CRITICAL SETS IN PARAMETRIC OPTIMIZATION

H.Th. JONGEN, P. JONKER and F. TWILT

Twente University of Technology, Department of Applied Mathematics, P.O. Box 217, 7500 AE Enschede, The Netherlands

We deal with one-parameter families of optimization problems in finite dimensions. The constraints are both of equality and inequality type. The concept of a 'generalized critical point' (g.c. point) is introduced. In particular, every local minimum, Kuhn-Tucker point, and point of Fritz John type is a g.c. point. Under fairly weak (even generic) conditions we study the set Σ consisting of all g.c. points. Due to the parameter, the set Σ is pieced together from one-dimensional manifolds. The points of Σ can be divided into five (characteristic) types. The subset of 'nondegenerate critical points' (first type) is open and dense in Σ (nondegenerate means: strict complementarity, nondegeneracy of the corresponding quadratic form and linear independence of the gradients of binding constraints). A nondegenerate critical point is completely characterized by means of four indices. The change of these indices along Σ is presented. Finally, the Kuhn-Tucker subset of Σ is studied in more detail, in particular in connection with the (failure of the) Mangasarian-Fromowitz constraint qualification.

Key words: Generalized Critical Point, Critical Point, Linear Index, Quadratic Index, Kuhn-Tucker Set, Mangasarian-Fromowitz Constraint Qualification, Parametric Optimization.

1. Introduction

Let $C^k(\mathbb{R}^n, \mathbb{R})$, $k \geq 1$, denote the space of real valued, k-times continuously differentiable functions defined on the n-dimensional Euclidean space $\mathbb{R}^n$, $n \geq 1$. Given finite index sets I, J, a differentiable optimization problem $\mathcal{P}$ has the following standard formulation:

$$\mathcal{P}: \quad \text{Minimize } f \text{ on } M, \tag{1}$$

where

$$M = \{x \in \mathbb{R}^n \mid h_i(x) = 0, g_j(x) \geq 0, i \in I, j \in J\}, \tag{2}$$

and $f, h_i, g_j \in C^2(\mathbb{R}^n, \mathbb{R})$, $i \in I, j \in J$.

The function f is the objective function, $h_i(g_j)$ are the (in)equality constraints and M is the feasible set.

Our paper is a study of *one-parameter families* of optimization problems of the type (1). The motivation for this is manifold and we start with a short exposition of our incentives.

Currently, there is a growing interest in the subject of sensitivity and stability analysis of mathematical programming problems. Important contributions are contained in [3, 4]. One-parameter families play a special role in this area. In fact, within a path-connected set of perturbation parameters two problems can be joined

by means of a one-parameter family. Moreover, a detailed description of the generic
behaviour of problem $\mathscr{P}$ depending on one-parameter can be given, as will be shown
in this paper. However, such a description becomes extremely difficult if more
parameters are involved. The latter fact is related with complicated phenomena in
the theory of singularities of higher 'singularity-codimension'; see [2] and, in
particular, also [15]. Of course, there are many problems where the number of
parameters is essentially greater than one, such as, for example problems of vector-
optimization. But also in this case one-parameter families play a role as being
one-dimensional sections of the total parametric problem; see [5]. If we look at
one-parameter families as deformations of one problem into another one, we have
an intimate relation with continuation, resp. homotopy methods; see [1] for an
extensive survey. Finally, interpreting the parameter as time, we obtain an insight
into the dynamic behaviour of problem $\mathscr{P}$.

In [9] we studied the generic behaviour of the feasible set depending on one
parameter, from both a local and global point of view. Now we proceed by
investigating the structure of the set of 'critical' points (for the special case of
equality constraints only, see [11]). Our concept of a 'critical' point will be a quite
general one. But it turns out that it is very suitable when studying parametric
problems.

For $\phi \in C^1(\mathbb{R}^n, \mathbb{R})$ let $D\phi(x)$ denote the *row vector* of the first partial derivatives
at x and let $J_0(x)$ denote the index set of active (=binding) inequality constraints:

$$J_0(x) = \{j \in J \mid g_j(x) = 0\}. \tag{3}$$

Definition 1.1. A point $\bar{x} \in \mathbb{R}^n$ is called a *generalized critical point* (g.c. point) for $\mathscr{P}$
(or for $f|_M$), if $\bar{x}$ belongs to M and, moreover, if the set of vectors

$$\{Df, Dh_i, Dg_j, i \in I, j \in J_0(\bar{x})\}|_{x=\bar{x}} \tag{4}$$

is linearly dependent.

If $\bar{x}$ is a g.c. point for $\mathscr{P}$, then obviously there exist real numbers $\lambda, \lambda_i, \mu_j, i \in I$,
$j \in J_0(\bar{x})$, not all vanishing, such that

$$\lambda Df = \sum_{i \in I} \lambda_i Dh_i + \sum_{j \in J_0(\bar{x})} \mu_j Dg_j|_{x=\bar{x}}. \tag{5}$$

In case that $(\lambda, \lambda_i, \mu_j)$ in (5) can be chosen such that $\lambda > 0$, $\mu_j \geq 0$, $j \in J_0(\bar{x})$ resp.
$\lambda \geq 0$, $\mu_j \geq 0$, $j \in J_0(\bar{x})$, the point $\bar{x}$ is usually called a *Kuhn–Tucker point*, resp. a
point of Fritz John type. In particular, a local minimum $\bar{x}$ for $\mathscr{P}$ is always a point
of Fritz John type (cf. [6]). However, it need not be a Kuhn–Tucker point, unless
some constraint qualification is satisfied. The simplest constraint qualification is
linear independence of the set $\{Dh_i, Dg_j, i \in I, j \in J_0(\bar{x})\}|_{x=\bar{x}}$. In the latter case, the
number λ in (5) must be unequal to zero and, moreover, the set M is locally
C^2-diffeomorphic to $\mathbb{R}^p \times \mathbb{H}^q$, where $\mathbb{H}^q$ is the nonnegative orthant in $\mathbb{R}^q$ and
$p = n - |I| - |J_0(\bar{x})|$, $q = |J_0(\bar{x})|$ (cf. [10]). This gives rise to the following definition.

Definition 1.2. The set M in (2) is called *regular* at $\bar{x}$, $\bar{x} \in M$, if $\{Dh_i, Dg_j, i \in I,$ $j \in J_0(\bar{x})\}|_{x=\bar{x}}$ is a linearly independent set. Moreover, a generalized critical point $\bar{x}$ is called a *critical point* (for $\mathscr{P}$ or $f|_M$) if M is regular at $\bar{x}$.

In order to give an outline of the paper we need one more definition ('nondegenerate' critical point). For $\phi \in C^2(\mathbb{R}^n, \mathbb{R})$ we denote by $D^2\phi(x)$ the (symmetric) matrix of the second order partial derivatives at x (i.e. the Hessian). If A is a symmetric $n \times n$ matrix and L a linear subspace of $\mathbb{R}^n$, then by $A|_L$ we mean some matrix of the family $\mathscr{V}$, $\mathscr{V} = \{V^T A V \,|\, V$ is a matrix with n rows, whose columns form a basis for $L\}$. In view of Sylvester's theorem (cf. [13]), the number of negative, resp. zero, positive eigenvalues of $V^T A V$ does not depend on the incidental choice of V. Therefore, the number of negative, resp. zero, positive eigenvalues of $A|_L$ is defined to be the corresponding number of $V^T A V$, where $V^T A V \in \mathscr{V}$. Furthermore, $A|_L$ is said to be nonsingular if $V^T A V$ is nonsingular, where $V^T A V \in \mathscr{V}$.

