\begin{array}{ll} & \min c^T(x, y) \\ \text{s.t. } y \text{ solves } Q(x): & \min\limits_y y^T[C_1, \ldots, C_n]x + \frac{1}{2}y^T[C_{n+1}, \ldots, C_{n+m}]y + d^T y \\ & \qquad \text{s.t. } B(x, y) \geq b. \end{array}$$

Applying Theorem 5.2.2 we conclude that, for almost all linear perturbations of the functions $c^T(x, y)$, $y^T[C_1, \ldots, C_n]x + \frac{1}{2}y^T[C_{n+1}, \ldots, C_{n+m}]y + d^T y$ and

$B(x,y)-b$, the critical points of the corresponding problem P_{KKTLEC}, $(\overline{z},\overline{\lambda})$, satisfy one of the following two conditions (corresponding to $\alpha \neq 0$ and $\alpha = 0$):

- $(\overline{z},\overline{\lambda})$ is a non-degenerate critical point in the MPCC-sense, see Definition 4.3.3.

- If $(\overline{z},\overline{\lambda})$ is a critical point where the multiplier α corresponding to $C^T z + d - BY^T\lambda = 0$ is the zero vector, then $\overline{z}$ is a non-degenerate critical point of $\min c^T z$ s.t. $Bz \geq b$ and there is some critical point $(\overline{z},\lambda^*)$ of P_{KKTLEC} such that $rank(BY_{J_0(\overline{z},\lambda^*)}(\overline{z})) = |J_0(\overline{z},\lambda^*)| = m$.

Introducing the polyhedron

$$\mathcal{R} = \left\{ (z,\lambda) \in \mathbb{R}^{n+m} \times \mathbb{R}^l \;\middle|\; \begin{array}{rcl} Bz & \geq & b, \\ Cz + d - BY^T\lambda & = & 0, \\ \lambda & \geq & 0. \end{array} \right\} \tag{6.5.5}$$

we can prove our next result.

Theorem 6.5.1 *For almost every (c,d,B,b), any critical point $(\overline{z},\overline{\lambda})$ of problem (6.5.2) (see (6.5.4)) with associated multipliers (α,β,γ) is of one of the following two types:*

LEC-1: If $\alpha \neq 0$ then MPCC-LICQ, MPCC-SC and MPCC-SOC are fulfilled w.r.t. the corresponding problem (6.5.2) and $(\overline{z},\overline{\lambda})$ is a non-degenerate vertex of the polyhedron $\mathcal{R}$.

LEC-2: If $\alpha = 0$ then $rank[BY_{J_0(\overline{z},\overline{\lambda})\cup J\Lambda_0(\overline{z},\overline{\lambda})}] = m$ and $\beta_i \neq 0$ for all $i \in J\Lambda_0(\overline{z},\overline{\lambda}) \cup J_0(\overline{z},\overline{\lambda})$. Moreover, there is some λ^* such that $rank[BY_{J_0(\overline{z},\lambda^*)}] = m = |J_0(\overline{z},\lambda^*)|$. If λ^* is such that $rank[BY_{J_0(\overline{z},\lambda^*)}] = |J_0(\overline{z},\lambda^*)| = m$, then MPCC-SOC holds and $(\overline{z},\lambda^*)$ is a non-degenerate vertex of the polyhedron $\mathcal{R}$.

Proof. The proof follows the same lines as the proof of the related Theorem 5.2.2. Here we will only present the main ideas and the consequences of conditions BL-1 and BL-2 in this case.
For critical points $(\overline{z},\overline{\lambda})$ with associated multiplier $\alpha \neq 0$, as in Theorem 5.2.2, we obtain that MPCC-LICQ, MPCC-SC and MPCC-SOC are fulfilled for almost all data (c,d,B,b). These conditions and the linearity of the involved functions imply that the active constraints generate the whole space $\mathbb{R}^{n+m+l}$. Together with the fulfillment of the MPCC-LICQ condition it follows that the number of active constraints is exactly $n+m+l$. So $(\overline{z},\overline{\lambda})$ is a non-degenerate vertex of $\mathcal{R}$.
If $\alpha = 0$ holds at the critical point $(\overline{z},\overline{\lambda})$, we can analogously, prove that for almost every (c,d,B,b) the vector $\overline{z}$ is a non-degenerate critical point of the problem:

$$\begin{array}{c} \min c^T z \\ \text{s.t. } Bz \geq b \end{array}$$

and that there exists a corresponding solution $\lambda = \overline{\lambda}$ of the system:

$$\begin{array}{rcl} C\overline{z} + d - BY^T_{[J_0 \cup J\Lambda_0](\overline{z})}\lambda & = & 0, \\ \lambda & \geq & 0, \end{array} \tag{6.5.6}$$

with $rank(BY_{J_0(\overline{z},\overline{\lambda})} = |J_0(\overline{z},\overline{\lambda})| = m$. Moreover, for these points the *MPCC-SOC* condition is satisfied.

Now we prove that, under these generic conditions, the critical point $(\overline{z},\overline{\lambda})$ is a vertex of $\mathcal{R}$. Again since BY_{J_0} is a regular matrix, the MPCC-SOC is equivalent to the regularity of matrix

$$\begin{pmatrix} 0 & B^T_{J_0 \cup J\Lambda_0} \\ B_{J_0 \cup J\Lambda_0} & 0 \end{pmatrix}$$

This implies $|J_0 \cup J\Lambda_0| = n+m$. As $|J_0| = m$ holds, it follows $|J\Lambda_0| = n$ and $(\overline{z},\overline{\lambda})$ is a point where $n+m+l$ constraints are active. So, $(\overline{z},\overline{\lambda})$ is a non-degenerate vertex of $\mathcal{R}$.