For an $r \times q$ matrix B, the set Ker B will be

$$\text{Ker } B = \{\xi \in \mathbb{R}^q \,|\, B\xi = 0\}. \tag{6}$$

Definition 1.3. Let M be *regular* at $\bar{x}$ and let $\bar{x}$ be a critical point for $f|_M$, i.e. there exist (Lagrange parameters) $\bar{\lambda}_i$, $\bar{\mu}_j$, $i \in I$, $j \in J_0(\bar{x})$, such that

$$Df = \sum_{i \in I} \bar{\lambda}_i Dh_i + \sum_{j \in J_0(\bar{x})} \bar{\mu}_j Dg_j|_{x=\bar{x}}. \tag{7}$$

The critical point $\bar{x}$ is called *nondegenerate* if the following two conditions hold:

$ND1$: $\bar{\mu}_j \neq 0, j \in J_0(\bar{x})$,

$ND2$: $D^2 L(\bar{x})|_T$ is nonsingular,

where (L = Lagrange function, T = tangent space)

$$L(x) = f(x) - \sum_{i \in I} \bar{\lambda}_i h_i(x) - \sum_{j \in J_0(\bar{x})} \bar{\mu}_j g_j(x), \tag{8}$$

$$T = \bigcap_{i \in I} \text{Ker } Dh_i(\bar{x}) \cap \bigcap_{j \in J_0(\bar{x})} \text{Ker } Dg_j(\bar{x}). \tag{9}$$

The linear index LI, resp. linear coindex LCI, is defined to be the number of $\bar{\mu}_j$ in (7) which are negative, resp. positive. The quadratic index QI, resp. quadratic coindex QCI, is defined to be the number of negative, resp. positive eigenvalues of $D^2 L(\bar{x})|_T$, with L, T as in (8), (9).

The numbers LI, LCI, QI, QCI at a nondegenerate critical point $\bar{x}$ completely characterize the local behaviour of $f|_M$ and in this sense they are intrinsic (cf. [10] for a detailed exposition). In particular, if LI = QI = 0, resp. LCI = QCI = 0, then $\bar{x}$ is a local minimum, resp. local maximum for $f|_M$. In all other cases, $\bar{x}$ is a certain kind of saddle-point. If LI = 0, then $\bar{x}$ is a Kuhn–Tucker point and in that case the quadratic index is the appropriate generalization of the so called Morse-index (cf. [10, 14]).

In the sequel we consider the following setting of a one-parameter family of problems, where a vector $z \in \mathbb{R}^{n+1}$ is always partitioned as $z = (x, t)$, $x \in \mathbb{R}^n$, $t \in \mathbb{R}$ (t = parameter):

$$\mathcal{P}(t): \quad \text{Minimize } f(\cdot, t) \text{ on } M(t) \quad (t \in \mathbb{R}) \tag{10}$$

where

$$M(t) = \{x \in \mathbb{R}^n \mid h_i(x, t) = 0, g_j(x, t) \geqslant 0, i \in I, j \in J\}, \tag{11}$$

$$I = \{1, \ldots, m\}, \quad m < n, \quad J = \{1, \ldots, s\}, \tag{12}$$

$$f, h_i, g_j \in C^3(\mathbb{R}^{n+1}, \mathbb{R}), \quad i \in I, \ j \in J.$$

A point $\bar{z} \in \mathbb{R}^{n+1}$ is called a (generalized, nondegenerate) critical point for $\mathcal{P}(\cdot)$ if $\bar{x}$ is a (generalized, nondegenerate) critical point for $\mathcal{P}(\bar{t})$, where $\bar{z} = (\bar{x}, \bar{t})$. The set of all generalized critical points will be denoted by Σ, so

$$\Sigma = \{z \in \mathbb{R}^{n+1} \mid x \text{ is g.c. point for } \mathcal{P}(t), \text{ where } z = (x, t)\}. \tag{13}$$

Roughly speaking, for most problems $\mathcal{P}(\cdot)$ it turns out that most points of Σ are nondegenerate critical points (points of *Type* 1 in the subsequent terminology). This observation is in a certain sense connected with an interesting study in [16]. However, one cannot expect that all points of Σ are nondegenerate critical points. In fact, for some parameter values t it might happen that at some points x the regularity of the set $M(t)$ is not satisfied (if the number of active constraints is less than $(n+1)$ we are dealing with a point of *Type* 4; if this number equals $(n+1)$ we will have a point of *Type* 5). So, this is one 'degeneracy' phenomenon. Other degeneracies can take place at critical points (so $M(t)$ is regular) where one of the conditions ND1, ND2 does not hold. Note, if ND1 fails to hold, then the 'strict complementarity' is not satisfied (*Type* 2). Further, if ND2 is not satisfied, then one of the eigenvalues of $D^2L(\bar{x})|_T$ vanishes (*Type* 3). We emphasize that the degeneracies (Type 2, 3, 4, 5) are stable, i.e. they remain present under 'C^3-perturbations' of the problem-defining functions f, h_i, g_j. The set Σ will generically be one-dimensional. Walking along Σ we will meet at certain (discrete) points a degeneracy (Type 2, 3, 4, 5). A basic question then arises: if we pass a point of Type 2, 3, 4, 5, what effect does this have on the local structure of the nondegenerate critical points before and after passing; i.e. how does the set of indices LI, LCI, QI, QCI change?

An important related study on Kuhn-Tucker points is done in [12], where it is assumed that the Mangasarian-Fromowitz constraint qualification holds (cf. Section 4). Finally we remark that 'Kuhn-Tucker' branches of the set Σ might be connected by means of non-Kuhn-Tucker branches (see [8]). In a certain sense this is analogous to the complexification-effect of a one-parameter family of real analytic systems (see [1, Section 4]).

The paper is organized as follows. In Section 2 we state a genericity theorem which clarifies the concept 'most problems $\mathcal{P}(\cdot)$'. In Section 3 we describe in five subsections the local behaviour of the set Σ in a neighbourhood of each of the

points of Type 1-5 (including the change of the indices LI, LCI, QI, QCI). Finally, in Section 4 we concentrate on the (closure of the) Kuhn-Tucker subset and discuss those situations where the Mangasarian-Fromowitz constraint qualification is not satisfied.

2. The genericity theorem

In this section we adopt the notation of Section 1. The space $C^3(\mathbb{R}^{n+1}, \mathbb{R})$ will be endowed with the strong (or Whitney-) C_S^3-topology (cf. [7]), the C_S^3-topology of the product of a finite number of copies of $C^3(\mathbb{R}^{n+1}, \mathbb{R})$ being the induced product topology. A typical C_S^3 base-neighbourhood $\mathcal{N}_\varepsilon$ of the zero function in $C^3(\mathbb{R}^{n+1}, \mathbb{R})$ is induced by means of a continuous positive function $\varepsilon : \mathbb{R}^{n+1} \to \mathbb{R}$ as follows:

$$\mathcal{N}_\varepsilon = \left\{ \phi \in C^3(\mathbb{R}^{n+1}, \mathbb{R}) \, \middle| \, |\phi(z)| + \Sigma \left| \frac{\partial}{\partial z_i} \phi(z) \right| + \Sigma \left| \frac{\partial^2}{\partial z_i \partial z_j} \phi(z) \right| \right.$$
$$\left. + \Sigma \left| \frac{\partial^3}{\partial z_i \partial z_j \partial z_k} \phi(z) \right| < \varepsilon(z) \text{ for all } z \in \mathbb{R}^{n+1} \right\}.$$

A typical C_S^3 base-neighbourhood of $f \in C^3(\mathbb{R}^{n+1}, \mathbb{R})$ will be the set $f + \mathcal{N}_\varepsilon$.

In Section 1 we intuitively introduced points of Type 1-5. These types will be made precise in the next section. Taking this for granted at this stage we can proceed with the statement of the genericity theorem.

Definition 2.1. Let the subset $\mathcal{F}$ of $C^3(\mathbb{R}^{n+1}, \mathbb{R})^{1+m+s}$ be defined as follows: $(f, h_1, \ldots, h_m, g_1, \ldots, g_s) \in \mathcal{F}$ if each point of Σ (cf. (13)) is one of the Types 1, 2, 3, 4, 5.

Theorem 2.1 (*Genericity*). *The set $\mathcal{F}$ is C_S^3 open and dense in $C^3(\mathbb{R}^{n+1}, \mathbb{R})^{1+m+s}$.*

The dense-part of the proof of Theorem 2.1 is tedious and based on transversal approximation. A sketch of the proof in case $J = \emptyset$ (i.e. without inequality constraints) is given in [11] and that proof can be extended without severe difficulties to a proof of the present situation. The open-part of the proof of Theorem 2.1 can be accomplished by means of continuity arguments (note that C^3-differentiability is needed in view of points of Type 3). We define:

$$\Sigma^i = \{ z \in \Sigma \, | \, z \text{ is of Type } i \}, \quad i = 1, \ldots, 5. \tag{14}$$

Theorem 2.2. *Let $(f, h_1, \ldots, h_m, g_1, \ldots, g_s)$ belong to $\mathcal{F}$. Then we have: Σ^1 is open and dense in Σ and, for $i = 2, 3, 4, 5$, the set Σ^i is a discrete point set.*

The proof of Theorem 2.2 is a straightforward consequence of the local structure of the set Σ in a neighbourhood of the points of Type 1-5, as will be explained in the next section.

3. Analysis of the five types

In this section we will make an analysis on the points of Type 1-5 which we intuitively introduced in Section 1. To this aim we divide this section into five subsections, each of them being dedicated to one single type. Every subsection is organized as follows. Firstly, we state a number of conditions which are necessary to describe the type under consideration, and we introduce so-called *characteristic numbers* which determine the essence of the type. Based on this information we proceed with a local description of the set Σ and the index-relations involved.

The notation of the foregoing sections will be adopted. In particular, a vector $z \in \mathbb{R}^{n+1}$ will always be partitioned as $z = (x, t)$, $x \in \mathbb{R}^n$, $t \in \mathbb{R}$, and $J_0(z)$ denotes the index set of active (=binding) inequality constraints. For $\phi \in C^k(\mathbb{R}^{n+1}, \mathbb{R})$, $D_x \phi$ stands for the row vector of first partial derivatives with respect to x; $D_t \phi$, $D_x^2 \phi$ are to be interpreted analogously.

3.1. Points of type 1

A generalized critical point $\bar{z} = (\bar{x}, \bar{t})$ is of Type 1 if $\bar{z}$ is a nondegenerate critical point, i.e. $\bar{x}$ is a nondegenerate critical point for $\mathcal{P}(\bar{t})$.

Characteristic numbers: LI, LCI, QI, QCI.

Let $\bar{z} = (\bar{x}, \bar{t})$ be a point of Type 1. Then, in an $(\mathbb{R}^{n+1} -)$ neighbourhood of $\bar{z}$ we can parametrize the set Σ by means of the parameter t, whereas the indices LI, LCI, QI, QCI remain (locally) constant. To see this, we use the implicit function theorem. Without loss of generality we may assume, in case $J_0(\bar{z}) \neq \emptyset$, that $J_0(\bar{z}) = \{1, \ldots, p\}$. Let λ, resp. μ, be an m, resp. p vector, with components λ_i, resp. μ_j. Consider the map $\mathcal{T}: \mathbb{R}^{n+m+p+1} \to \mathbb{R}^{n+m+p}$,

$$
\mathcal{T}: \begin{pmatrix} x \\ \lambda \\ \mu \\ t \end{pmatrix} \mapsto \begin{pmatrix} D_x^T f(x, t) - \sum_{i=1}^{m} \lambda_i D_x^T h_i(x, t) - \sum_{j=1}^{p} \mu_j D_x^T g_j(x, t) \\ h_i(x, t), \quad i = 1, \ldots, m \\ g_j(x, t), \quad j = 1, \ldots, p \end{pmatrix} \tag{15}
$$

The map $\mathcal{T}$ in (15) is of class C^2 since f, h_i, g_j are of class C^3. Let $(\bar{\lambda}_i)$, $(\bar{\mu}_j)$ be the Lagrange parameters at $\bar{x}$ as a critical point for $\mathcal{P}(\bar{t})$ (cf. (7)). Then $\mathcal{T}$ vanishes at the point $(\bar{x}, \bar{\lambda}, \bar{\mu}, \bar{t})$. The partial derivative of $\mathcal{T}$ at $(\bar{x}, \bar{\lambda}, \bar{\mu}, \bar{t})$ with respect to (x, λ, μ) has the following typical blockstructure:

$$
\left(\begin{array}{c|c} D_x^2 L(\bar{z}) & -B \\ \hline B^T & 0 \end{array} \right), \tag{16}
$$

where

$$
L(z) = f(z) - \sum_{i=1}^{m} \bar{\lambda}_i h_i(z) - \sum_{j=1}^{p} \bar{\mu}_j g_j(z) \tag{17}
$$

and $B = (D_x^T h_1 | \cdots | D_x^T g_p)|_{z=\bar{z}}$.

Note that the matrix B in (16) has rank $(m + p)$. Then, from condition ND2 (cf. Definition 1.3) it follows that the matrix in (16) is nonsingular (see [10]). Now we can apply the implicit function theorem and obtaîn C^2-mappings $x(t)$, $\lambda(t)$, $\mu(t)$ (in an open neighbourhood of $\bar{t}$) such that $\mathcal{T}(x(t), \lambda(t), \mu(t), t) \equiv 0$. Note that all components of $\bar{\mu}$ are unequal to zero (condition ND1) and that $\mu(\bar{t}) = \bar{\mu}$. Consequently, for t sufficiently close to $\bar{t}$ we see that no component of $\mu(t)$ vanishes. In an open neighbourhood of $\bar{z}$ we can parametrize the set Σ by means of a unique C^2-map $t \mapsto (x(t), t)$. Hence, Σ^1 is a one-dimensional C^2-manifold and the indices LI, LCI, QI, QCI are constant on every connected component of Σ^1.

3.2. Points of Type 2

A generalized critical point $\bar{z} = (\bar{x}, \bar{t})$ is of Type 2 if the following conditions A1–A6 hold:

A1. $\bar{x}$ is a critical point for $\mathcal{P}(\bar{t})$.