□

Now we consider local minimizers $(\overline{z},\overline{\lambda})$ in the generic situation of Theorem 6.5.1 and apply the results obtained in Section 5.2 to the present problem P_{KKTLEC}.

Corollary 6.5.1 *A local minimizer $(\overline{z},\overline{\lambda})$ of problem P_{KKTLEC} where MPCC-LICQ, MPCC-SC and MPCC-SOC holds is a local minimizers of order 1.*

Proof. From the proof of Theorem 6.5.1 it follows that $(\overline{z},\overline{\lambda})$ is a non degenerate critical vertex of $\mathcal{R}$, *i.e.* precisely $n+m+l$ gradients are active. In view of MPCC-SC, the assumptions of Theorem 4.4.3 are satisfied.

□

Under the assumptions of Corollary 6.5.1, for the original problem P_{LEC} in (6.5.1) we obtain the following result.

Corollary 6.5.2 *Let $(\overline{z},\overline{\lambda})$ be a local minimizer of problem P_{KKTLEC} where MPCC-LICQ, MPCC-SC and MPCC-SOC hold. Then z is a local minimizers of order 1 of problem P_{LEC}.*

Proof. Arguing as in the proof of Corollary 5.2.3, there is a neighborhood $V(\overline{z}) \times V(\overline{\lambda})$ of $(\overline{z},\overline{\lambda})$ such that for any $z \in M_{LEC} \cap V(\overline{z})$ we have $(z,\lambda) \in M_{\text{KKTLEC}} \cap [V(\overline{z}) \times V(\overline{\lambda})]$ for some $\lambda \in \mathbb{R}^l$. But by Corollary 6.5.1, the point $(\overline{z},\overline{\lambda})$ is a local minimizer of order 1 of problem P_{KKTLEC}. So, there is some $\kappa,\ \kappa > 0$ such that for all $(z,\lambda) \in M_{\text{KKTLEC}} \cap [V(\overline{z}) \times V(\overline{\lambda})]$,

$$c^Tz - c^T\overline{z} \geq \kappa \left[\|z-\overline{z}\| + \|\lambda - \overline{\lambda}\|\right] \geq \kappa\|z-\overline{z}\|.$$

This shows that $\overline{z}$ is a local minimizer of order 1 for problem P_{LEC}.

□

In the generic case, the assumptions of the previous corollaries hold for a local minimizer $(\overline{z}, \overline{\lambda})$ if the associated multiplier satisfies $\alpha \neq 0$. In the case of a local minimizer $(\overline{z}, \overline{\lambda})$ with $\alpha = 0$, there exists $\lambda^* \geq 0$ solving system (6.5.6) and $rank(BY_{J_0(\overline{z}, \lambda^*)}) = |J_0(\overline{z}, \lambda^*)| = m$. The point $(\overline{z}, \lambda^*)$ is a critical point and a vertex of $\mathcal{R}$, but as in the situation of bilevel problems (see Chapter 5) it may fail to be a local minimizer of problem P_{KKTLEC}. Moreover the set of local minimizers may be non-closed. The following example illustrates these facts, see also Example 5.2.1:

$$\begin{aligned} &\min -x + y \\ &\text{s.t. } v - y \geq 0, \ \forall v \in Y(x) \end{aligned}$$

$$Y(x) = \left\{ y \;\middle|\; \begin{array}{rcl} y & \geq & 0, \\ x & \geq & 0. \end{array} \right\}$$

By writing $(z_1, z_2) = z = (x, y)$, the corresponding P_{KKTLEC} problem is of the form:

$$\begin{array}{rcl} \min -z_1 + z_2 & & \\ 1 - \lambda_2 & = & 0 \\ z_1 & \geq & 0 \\ z_2 & \geq & 0 \\ \lambda_1 & \geq & 0 \\ \lambda_2 & \geq & 0 \\ z_1\lambda_1 & = & 0 \\ z_2\lambda_2 & = & 0 \end{array} \tag{6.5.7}$$

and coincides with the MPCC problem corresponding to the KKT relaxation of the BL problem in Example 5.2.1. As already analyzed in that example, the points $(0, 0, \lambda_1, 1),\ \lambda_1 > 0$ are local minimizers of the problem (6.5.7) with $\alpha = 0$. However, $(0, 0, 0, 1)$ is not a local minimizer. Note that the set of local minimizers is non-closed.

If $(\overline{z}, \overline{\lambda})$ is a vertex of $\mathcal{R}$ and a local minimizer of the problem P_{KKTLEC} with $\alpha = 0$, by applying the same arguments as in the proof of Corollary 5.2.5 we obtain that generically $\overline{z}$ is a non-degenerate local minimizer of the corresponding linear problem (see (5.2.20) for the general case)

$$P_{J_0(\overline{z}, \overline{\lambda})}: \qquad \begin{aligned} &\min c^T z \\ &\text{s.t. } z \in M_{J_0(\overline{z}, \overline{\lambda})}, \end{aligned} \tag{6.5.8}$$

$$M_{J_0(\overline{z}, \overline{\lambda})} = \left\{ z \in \mathbb{R}^{n+m} \;\middle|\; \begin{array}{rcl} B_{J\Lambda_0(\overline{z}, \overline{\lambda})} z & \geq & b_{J\Lambda_0(\overline{z}, \overline{\lambda})}, \\ B_{J_0(\overline{z}, \overline{\lambda})} z & = & b_{J_0(\overline{z}, \overline{\lambda})}. \end{array} \right\}$$