A2. $J_0(\bar{z}) \neq \emptyset$.

After renumbering we may assume that $J_0(\bar{z}) = \{1, \ldots, p\}$, $p \geq 1$. Then, we have (cf. A1 and (7))

$$D_x f = \sum_{i=1}^m \bar{\lambda}_i D_x h_i + \sum_{j=1}^p \bar{\mu}_j D_x g_j \big|_{z = \bar{z}}. \tag{18}$$

A3. In (18) exactly one of the Lagrange parameters $\bar{\mu}_j$ vanishes.

After renumbering we may assume that $\bar{\mu}_p = 0$ and $\bar{\mu}_j \neq 0$, $j = 1, \ldots, p - 1$. Let the Lagrange function L be defined as in (17) and put

$$T = \bigcap_{i \in I} \text{Ker } D_x h_i(\bar{z}) \cap \bigcap_{j \in J_0(\bar{z})} \text{Ker } D_x g_j(\bar{z}), \tag{19}$$

$$\tilde{T} = \bigcap_{i \in I} \text{Ker } D_x h_i(\bar{z}) \cap \bigcap_{j \in J_0(\bar{z}) \setminus \{p\}} \text{Ker } D_x g_j(\bar{z}). \tag{20}$$

A4. $D_x^2 L(\bar{z})|_T$ is nonsingular.

A5. $D_x^2 L(\bar{z})|_{\tilde{T}}$ is nonsingular.

Let B be an $n \times r$-matrix of rank r. By $B^\dagger$ we denote the matrix $(B^T B)^{-1} B^T$. In fact, $B^\dagger$ is the Moore–Penrose inverse of B.

Let W be a matrix with n rows, whose columns form a basis for the linear space $\tilde{T}$. Put $\Phi = (h_1, \ldots, h_m, g_1, \ldots, g_{p-1})^T$ and define the $n \times 1$-vectors:

$$\alpha = -((D_x^T \Phi)^\dagger)^T \cdot D_t \Phi, \tag{21}$$

$$\beta = -W(W^T \cdot D_x^2 L \cdot W)^{-1} W^T \{D_x^2 L \cdot \alpha + D_t D_x^T L\}. \tag{22}$$

In (21), (22), D_t stands for $\partial/\partial t$ and all partial derivatives are evaluated at $\bar{z}$. Next, we put

$$\gamma = D_x g_p(\bar{z})(\alpha + \beta) + D_t g_p(\bar{z}) \tag{23}$$

Note that β (and thus γ as well) is independent of the choice of the matrix W.

A6. $\gamma \neq 0$.

Let δ_1, resp. δ_2 denote the number of negative eigenvalues of $D_x^2 L(\bar{z})|_{\tilde{T}}$, resp. $D_x^2 L(\bar{z})|_T$, and put $\delta = \delta_1 - \delta_2$.

Characteristic numbers: sign(γ), δ.

In the special case that $I = \emptyset$ and $p = 1$, we have $\tilde{T} = \mathbb{R}^n$ (cf. A5) and we just delete all entries of Φ in (21).

We proceed with an analysis of the set Σ in a neighbourhood of $\bar{z}$.

The simplest example to have in mind is the following (one-dimensional): $f(x, t)$ is one of the four functions $\pm(x \pm t)^2$, there is only one constraint, namely the constant inequality constraint $x \geq 0$, and the point $\bar{z}$ is the origin in $\mathbb{R}^2$. Note that, as t passes zero, an unconstrained critical point is moving into the feasible set or out of the feasible set. (A similar higher dimensional example is obtained by choosing $f(x, t)$ from the functions $\pm(x_n \pm t)^2 + \sum_{j=1}^{n-1} \pm x_j^2$ and taking only one constraint, namely $x_n \geq 0$ into account).

It will turn out that, locally, Σ is pieced together from two curves which intersect at $\bar{z}$. In fact, since $\bar{\mu}_p$ vanishes, it follows that $\bar{x}$ is a critical point both for problem $\mathcal{P}(\bar{t})$ as well as for the problem $\tilde{\mathcal{P}}(\bar{t})$, where $\tilde{\mathcal{P}}(\bar{t})$ differs from $\mathcal{P}(\bar{t})$ only in the fact that the inequality constraint g_p is *deleted*. The Lagrange parameters $\bar{\mu}_1, \ldots, \bar{\mu}_{p-1}$ are unequal to zero.

Together with condition A5 it follows that $\bar{x}$ is a nondegenerate critical point for $\tilde{\mathcal{P}}(\bar{t})$. Now we can apply the result of Section 3.1 to problem $\tilde{\mathcal{P}}(\cdot)$. Let $\tilde{\Sigma}$ denote the set of g.c. points for problem $\tilde{\mathcal{P}}(\cdot)$. Then, in a neighbourhood of $\bar{z}$ the set $\tilde{\Sigma}$ is a one-dimensional manifold, parametrized by means of a unique C^2-map $t \mapsto (\tilde{x}(t), t)$. The latter curve belongs to Σ as far as $\psi(t)$ is nonnegative, where $\psi(t) := g_p(\tilde{x}(t), t)$, and it traverses the zero set '$g_p = 0$' transversally iff $(d\psi/dt)(\bar{t}) \neq 0$. A few calculations show that $(d\tilde{x}/dt)(\bar{t}) = \alpha + \beta$ (cf. (21), (22)). Hence $(d\psi/dt)(\bar{t}) = \gamma$, with γ as in (23). So, if we walk along $\tilde{\Sigma}$ as t increases, then at $t = \bar{t}$ we leave (enter) the feasible set $M(t)$ (cf. (11)) according to sign(γ) = -1 ($+1$).

Next, we consider a problem $\hat{\mathcal{P}}(\bar{t})$ which differs from $\mathcal{P}(\bar{t})$ only in the fact that g_p is treated as an *equality* constraint. Of course, $\bar{x}$ is also a critical point for $\hat{\mathcal{P}}(\bar{t})$ and, moreover, $\bar{x}$ is a nondegenerate critical point for $\hat{\mathcal{P}}(\bar{t})$. In fact, the nondegeneracy condition ND1 (cf. Definition 1.3) is satisfied, since the vanishing Lagrange parameter $\bar{\mu}_p$ does not anymore correspond to an inequality constraint (by the very definition of $\hat{\mathcal{P}}(\bar{t})$). Condition ND2 holds in view of A4. Let $\hat{\Sigma}$ be the set of g.c. points for problem $\hat{\mathcal{P}}(\cdot)$. Again we can apply the result of Section 3.1 and hence, in a neighbourhood of $\bar{z}$ the set $\hat{\Sigma}$ is a one-dimensional manifold, parametrized by means of a unique C^2-map $t \mapsto (\hat{x}(t), t)$. (Note that $g_p(\hat{x}(t), t) \equiv 0$). Since the curve $\tilde{\Sigma}$ traverses the zero set '$g_p = 0$' at $\bar{z}$ transversally, it follows that $\tilde{\Sigma}$ and $\hat{\Sigma}$ intersect at $\bar{z}$ under a nonvanishing angle. Obviously, in a neighbourhood of $\bar{z}$, the set Σ consists of $\hat{\Sigma}$ and that part of $\tilde{\Sigma}$ on which g_p is nonnegative. Altogether, the set Σ has a local structure as depicted in Fig. 1c,d, and note that $\Sigma \setminus \{\bar{z}\}$ consists of nondegenerate critical points.