So, generically, $\overline{z}$ is a local minimizer of order 1 of problem (6.5.8). Again by Corollary 5.2.5, for all z near $\overline{z}$ with $(z,\lambda) \in M_{\text{KKTLEC}}$ for some $\lambda \in \mathbb{R}^l$, we can conclude that z is feasible for (6.5.8) such that locally $c^T z - c^T \overline{z} \geq \kappa\|z - \overline{z}\|$ holds. As a consequence we obtain:

Corollary 6.5.3 *Generically all local minimizers $(\overline{z}, \overline{\lambda})$ of P_{KKTLEC} with multiplier $\alpha = 0$ and such that $(\overline{z}, \overline{\lambda})$ is a vertex of $\mathcal{R}$, are local minimizers of order 1. Moreover $\overline{z}$ is a local minimizer of problem $P_{J_0(\overline{z},\overline{\lambda})}$ in (6.5.8).*

Even under the conditions of Corollary 6.5.3 we can not guarantee that $\overline{z}$ is a local minimizer of the original problem P_{LEC} as shown in the following example:

Example 6.5.1

$$\begin{array}{rrcl}
\min x_1 + x_2 - x_3 - 2y & & & \\
s.t. & x_1 + y & \geq & 0, \\
& x_2 - y & \geq & 0, \\
& x_3 + y & \geq & 0, \\
& x_3 & \geq & 0,
\end{array}$$

$$(v - y) \geq 0, \ \forall v \in Y(x) = \left\{ v \in \mathbb{R} \ \middle| \ \begin{array}{rcl} x_1 + v & \geq & 0, \\ x_2 - v & \geq & 0, \\ x_3 + v & \geq & 0. \end{array} \right\}$$

In this case the problem P_{KKTLEC} is (we again set $z = (x, y)$):

$$\begin{array}{lrrcl}
P_{\text{KKTLEC}} & & \min z_1 + z_2 - z_3 - 2z_4 & & \\
& \text{s.t.} & 1 - \lambda_1 + \lambda_2 - \lambda_3 & = & 0, \\
& & z_1 + z_4 & \geq & 0, \\
& & z_2 - z_4 & \geq & 0, \\
& & z_3 + z_4 & \geq & 0, \\
& & z_3 & \geq & 0, \\
& & \lambda_1, \lambda_2, \lambda_3, \lambda_4 & \geq & 0, \\
& & \lambda_1(z_1 + z_4) & = & 0, \\
& & \lambda_2(z_2 - z_4) & = & 0, \\
& & \lambda_3(z_3 + z_4) & = & 0, \\
& & z_3\lambda_4 & = & 0.
\end{array}$$

As can easily be seen, the vertex $(\overline{z}, \lambda) = (0,0,0,0,0,0,1)$ of $\mathcal{R}$ is a local minimizer of P_{KKTLEC}. However the vertex $(\overline{z}, \overline{\lambda}) = (0,0,0,0,1,0,0)$ with the same z value, is not. Note that, in fact, the point $(0,0,0,0)$ is not a local minimizer of the original problem, since the points $(x,y) = (0,0,x_3,0)$, $x_3 > 0$, are also feasible points of P_{LEC} with smaller objective function value.

This example shows the possible bad behavior of the local minimizers due to the disjunctive structure of the set M_{KKTLEC}. However we can state the following result

Proposition 6.5.2 *Let for all vertex solutions $\overline{\lambda}$ of system (6.5.6) $(\overline{z}, \overline{\lambda})$ be a local minimizer of P_{KKTLEC} with multipliers (α, β, γ), $\alpha = 0$, $\gamma = 0$. Let us assume that BY_{I_0} is a regular matrix for all $I_0 \subset \{1, \ldots, l\}$, $|I_0| = m$ (which is generically fulfilled). Then $\overline{z}$ is a local minimizer of P_{LEC} of order 1 i.e. for all $z \in M_{LEC}$ near $\overline{z}$, it holds that $c^T z - c^T \overline{z} \geq \kappa \|z - \overline{z}\|$ for some $\kappa > 0$.*

Proof. Let

$$\mathcal{J}_0(\overline{z}) = \{J_0 \subset \{1, \ldots, l\} | \ J_0 = J_0(\overline{z}, \lambda^*) \text{ for some vertex solution } \lambda^* \text{ of (6.5.6)}\}.$$

If $z_k \to \overline{z}$, with $z_k \in M_{LEC}$, by taking associated lower level multiplier λ_k such that $(z_k, \lambda_k) \in M_{\text{KKTLEC}}$ and $BY_{J_0(z_k, \lambda_k)}$ has full rank $|J_0(z_k, \lambda_k)|$, it can be shown that $z_k \in M_{J_0}$, see (6.5.8) for some $J_0 \in \mathcal{J}_0(\overline{z})$. So, around $\overline{z}$,

$$M_{\text{KKTLEC}}|_{\mathbb{R}^{n+m}} \subset \bigcup_{J_0 \in \mathcal{J}_0(\overline{z})} M_{J_0} \tag{6.5.9}$$

holds, where $M_{\text{KKTLEC}}|_{\mathbb{R}^{n+m}}$ is the projection of the set M_{KKTLEC} onto the z-space $\mathbb{R}^{n+m}$.

Let $(\overline{z}, \lambda^*) \in M_{\text{KKTLEC}}$ be such that λ^* is a vertex solution of the system (6.5.6). By assumption it is a local minimizer of order 1 of problem $P_{J_0(\overline{z}, \lambda^*)}$, see Corollary 6.5.3. Since $M_{EC} = M_{\text{KKTLEC}}|_{\mathbb{R}^{n+m}}$, the result now follows from (6.5.9). $\square$

If P_{KKTLEC}, has a global minimizer, we can prove the existence of a vertex $(\overline{z}, \overline{\lambda})$ of $\mathcal{R}$, which is a global minimizer of the problem. Indeed, using the disjunctive analysis we can guarantee that if problem P_{KKTLEC} has a global minimizer $(\overline{z}, \overline{\lambda})$ then, in particular, it should be a minimizer of:

$$\begin{array}{llrcl} P(J\Lambda_0^0) & & & \min c^T z & \\ & \text{s.t.} & B_j z \geq b_j, & \lambda = 0, & j \in J\Lambda_0^0 \cup \Lambda_0, \\ & & B_j z = b_j, & \lambda \geq 0, & j \in J_0 \cup J\Lambda_0 \setminus J\Lambda_0^0, \\ & & Cz + d - BY\lambda = 0, & & \end{array}$$

for all $J\Lambda_0^0 \subset J\Lambda_0(\overline{z}, \overline{\lambda})$. As $P(J\Lambda_0^0)$ is a linear programming problem with a finite solution, there is a vertex of the feasible set which is also a minimizer of $P(J\Lambda_0^0)$ and feasible for problem P_{KKTLEC}. So, it will be a global minimizer of P_{LEC} too.

6.5.2 Algorithm.

In this subsection we will sketch a projection algorithm for solving the MPCC problem P_{KKTLEC} obtained by applying the KKT-approach to the linear equilibrium program P_{LEC}. It is based on the genericity analysis of the previous subsection. Roughly speaking this algorithm proceeds by descent steps in the direction of the projection of the objective function onto a polyhedral subset of the feasible set.

The feasible set can be described in a disjunctive form as

$$M_{\text{KKTLEC}} = \cup_{I_0 \subset \{1,\dots,l\}} M_{I_0}$$

where

$$M_{I_0} = \left\{ (z,\lambda) \in \mathbb{R}^{n+m+l} \;\middle|\; \begin{array}{rcllcll} [Bz]_i &=& b_i, & \lambda_i &\geq& 0, & i \in I_0, \\ [Bz]_i &\geq& b_i, & \lambda_i &=& 0, & i \in \{1,\dots,l\} \setminus I_0, \\ Cz + d - BY\lambda &=& 0. \end{array} \right\}$$

In particular any feasible point $(\overline{z},\overline{\lambda}) \in M_{\text{KKTLEC}}$ will be included in the polyhedron:

$$R(\overline{z},\overline{\lambda}) = \left\{ (z,\lambda) \in \mathbb{R}^{n+m+l} \;\middle|\; \begin{array}{rcllcll} [Bz]_i &=& b_i, & \lambda_i &\geq& 0, & i \in J_0(\overline{z},\overline{\lambda}), \\ [Bz]_i &=& b_i, & \lambda_i &=& 0, & i \in J\Lambda_0(\overline{z},\overline{\lambda}), \\ [Bz]_i &\geq& b_i, & \lambda_i &=& 0, & i \in \Lambda_0(\overline{z},\overline{\lambda}), \\ Cz + d - BY\lambda &=& 0, \end{array} \right\}$$

which is a face of the polyhedron M_{I_0} if $J_0 \subset I_0$ and $\Lambda_0 \subset \{1,\dots,l\} \setminus I_0$.

The idea of the algorithm is the following. First compute a feasible solution (z_0,λ_0). Then, we consider the corresponding active index sets $J_0 = J_0(z_0,\lambda_0)$, $J\Lambda_0 = J\Lambda_0(z_0,\lambda_0)$, $\Lambda_0 = \Lambda_0(z_0,\lambda_0)$ and the associated polyhedron $R(\overline{z},\overline{\lambda})$ of dimension $k = n - |J\Lambda_0|$.

Next we compute the projection s of the objective gradient $-(c,0)$ (in variables (z,λ)) onto $R(\overline{z},\overline{\lambda})$. The new point (z_1,λ_1) is computed by moving in this descent direction s, as long as feasibility is maintained.

At (z_1,λ_1) a new constraint will become active and the set $J\Lambda_0$ will increase because either an index of $J_0(z_0,\lambda_0)$ or of $\Lambda_0(z_0,\lambda_0)$ will enter $J\Lambda_0(z_1,\lambda_1)$. We consider the adjacent face (of M_{I_0}) of dimension k, corresponding to the deactivation of an active constraint. For instance if $i^* \in J_0(z_0,\lambda_0) \cap J\Lambda_0(z_1,\lambda_1)$, in the associated adjacent face of M_{I_0} only the index i^* changes form J_0 to Λ_0. If the projected objective function decreases in this face we repeat the procedure with this k-dimensional face. In the other case, we move to the face of smaller dimension $R(z_1,\lambda_1)$. This is repeated until a vertex $(\overline{z},\overline{\lambda})$ of $R(\overline{z},\overline{\lambda})$ (and hence of $\mathcal{R}$, see (6.5.5)) is reached. Then we try to move to an adjacent vertex with smaller objective value. As the objective function always decreases and there only exists a finite number of faces and vertices, the method must find a local minimizer of P_{KKTLEC} after finitely many steps.