We proceed with the calculation of the index-relations between the branches of nondegenerate critical points in $\Sigma \setminus \{\bar{z}\}$. Let us consider the (above defined) curve
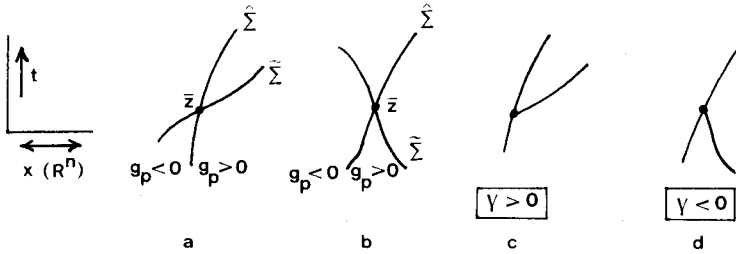
Fig. 1.

$t \mapsto (\hat{x}(t), t)$ once more, i.e. we treat g_p as an equality constraint. For each t we can calculate the Lagrange parameter corresponding to g_p, say $\mu_p(t)$, at the point $\hat{x}(t)$, viewed at as a critical point for $\hat{\mathscr{P}}(t)$. Since all data are of class C^3, $\mu(\cdot)$ will be of class C^2. Now, if we look at g_p as an inequality constraint, then, except for $t = \bar{t}$, the point $\hat{x}(t)$ is a nondegenerate critical point for $\mathscr{P}(t)$. In fact, the only difference between two critical points $\hat{x}(t_1)$, $\hat{x}(t_2)$ for $\mathscr{P}(t)$, where $t_1 < \bar{t} < t_2$, lies in the fact that one Lagrange parameter $(\mu_p(t))$ belonging to an inequality constraint (g_p) changes sign. In particular, we will show the more general relation:

$$\text{sign}(\gamma) \cdot \text{sign}\left(\frac{d\mu_p}{dt}(\bar{t})\right) = +1 \ (\text{resp. } -1) \quad \text{iff} \quad \delta = 1 \ (\text{resp. } \delta = 0), \quad (24)$$

where $\text{sign}(\gamma)$ and δ are the characteristic numbers.

Taking (24) for granted at this stage and noting that the dimensions of $\tilde{T}$ and T (cf. (19), (20)) differ by one, the index-relations are easily derived. In Fig. 2 we made a picture of all four possibilities; the 4-vectors stand for (LI, LCI, QI, QCI). Note that $LI + LCI + QI + QCI = n - m$, whereas $LI + LCI$ equals p or $p - 1$.
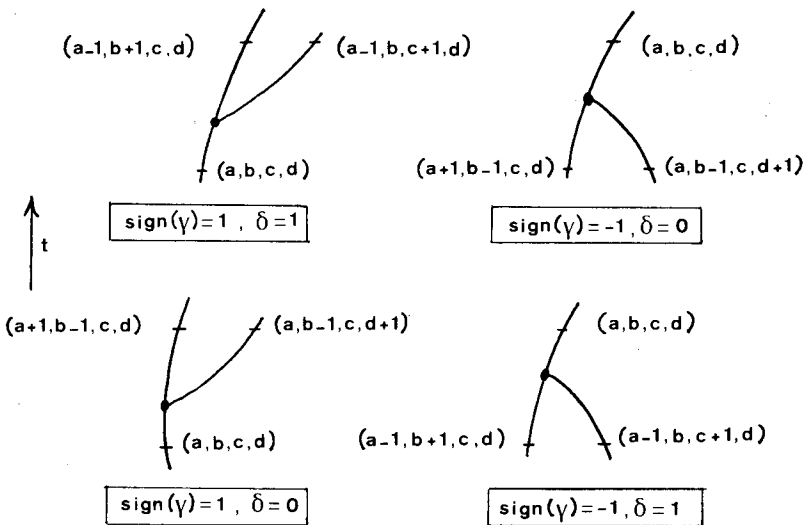

Fig. 2.

In order to derive (24) we start with the following observation: the inequality constraints $g_1, \ldots, g_{p-1}$ remain active (locally) along $\tilde{\Sigma}$ and $\hat{\Sigma}$. Consequently, the only inequality constraint function that really counts is the function g_p. So, we can simplify the verification of (24) in local coordinates in a neighborhood of the origin in $\mathbb{R}^{q+1}$, where $q = \dim \tilde{T} = n - m - p + 1$. In these new coordinates, say $(y, u) = \Psi(x, t)$, 'u-hyperplanes' correspond to 't-hyperplanes' and the orientation of t is preserved.

In fact, choose vectors $\xi_i \in \mathbb{R}^n$, $i = 1, \ldots, q-1$, which form together with the set $\{D_x^T h_i(\bar{z}), D_x^T g_j(\bar{z}), i = 1, \ldots, m, j = 1, \ldots, p\}$ a basis for $\mathbb{R}^n$, and consider the local C^3-coordinate transformation in $\mathbb{R}^{n+1}$, $(y_1, \ldots, y_n, u) = \Psi(x_1, \ldots, x_n, t)$, defined by

$$y_i = \xi_i^T(x - \bar{x}), \quad i = 1, \ldots, q-1,$$

$$y_q = g_p(x, t),$$

$$y_{q+j} = g_j(x, t), \quad j = 1, \ldots, p-1,$$

$$y_{q+p+i-1} = h_i(x, t), \quad i = 1, \ldots, m,$$

$$u = t - \bar{t}.$$

Note that $(\bar{x}, \bar{t})$ is sent to the origin, and that, for t close to $\bar{t}$ and x close to $\bar{x}$, the intersection of the feasible set $M(t)$ with the zero set (t fixed): $g_1 = \cdots = g_{p-1} = h_1 = \cdots = h_m = 0$, is transformed to the following set of dimension q (*constant* with respect to u):

$$\{(y_1, \ldots, y_q) \in \mathbb{R}^q \mid y_q \geqslant 0\}.$$

Let us denote the transformed function $f(x, t)$ by $g(y_1, \ldots, y_q, u)$. Instead of the critical points $\tilde{x}(t)$, resp. $\hat{x}(t)$ we obtain vector functions $(\tilde{y}_1(u), \ldots, \tilde{y}_q(u))$, resp. $(\hat{y}_1(u), \ldots, \hat{y}_{q-1}(u), 0)$ satisfying:

$$\frac{\partial g}{\partial y_i}(\tilde{y}_1(u), \ldots, \tilde{y}_q(u), u) \equiv 0, \quad i = 1, \ldots, q,$$

$$\frac{\partial g}{\partial y_i}(\hat{y}_1(u), \ldots, \hat{y}_{q-1}(u), 0, u) \equiv 0, \quad i = 1, \ldots, q-1. \tag{25}$$

moreover, $\mu_p(t)$ (recall (24)) becomes $\mu(u)$, where

$$\mu(u) := \frac{\partial g}{\partial y_q}(\hat{y}_1(u), \ldots, \hat{y}_{q-1}(u), 0, u), \tag{26}$$

and the corresponding γ, say γ', becomes $(d\tilde{y}_q/du)(0)$.

Put $\tilde{G} = D_y^2 g(0)$ and let $\hat{G}$ be the $(q-1) \times (q-1)$ matrix obtained from $\tilde{G}$ by deleting the last column and last row. In view of (A4), (A5) we see that $\tilde{G}$ and $\hat{G}$ are both nonsingular. Furthermore, we put $g_{u,i} = (\partial^2 g / \partial u \partial y_i)(0)$, $i = 1, \ldots, q$, and $g_{i,j} = (\partial^2 g / \partial y_i \partial y_j)(0)$.

Differentiation of (25), (26) yields:

$$\frac{d\mu}{du}(0) = [-(g_{1,q}, \ldots, g_{q-1,q}) \cdot \hat{G}^{-1}, 1] \cdot (g_{u,1}, \ldots, g_{u,q})^{\mathrm{T}},$$

$$\gamma' = -e_q^{\mathrm{T}} \cdot \tilde{G}^{-1} \cdot (g_{u,1}, \ldots, g_{u,q})^{\mathrm{T}}, \quad \text{where } e_q = (0, \ldots, 0, 1)^{\mathrm{T}}. \tag{27}$$

A short calculation shows:

$$e_q^{\mathrm{T}} \tilde{G}^{-1} = [e_q^{\mathrm{T}} \tilde{G}^{-1} e_q][-(g_{1,q}, \ldots, g_{q-1,q}) \cdot \hat{G}^{-1}, 1]. \tag{28}$$

Now, substitution of (28) in (27) yields the relation:

$$\gamma' = -\frac{d\mu}{du}(0) \cdot [e_q^{\mathrm{T}} \tilde{G}^{-1} e_q],$$

and hence

$$\text{sign}(\gamma') \cdot \text{sign}\left(\frac{d\mu}{du}(0)\right) = -\text{sign}[e_q^{\mathrm{T}} \tilde{G}^{-1} e_q]. \tag{29}$$

Let the symbol $\#$ denote 'the number of negative eigenvalues'.
   Note that we have:

$$\text{sign}[e_q^{\mathrm{T}} \tilde{G}^{-1} e_q] = +1(-1) \quad \text{iff} \quad \# \tilde{G}^{-1}|_{\text{span}\{e_q\}} = 0(1). \tag{30}$$

The characteristic number δ becomes

$$\delta = \# \tilde{G} - \# \hat{G}. \tag{31}$$

We have to show (cf. (24))

$$\text{sign}(\gamma') \cdot \text{sign}\left(\frac{d\mu}{du}(0)\right) = +1 \; (-1) \quad \text{iff} \quad \delta = 1 \; (0). \tag{32}$$

In view of (29), (30), (31), it suffices to show the following equality:

$$\# \tilde{G} - \# \hat{G} = \# \tilde{G}^{-1}|_{\text{span}\{e_q\}}. \tag{33}$$

The equality (33) is a direct consequence of the following lemma (recall the definition of $\tilde{G}$ and $\hat{G}$). Its proof is given in [11], in case dim $L = n - 1$. However, the idea of that proof fits the general case as well, and therefore we delete the proof here.

Lemma 3.2.1. *Let A be a nonsingular symmetric $n \times n$ matrix, L a linear subspace of $\mathbb{R}^n$ and $L^\perp$ its orthogonal complement. Suppose that $A^{-1}|_{L^\perp}$ is nonsingular. Then we have*
   a. $A|_L$ *is nonsingular,*
   b. $\#(A) = \#(A|_L) + \#(A^{-1}|_{L^\perp})$.

3.3. Points of Type 3

A generalized critical point $\bar{z} = (\bar{x}, \bar{t})$ is of Type 3 if the following conditions B1–B4 hold:

B1. $\bar{x}$ is a critical point for $\mathcal{P}(\bar{t})$.

After renumbering we may assume in case $J_0(\bar{z}) \neq \emptyset$ that $J_0(\bar{z}) = \{1, \dots, p\}$. From B1 and (7) we see that the critical point relation (18) holds.

B2. In (18) we have $\bar{\mu}_j \neq 0$, $j = 1, \dots, p$.

Let the Lagrange function L be defined as in (17) and let the tangent space T be as in (19).

B3. Exactly one eigenvalue of $D_x^2 L(\bar{z})|_T$ vanishes.

Let V be a matrix with n rows, whose columns form a basis for the tangent space T. According to B3, let w be a nonvanishing vector such that $V^T D_x^2 L(\bar{z}) V w = 0$, and put $v = Vw$. Put $\Phi = (h_1, \dots, h_m, g_1, \dots, g_p)^T$ and define (the symbol † denoting as in (21) the Moore–Penrose inverse):

$$\beta_1 = v^T (D_x^3 L \cdot v) v - 3 v^T D_x^2 L \cdot ((D_x^T \Phi)^\dagger)^T (v^T D_x^2 \Phi v), \tag{34}$$

$$\beta_2 = D_t (D_x L \cdot v) - D_t^T \Phi \cdot (D_x^T \Phi)^\dagger \cdot D_x^2 L v, \tag{35}$$

where

$$v^T (D_x^3 L \cdot v) v = \sum_{i,j,k=1}^{n} (\partial^3 / \partial x_i \partial x_j \partial x_k) L \cdot v_i v_j v_k,$$

$$v^T D_x^2 \Phi v = (v^T D_x^2 h_1 v, \dots, v^T D_x^2 g_p v)^T,$$

all partial derivatives being evaluated at $\bar{z}$. In case that $I = J_0(\bar{z}) = \emptyset$ we have $T = \mathbb{R}^n$ and we delete the entries of Φ in (34), (35).

Next we define:

$$\beta = \beta_1 \cdot \beta_2. \tag{36}$$

B4. $\beta \neq 0$.

We note that v is determined up to a scalar and hence, β is determined up to a *positive* scalar.

Let α denote the number of negative eigenvalues of $D_x^2 L(\bar{z})|_T$.

Characteristic numbers: α, $\text{sign}(\beta)$.

Before we proceed with a local analysis of the set Σ, let us firstly present the easiest example. In fact take the one-dimensional, unconstrained case: $f(x, t) = \frac{1}{3} x^3 + tx$. The point $\bar{z}$ is now the origin and for the set Σ we obtain: $\Sigma = \{(x, t) \mid x^2 + t = 0\}$. Obviously, Σ is a parabola. Note that $(\partial^2 f / \partial x^2)(0) = 0$. This corresponds with condition B3. However, both $(\partial^3 f / \partial x^3)$ and $(\partial^2 f / \partial t \partial x)$ do not vanish at the origin (cf. β_1, β_2 in (34), (35)). In fact, condition B4 is a generalization of this.

Now we consider the general case. From condition B1 and Definition 1.2 we see that the set $\{D_x h_i(\bar{z}),\ D_x g_j(\bar{z}),\ i \in I,\ j \in J_0(\bar{z})\}$ is linearly independent, and from B2 we learn that all Lagrange parameters $\bar{\mu}_j,\ j \in J_0(\bar{z})$ are unequal to zero. From this we conclude that—in some neighbourhood $\mathcal{O}$ of $\bar{z}$—the set Σ consists entirely of critical points and, moreover, that $J_0(z) = J_0(\bar{z})$ for all $z \in \Sigma \cap \mathcal{O}$. The latter fact follows from the observation that the set $\{D_x f,\ D_x h_i,\ i \in I,\ D_x g_j, j \in \tilde{J}\}|_{\bar{z}}$ is linearly independent if $\tilde{J} \subset J_0(\bar{z})$ and $\tilde{J} \neq J_0(\bar{z})$. But then, we may—locally—consider the inequality constraint functions $g_j,\ j \in J_0(\bar{z})$ as equality constraint functions in order to describe the structure of the set Σ.