Conceptual algorithm

step 0: Find a feasible solution (z,λ) of P_{KKTLEC}. Compute the active index sets $J_0 = \{i \mid [Bz]_i = b_i, \lambda_i > 0\}$, $J\Lambda_0 = \{i \mid [Bz]_i = b_i, \lambda_i = 0\}$ and $\Lambda_0 = \{i \mid [Bz]_i > b_i, \lambda_i = 0\}$.

step 1: Let $s = (s_z, s_\lambda)$ be the output of procedure *Projection* with index sets $J_0, J\Lambda_0, \Lambda_0$, see (6.5.11).

while $\mid J\Lambda_0 \mid < n$

step 2: Let $\bar{t}$ be the solution of

$$\max \left\{ t \;\middle|\; \begin{array}{rcl} B_{\Lambda_0}[z + \bar{t}s_z] & \geq & b_{\Lambda_0}, \\ \lambda_i + \bar{t}s_{\lambda i} & \geq & 0,\ i \in J_0, \\ t & \geq & 0. \end{array} \right\} \tag{6.5.10}$$

If $\bar{t} < \infty$, update $[z, \lambda] = [z, \lambda] + \bar{t}[s_z, s_\lambda]$. Else exit because the objective function is not bounded.

Consider the following possible cases:

step 3: In case $B_i[z + \bar{t}s_z] = b_i$ for some $i \in \Lambda_0$, find output s of *Projection* with $J_0 = J_0 \cup \{i\}\,,\ J\Lambda_0 = J\Lambda_0,\ \Lambda_0 = \Lambda_0 \setminus \{i\}$.

1. If $s_{\lambda,i} > 0$, update index sets: $J_0 = J_0 \cup \{i\}\,,\ J\Lambda_0 = J\Lambda_0,\ \Lambda_0 = \Lambda_0 \setminus \{i\}$ and goto 2. Else $J_0 = J_0\ J\Lambda_0 = J\Lambda_0 \cup \{i\}\ \Lambda_0 = \Lambda_0 \setminus \{i\}$.
2. Find s resulting from the procedure *Projection* with the new active index sets.

step 4: If $\lambda_i + \bar{t}s_{\lambda i} = 0$ for some $i \in J_0$, find s output of *Projection* with $J_0 = J_0 \setminus \{i\}\,,\ J\Lambda_0 = J\Lambda_0,\ \Lambda_0 = \Lambda_0 \cup \{i\}$

1. If $B_i s > 0$, update the index sets as $J_0 = J_0 \setminus \{i\}\,,\ J\Lambda_0 = J\Lambda_0$, $\Lambda_0 = \Lambda_0 \cup \{i\}$ and goto 2. Else $J_0 = J_0 \setminus \{i\}\ J\Lambda_0 = J\Lambda_0, \cup \{i\}\ \Lambda_0 = \Lambda_0$.
2. Find s resulting from the procedure *Projection* with the new active index sets.

end while.
Let (z, λ) be the computed vertex of $\mathcal{R}$.

step 5 If there exists a neighboring vertex where the objective function decreases, move to it and goto 5. Else stop, a local minimizer of P_{KKTLEC} was obtained.

We add some comments on the different steps of the algorithm.
The procedure *Projection* computes the projection of $[-c, 0_{|J_0|}]$ onto the subspace defined by:

$$\begin{array}{rcl} B_{J_0}s_z & = & 0, \\ B_{J\Lambda_0}s_z & = & 0, \\ Cs_z - BY_{J_0}^T s_\lambda & = & 0. \end{array} \tag{6.5.11}$$

In step 0, we try to find an initial point by solving:

$$\begin{array}{rcl} \min c^T z & & \\ Cz + d & = & 0 \\ Bz & \geq & b \end{array} \tag{6.5.12}$$

The feasible set of this problem is obviously contained in M_{KKTLEC}. Of course, this problem may be infeasible although P_{KKTLEC} may have feasible solutions (take for instance problem (6.5.7)). If a finite solution was found, the algorithm start with $(z, 0)$ as initial feasible point.

Note that generically in (c, d, b) the solutions of problem (6.5.12) are isolated and non-degenerate vertices of the polyhedron

$$\left\{ z \in \mathbb{R}^n \times \mathbb{R}^m \;\middle|\; \begin{array}{rcl} Cz + d & = & 0, \\ Bz & \geq & 0 \end{array} \right\}.$$

So at the starting solution $(z_0, 0)$, generically we will have $|J\Lambda_0(z_0, 0)| = n$ and $(z_0, 0)$ is a vertex of the polyhedron $\mathcal{R}$ and only step 5 will be performed. However we included the other steps for the case that the starting point is computed in a different way.

Roughly speaking, in step 5 at a vertex (x, λ) of $\mathcal{R}$, we first calculate a new direction $s = (s_z, s_\lambda)$ obtained by deactivating one active constraint. The possible active indexes are $i \in J_0 \cup J\Lambda_0$ for $[Bz]_i = b_i$ and $i \in J\Lambda_0 \cup \Lambda_0$ for $\lambda_i = 0$. In each case the new direction is calculated as follows:

Case	$[Bz]_i = b_i, i \in J_0$	$[Bz]_i = b_i,\, i \in J\Lambda_0$
Directions computed by	$\begin{bmatrix} B_{J_0} & 0 \\ B_{J\Lambda_0} & 0 \\ C & -BY_{J_0} \end{bmatrix} s = \begin{bmatrix} e_i^{\vert J_0\vert} \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} B_{J_0} & 0 \\ B_{J\Lambda_0} & 0 \\ C & -BY_{J_0} \end{bmatrix} s = \begin{bmatrix} 0_{\vert J_0\vert} \\ e_i^n \\ 0 \end{bmatrix}$

Case	$\lambda_i = 0,\, i \in J\Lambda_0$	$\lambda_i = 0,\, i \in \Lambda_0$
Directions computed by	$\begin{bmatrix} B_{J_0} & 0 \\ B_{J\Lambda_0} & 0 \\ C & -B_{J_0} \end{bmatrix} s = \begin{bmatrix} 0_{\vert J_0\vert} \\ 0_n \\ BY_i \end{bmatrix}$	$\begin{bmatrix} BY_{J_0} & 0 \\ B_{J\Lambda_0} & 0 \\ C & BY_{J_0} \end{bmatrix} s = \begin{bmatrix} 0 \\ 0 \\ -B_{\Lambda_0 i} \end{bmatrix}$

here e_j^l is the j^{th}-unit vector in $\mathbb{R}^l$.
If in this direction the objective function decreases, *i.e.* $c^T s_z < 0$, and for t^*

solution of the problem (6.5.10) the point $(z+t^* s_z, \lambda+t^* s_\lambda)$ is feasible, we change to this new vertex. If for all possible $n+l$ active constraints this procedure fails, we have found a local minimizer of P_{KKTLEC}.

Chapter 7

Final remarks

In the thesis we have examined the KKT approach for solving variational inequalities, bilevel problems and equilibrium constrained problems.

The KKT approach leads to a corresponding MPCC problem with a particular structure. As a solution method for the resulting MPCC programs, we have studied the parametric smoothing approach P_τ, where τ is the smoothing parameter. We proved that for standard MPCC programs this method converges under generic assumptions with a rate $\mathcal{O}(\sqrt{\tau})$. We have shown that generically the parametric problem P_τ is JJT regular which allows to apply pathfollowing solution strategies. We mainly concentrated on the theoretical basis of the method. Some further practical research will be:

- A practical implementation of the smoothing approach with the aid of pathfollowing strategies.

- Adapt the known pathfollowing methods to track the set of solutions of a one-parametric MPCC around the generic singularities.

For bilevel problems, we found out that unfortunately the MPCC problem obtained by the KKT approach is not MPCC-regular. So we adapted the general smoothing scheme in order to cover all possible candidate minimizers. The KKT approach has the disadvantage that the minimizer of the KKT formulation may not be a solution of the original bilevel problem (even if it is a feasible point). So we can conclude that in practice instead of the KKT relaxation a FJ relaxation should be used. We expect that our structural and genericity results for the KKT approach can directly be carried over to the FJ approach. For the FJ approach, research can be done in the following directions:

- Generic properties of the corresponding MPCC problem.

- Detailed analysis on the relation between the FJ relaxation and the original bilevel problem

- A numerical algorithm for this approach with satisfactory convergence properties under generic conditions.

The general equilibrium constrained program has a more difficult structure than a bilevel problem. As a matter of fact, different from BL, for general equilibrium constrained programs the constraint qualification required for the KKT method is not generically fulfilled. Only for the convex case with extra constraint qualifications in the lower level problem, the MPCC regularity holds generically. In the special linear case we obtained results similar to the bilevel case.

Open questions are:

- The generic structure of general equilibrium constrained problems.

- The case of programs with variational inequality constraints.

- Implementation of the projection algorithm and its comparison with the behavior of the corresponding method for solving linear BL.

Bibliography

[1] B. Bank, J. Guddat, D. Klatte, B. Kummer, and D. Tammer. *Nonlinear Parametric Optimization*. Birkhäuser Verlag, Basel, 1983.

[2] J.F. Bard. Convex two-level optimization. *Mathematical Programming*, 40(1):15–27, 1988.

[3] J.F. Bard. *Practical bilevel optimization. Algorithms and applications*. Nonconvex Optimization and its Applications, 30. Kluwer Academic Publishers, Dordrecht, 1998.

[4] M. S. Bazaraa, H. D. Sherali, and C.M. Shetty. *Non linear programming theory and Algorithms*. John Willey and Sons, 1993.

[5] H. Y. Benson, D. F. Shanno, and R.J. Vanderbei. Interior-point methods for non-convex programming: complementarity constraints. Technical report, Operations Research and Financial Engineering Department, Princeton University, 2002.

[6] H.Y. Benson, A. Sen, D. F. Shanno, and R. J. Vanderbei. Interior-point algorithms, penalty methods and equilibrium problems. Technical report, Operations Research and Financial Engineering, Princeton University, 2003. ORFE-03-02.

[7] S. I. Birbil, G. Bouza, J. B. G. Frenk, and G. Still. Equilibrium constrained optimization problems. *European Journal on Operational Research*, 169(13):1108–1128, 2006.

[8] S. Braun and J.E. Mitchell. A semidefinite programming heuristic for quadratic programming problems with complementarity constraints. *http://www.optimization-online.org/DB_HTML/2002/12/579.html*, 2002.

[9] C. Chen and O.L. Mangasarian. Smoothing methods for convex inequalities and linear complementarity problems. *Mathematical Programming*, 71(1):51–69, 1995.

[10] X. Chen and M. Fukushima. A smoothing method for a mathematical program with P-matrix linear complementarity constraints. *Computational Optimization and Applications*, 27(3):223–246, 2004.

[11] S. Dempe. Annotated bibliography on bilevel programming and mathematical programs with equilibrium constrains. *Optimization*, 52(3):333–359, 2003.

[12] D. Dentcheva, R. Gollmer, J. Guddat, and J.J. Rückmann. Pathfollowing methods in nonlinear optimization. II. Exact penalty methods. In *Approximation & Optimization, 8*, pages 200–230, Frankfurt am Main, 1995. Lang.

[13] F. Facchinei, J. Houyuan, and Q. Liqun. A smoothing method for mathematical programs with equilibrium constraints. *Mathematical Programming*, 85(1):107–134, 1999.