For equality-constrained problems, the present type has been studied extensively in [11, 'Type 2']. Therefore, we may restrict ourselves to a citation of those results within this context. In fact, in a neighbourhood of $\bar{z}$, the set Σ is a one-dimensional C^2-manifold. Moreover, the parameter t, viewed at as a function on Σ, has a (nondegenerate) local maximum, resp. local minimum, at $(\bar{x}, \bar{t})$ according to $\text{sign}(\beta) = +1$, resp. $\text{sign}(\beta) = -1$. Consequently, the set Σ can be approximated by means of a parabola, in a neighbourhood of $\bar{z}$. The orientation of this parabola as well as the index-relations are depicted in Fig. 3. In view of condition B2, the indices LI, LCI do not change when passing the point $\bar{z}$ along Σ. Moreover, apart from the degenerate critical point $\bar{z}$, we have $QI + QCI = n - |I| - |J_0(\bar{z})|$. Therefore, in Fig. 3 we only mention the change of QI.
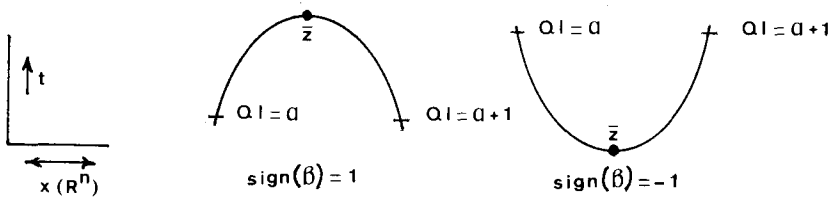

Fig. 3.

3.4. Points of Type 4

A generalized critical point $\bar{z} = (\bar{x}, \bar{t})$ is of Type 4 if the following conditions C1–C6 hold:

C1. $|I| + |J_0(\bar{z})| > 0$, $|\cdot|$ *denoting the cardinality.*
After renumbering we may assume in case $J_0(\bar{z}) \neq \emptyset$, that $J_0(\bar{z}) = \{1, \ldots, p\}$.

C2. $\dim.\ \text{span}\{D_x h_i(\bar{z}), D_x g_j(\bar{z}), i \in I, j \in J_0(\bar{z})\} = m + p - 1$.

C3. $m + p - 1 < n$.

From C2 we see that there exist $\lambda_i,\ \mu_j,\ i \in I,\ j \in J_0(\bar{z})$, not all vanishing such that

$$\sum_{i=1}^{m} \lambda_i D_x h_i(\bar{z}) + \sum_{j=1}^{p} \mu_j D_x g_j(\bar{z}) = 0. \tag{37}$$

Note that the numbers $\lambda_i,\ \mu_j$ in (37) are unique up to a common multiple.

C4. In case $p \neq 0$, we have $\mu_j \neq 0,\ j = 1, \ldots, p$, and we normalize the μ_j's by setting $\mu_p = 1$ (normalization).

We define furthermore

$$L(z) = \sum_{i=1}^{m} \lambda_i h_i(z) + \sum_{j=1}^{p} \mu_j g_j(z), \tag{38}$$

where λ_i, μ_j in (38) satisfy (37),

$$T = \bigcap_{i \in I} \text{Ker } D_x h_i(\bar{z}) \cap \bigcap_{j \in J_0(\bar{z})} \text{Ker } D_x g_j(\bar{z}). \tag{39}$$

Let W be a matrix with n rows, whose columns form a basis for T. Define

$$A = D_t L \cdot W^{\mathrm{T}} \cdot D_x^2 L \cdot W, \tag{40}$$

$$w = W^{\mathrm{T}} \cdot D_x^{\mathrm{T}} f, \tag{41}$$

all partial derivatives being evaluated at $\bar{z}$.

C5. A is nonsingular.

Finally, define

$$\alpha = w^{\mathrm{T}} A^{-1} w. \tag{42}$$

C6. $\alpha \neq 0$.

We remark that α is independent of the choice of the matrix W. Let β denote the number of *positive* eigenvalues of A. In case $p \neq 0$, let γ be the number of negative μ_j, $j \in \{1, \ldots, p-1\}$ and put $\delta = D_t L(\bar{z})$. Note, in particular, that $\gamma < p$.

Characteristic numbers: $\text{sign}(\alpha)$, β.

Characteristic numbers (*corresponding to* $\mu_p = 1$): γ, $\text{sign}(\delta)$.

We proceed with an analysis of the set Σ. Note that $\bar{x}$ is *not* a critical point for $\mathcal{P}(\bar{t})$. In fact, from (37) we see that the vectors $D_x h_i(\bar{z})$, $D_x g_j(\bar{z})$, $i \in I$, $j \in J_0(\bar{z})$, are *not* linearly independent. In the case $J_0(\bar{z}) = \emptyset$ we are dealing—locally—only with equality constraints. The equality-constrained case is treated extensively in [11, 'Type 3']. Now we will generalize the corresponding results obtained in [11] for the case that $J_0(\bar{z}) \neq \emptyset$. In accordance with the conditions C1–C6 we assume that $J_0(\bar{z}) = \{1, \ldots, p\}$ and, in addition, that $p \geq 1$. In particular, we use the normalization $\mu_p = 1$ (cf. condition C4).

Conditions C2 and C4 imply that for every $q \in J_0(\bar{z})$ the following set is linearly independent:

$$\{D_x h_i(\bar{z}), D_x g_j(\bar{z}), i \in I, j \in J_0(\bar{z}) \setminus \{q\}\}. \tag{43}$$

From condition C5 and (43) it follows that there exists a neighbourhood $\mathcal{O}$ of $\bar{z}$ such that $\{D_x h_i(z), D_x g_j(z), i \in I, j \in J_0(z)\}$ is linearly independent for all $z \in \mathcal{O} \setminus \{\bar{z}\}$. But then, all g.c. points—apart from $\bar{z}$—in some neighbourhood of $\bar{z}$ are critical points. Condition C6 implies in particular that $W^{\mathrm{T}} D_x^{\mathrm{T}} f \neq 0$. Consequently, $\bar{z}$ is not a critical point if we delete any g_q as a constraint function, where $q \in J_0(\bar{z})$ (cf. (43)). In particular we obtain that, in some neighbourhood of $\bar{z}$, at all points of Σ the active

set J_0 is *constant* $(=J_0(\bar{z}))$. But then, the local structure of Σ around $\bar{z}$ is in accordance
with the equality-constrained case as treated in [11], of course including the change
of the indices QI, QCI. As a consequence, locally around $\bar{z}$, the set Σ is a one-
dimensional C^2-manifold, and the parameter t, viewed at as a function on Σ, has
a (nondegenerate) local maximum, resp. local minimum, at $(\bar{x}, \bar{t})$ corresponding to
$\text{sign}(\alpha) = +1$, resp. $\text{sign}(\alpha) = -1$ (cf. [11], where the characteristic numbers have
different names). So, locally around $\bar{z}$, the set Σ can be approximated by means of
a parabola. It remains to compute the change of the *linear* indices LI, LCI, in
relation with the change of the indices QI, QCI. Let z be a critical point, close to
$\bar{z}$, and let $\hat{\mu}_j(z)$ be the Lagrange parameter corresponding to the inequality constraint
function g_j, $j \in J_0(\bar{z})$. Then, it is not difficult to see that $|\hat{\mu}_j(z)|$ tends to infinity as z
tends to $\bar{z}$. Furthermore, taking (37) and condition C4 into account, a moment of
reflection shows that, for $j \neq p$, the sign of $\hat{\mu}_j(z)$ equals $\text{sign}(\mu_j) \cdot \text{sign}(\hat{\mu}_p(z))$, with
μ_j as in (37). Therefore, we may reduce our considerations to the case $I = \emptyset$ and
$p = |J_0(\bar{z})| = 1$.