[14] M. Flegel and C. Kanzow. A Fritz John approach to first order optimality conditions for mathematical programming with equilibrium constraints. *Optimization*, 52(3):277–286, 2003.

[15] M. Fukushima, Z.Q Luo, and J.S. Pang. A globally convergent sequential quadratic programming algorithm for mathematical programs with linear complementarity constraints. *Computational Optimization and Applications*, 10(1):5–34, 1998.

[16] M. Fukushima and J.S. Pang. Convergence of a smoothing continuation method for mathematical programs with complementarity constraints. In *Lecture Notes in Economics and Mathematical Systems*, 477, pages 99–110, Berlin, 1999. Springer Verlag.

[17] M. Fukushima and P. Tseng. An implementable active-set algorithm for computing a B-stationary point of a mathematical program with linear complementarity constraints. *SIAM Journal on Optimization*, 12(3):724–739, 2002.

[18] R. Gollmer, U. Kausmann, D. Nowack, K. Wendler, and J. Bacallao Estrada. *Computerprogramm PAFO*. Humboldt-Universitaet, Institut fuer Mathematik, 2004.

[19] W. Gómez. On a generic class of regular one-parametric variational inequalities. *Mathematical Methods of Operations Research*, 51(1):179–201, 2000.

[20] W. Gómez. On generic quadratic penalty embeddings for nonlinear optimization problems. *Optimization*, 50(3-4):279–295, 2001.

[21] W. Gómez, J. Guddat, H. Th. Jongen, J.J. Rückmann, and C.Solano. *Curvas criticas y saltos en la optimizacion no lineal.* http://www.emis.de/monographs/curvas/index.html, 2000.

[22] J. Guddat, F. Guerra, and H. Th. Jongen. *Parametric optimization: singularities, pathfollowing and jumps*. Teubner and John Wiley, Chichester, 1990.

[23] M. Hirsch. *Differential Topology*. Springer Verlag, New York, 1976.

[24] X. Hu and D. Ralph. A note on sensitivity of value functions of mathematical programs with complementarity constraints. *Mathematical Programming, Series A*, 93:265–279, 2002.

[25] X. M. Hu and D. Ralph. Convergence of a penalty method for mathematical programming with complementarity constraints. *Journal of Optimization Theory and Applications*, 123(2):365–390, 2004.

[26] H. Jiang and D. Ralph. Smooth SQP methods for mathematical programs with nonlinear complementarity constraints. *SIAM Journal on Optimization*, 10(3):779–808, 2000.

[27] H.Th. Jongen, P. Jonker, and F. Twilt. Critical sets in parametric optimization. *Mathematical Programing*, 34(3):333–353, 1986.

[28] H.Th. Jongen, P. Jonker, and F. Twilt. On one-parametrer families of optimization problems: Equality constrains. *Journal of Optimization Theory and Applications*, 48(1):141–161, 1986.

[29] H.Th. Jongen, P. Jonker, and F. Twilt. *Nonlinear optimization in finite dimensions. Morse theory, Chebyshev approximation, transversality, flows, parametric aspects*. Nonconvex Optimization and its Applications, 47. Kluwer Academic Publishers, Dordrecht, (2000).

[30] H.Th. Jongen and J.J. Rückmann. On stability and deformation in semi-infinite optimization. In Rembert et al. Reemtsen, editor, *Semi-infinite programming*, number 25 in Nonconvex Optimization Applications, pages 29–67, Boston, 1998. Kluwer Academic Publishers.

[31] H.Th. Jongen and J.J. Rückmann. One-parameter families of feasible sets in semi-infinite optimization. *Journal of Global Optimization*, 14(2):181–203, 1999.

[32] H.Th. Jongen, F. Twilt, and G.-W. Weber. Semi-infinite optimization: structure and stability of the feasible set. *Journal of Optimization Theory and Applications*, 72(3):529–552, 1992.

[33] H.Th. Jongen and G. Zwier. On regular semi-infinite optimization. In *Lecture Notes in Economics and Mathematical Systems*, 259, pages 53–64, 1985.

[34] H.Th. Jongen and G. Zwier. On the local structure of the feasible set in semi-infinite optimization. In Brosowski Deutsch, editor, *Internationale Schriftenreihe zur Numerischen Mathematik*, 72, pages 185–202, Basel, 1985. Birkhäuser.

[35] G. Kassay. On equilibrium problems. Technical report, Babes-Bolyai University Cluj, Faculty of Mathematics and Computer Science, 2004.

[36] D Klatte and B. Kummer. *Nonsmooth equations in optimization. Regularity, calculus, methods and applications.* Nonconvex Optimization and its Applications,60. Kluwer Academic Publishers, Dordrecht, 2002.

[37] G.H. Lin and M. Fukushima. New relaxation method for mathematical programs with complementarity constraints. *Journal of Optimization Theory and Applications*, 118(1):81–116, 2003.

[38] G.H. Lin and M. Fukushima. Some exact penalty results for nonlinear programs and mathematical programs with equilibrium constraints. *Journal of Optimization, Theory and Applications*, 118(1):67–80, 2003.

[39] G. S. Liu and J. Z. Zhang. A new branch and bound algorithm for solving quadratic programs with linear complementarity constraints. *Journal of Computational and Applied Mathematics*, 146(1):77–87, 2002.

[40] X. Liu, G. Perakis, and J. Sun. A robust SQP method for mathematical programs with linear complementarity constraints. *http://www.optimization-online.org /DB_HTML/2003/09/737.html*, 2003.

[41] D.G. Luenberger. *Linear and nonlinear programming.* Addison-Wesley Inc., Reading, Massachusetts, Second edition, 1984.

[42] Z. Q. Luo, J. S. Pang, and D. Ralph. *Mathematical Programs with Equilibrium Constraints.* Cambridge University Press, Cambridge, U.K., 1996.

[43] J. Outrata, M. Kocvara, and J. Zowe. *Nonsmooth approach to optimization problems with equilibrium constraints.* Number 28 in Nonconvex Optimization and its Applications. Kluwer Academic Publishers, Dordrecht, 1998.

[44] J.S. Pang and M. Fukushima. Complementarity constraint qualications and simplied B-stationarity conditions for mathematical programs with equilibrium constraints. *Computational Optimization and Applications*, 13(1-3):111–136, 1999.

[45] M. Patriksson. *The traffic assignment problem: models and methods.* VSP, Utrecht, Netherlands, 1994.

[46] A. U. Raghunathan and L. T. Biegler. Interior point methods for mathematical programs with complementarity constraints (MPCCs). *http://www.optimization-online.org/ DB_HTML/2003/06/667html*, 2003.

[47] D. Ralph and S. J. Wright. Some properties of regularization and penalization schemes for MPEC's. *Optimization Methods and Software*, 19(5):527–556, 2004.

[48] G. Ravindran and M. Gowda Seetharama. Regularization of P_0-functions in box variational inequality problems. *SIAM Journal on Optimization*, 11(3):748– 760, 2000/01.

[49] H. Scheel and S. Scholtes. Mathematical programs with complementarity constraints: Stationarity, optimality and sensitivity. *Mathematics of Operations Research*, 25(1):1–22, 2000.

[50] R. Schmidt. Eine modifizierte standard Einbettung zur Behandlung von Gleichungs und Ungleichungs Restriktionen. Master's thesis, Humboldt Universitaet zu Berlin, 2000.

[51] S. Scholtes. Convergence properties of a regularization scheme for mathematical programs with complementarity constraints. *SIAM Journal on Optimization*, 11(4):918–936, 2001.

[52] S. Scholtes and M. Stoehr. Exact penalization of mathematical programs with equilibrium constraints. *SIAM Journal on Control Optimization*, 37(2):617–652, 1999.

[53] S. Scholtes and M. Stöhr. How stringent is the linear independence assumption for mathematical programs with stationary constraints? *Mathematics of Operations Research*, 26(4):851–863, 2001.

[54] H. Schramm and J. Zowe. A version of the bundle idea for minimizing a nonsmooth function: conceptual idea, convergence analysis, numerical results. *SIAM Journal on Optimization*, 2(1):121–152, 1992.

[55] K. Shimizu and E. Aiyoshi. A new computational method for Stackelberg and min-max problems by use of a penalty method. *Institute of Electrical and Electronics Engineers. Transactions on Automatic Control*, AC-26(2):460–466, 1981.

[56] M. V. Solodov. Merit functions and error bounds for generalized variational inequalities. *Journal of Mathematical Analysis and Applications*, 287(2):405–414, 2003.

[57] O. Stein. *On parametric semi-infinite optimization*. Shaker Verlag, Aachen, 1996.

[58] O. Stein. *Bi-level strategies in semi-infinite programming.* Nonconvex Optimization and its Applications, 71. Kluwer Academic Publishers, Boston, MA, 2003.

[59] O. Stein and G. Still. On generalized semi-infinite optimization and bilevel optimization. *European Journal of Operational Research*, 142(3):444–462, 2002.

[60] O. Stein and G. Still. Solving semi-infinite optimization problems with interior point techniques. *SIAM Journal on Control and Optimization*, 42(3):769–788, 2003.

[61] G. Still. Linear bilevel problems: Genericity results and an efficient method for computing local minima. *Mathematical Methods of Operations Research*, 55(3):383–400, 2002.

[62] G. Still and M. Streng. Survey paper: Optimality conditions in smooth nonlinear optimization. *Journal of Optimization Theory and Applications*, 90(3):483–516, 1996.

[63] J.J. Ye. Optimality conditions for optimization problems with complementarity constraints. *SIAM Journal on Optimization*, 9(2):374–387, 1999.

[64] J.J. Ye. Constraint qualifications and necessary optimality conditions for optimization problems with variational inequality constraints. *SIAM Journal on Optimization*, 10(4):943–962, 2000.

[65] J.-Z. Zhang and G.-S. Liu. A new extreme point algorithm and its application in PSQP algorithms for solving mathematical programs with linear complementarity constraints. *Journal of Global Optimization*, 19(4):345–361, 2001.

List of Figures

Index

About the author

Gemayqzel Bouza Allende was born on September 23^{rd}, 1977, in La Habana, Cuba. From 1982-1992 she attended primary and secondary school in El Cotorro, C. Habana. In 1995 she graduated from I.P.V.C.E. Vladimir Ilich Lenin, C. Habana. Then she started studying Mathematics at University of Havana. During the 1999-2000 academic year she spent a semester at Humboldt University. In 2000 she graduated with honors from the University of Havana with a thesis on parametric embeddings for Augmented Lagrangian. Since then she works at the Department of Applied Mathematics, teaching numerical analysis and nonlinear programming and doing research in Optimization. In 2001 she obtained her Master on Optimization (University of Havana).

In February 2004, she started her Ph.D. research under the supervision of Prof. Dr. Jürgen Guddat (Humboldt University) and Dr. Georg Still (University of Twente). Her research interest are Equilibrium constrained optimization problems and parametric optimization.