In the case $I = \emptyset$, $p = 1$, we have the following simplification: $L = g_1$, $T = \mathbb{R}^n$
(cf. C2), $W = \text{identity matrix}$ (without loss of generality), $w = D_x^T f$ and thus,
$A = D_t g_1 \cdot D_x^2 g_1$, $\alpha = D_x f [D_t g_1 \cdot D_x^2 g_1]^{-1} D_x^T f$. In particular, $D_x f \neq 0$ and $D_x^2 g_1$ is
nonsingular (at $z = \bar{z}$).

The critical point equation, with η denoting the corresponding Lagrange param-
eter, becomes

$$\left.\begin{array}{r} D_x f - \eta D_x g_1 = 0 \\ g_1 = 0 \end{array}\right\}, \quad z \neq \bar{z}. \tag{44}$$

Let the symbol $\#$ again denote 'the number of negative eigenvalues'. At a critical
point z $(z \neq \bar{z})$ the quadratic index QI equals $\#(D_x^2 f - \eta D_x^2 g_1|_{\text{Ker } D_x g_1})$, which, in
view of (44), equals

$$\#(D_x^2 f - \eta D_x^2 g_1|_{\text{Ker } D_x f}). \tag{45}$$

Note that in (45) the number $|\eta|$ tends to infinity as z tends to $\bar{z}$. Furthermore, η^{-1}
changes sign at $z = \bar{z}$. Since $D_x^2 g_1(\bar{z})$ is nonsingular and $\alpha \neq 0$, it follows from Lemma
3.2.1a that $D_x^2 g_1(\bar{z})|_{\text{Ker } D_x f(\bar{z})}$ is nonsingular. But then, for z close to $\bar{z}$, the number
in (45) equals

$$\#(-\eta D_x^2 g_1|_{\text{Ker } D_x f}). \tag{46}$$

From (46) it follows that the quadratic index QI of a critical point z, z close to $\bar{z}$,
corresponds to the following list (the partial derivatives being evaluated at $z = \bar{z}$):

$$\left.\begin{array}{l} 1.\ \eta > 0,\ D_t g_1 > 0:\ \text{QI} = \#(-B) \\ 2.\ \eta > 0,\ D_t g_1 < 0:\ \text{QI} = \#(B) \\ 3.\ \eta < 0,\ D_t g_1 > 0:\ \text{QI} = \#(B) \\ 4.\ \eta < 0,\ D_t g_1 < 0:\ \text{QI} = \#(-B) \end{array}\right\}, \quad B = D_t g_1 \cdot D_x^2 g_1|_{\text{Ker } D_x f}. \tag{47}$$

So it remains to relate $\#(\pm D_t g_1 \cdot D_x^2 g_1)$ with the number $\#(\pm D_t g_1 \cdot D_x^2 g_1|_{\text{Ker } D_x f})$.

This can be accomplished by means of Lemma 3.2.1b. In particular we have:

$$\#(D_t g_1 \cdot D_x^2 g_1) = \#(D_t g_1 \cdot D_x^2 g_1|_{\mathrm{Ker}\, D_x f}) + \#(D_x f[D_t g_1 \cdot D_x^2 g_1]^{-1} D_x^T f).$$
(48)

Note that $\#(D_x f[D_t g_1 \cdot D_x^2 g_1]^{-1} D_x^T f) = +1$ (resp. 0) according to $\mathrm{sign}(\alpha) = -1$ (resp. $+1$).

Now we return to the general case. Then, in (48), $D_t g_1 \cdot D_x^2 g_1$ has to be replaced by $D_t L \cdot W^T \cdot D_x^2 L \cdot W$, etc. A few calculations show that the index-relations are as depicted in Fig. 4, where the parabolas are local approximations of the set Σ. If we
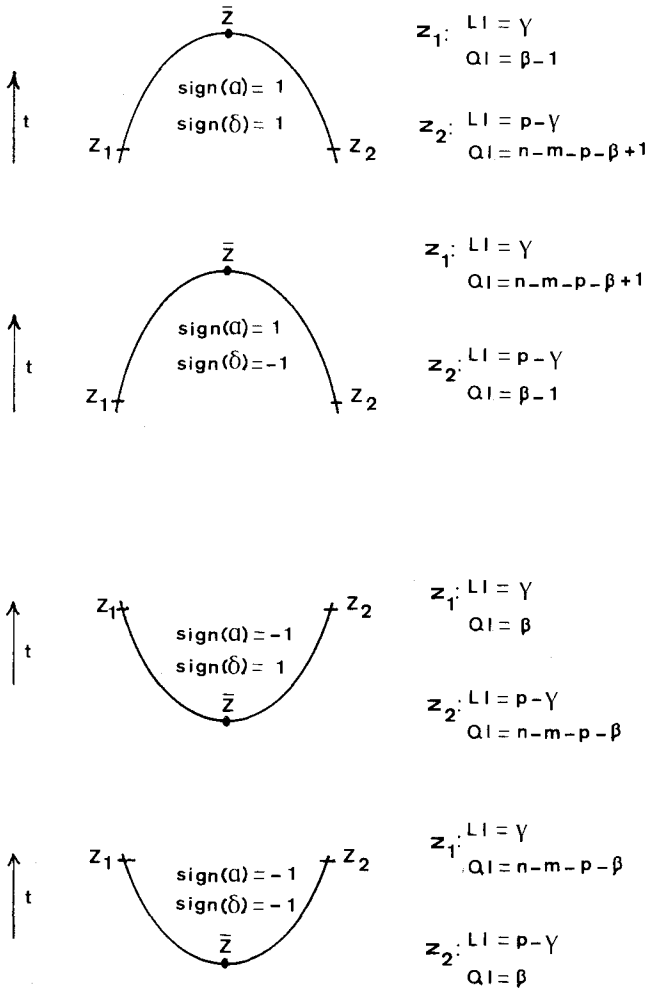
$p = |J_0(\bar{z})|$, $m = |I|$, normalization: $\mu_p = 1$

$LI + LCI = p$, $QI + QCI = n - m - p$

Fig. 4.

would replace f by the function $-f$, this has the following effect on Fig. 4: interchange the index-entries corresponding to z_1, z_2.

3.5. Points of Type 5

A generalized critical point $\bar{z} = (\bar{x}, \bar{t})$ is of Type 5 if the following conditions D1–D4 hold:

D1. $|I| + |J_0(\bar{z})| = n + 1$.   (49)

D2.   The set $\{Dh_i(\bar{z}), Dg_j(\bar{z}), i \in I, j \in J_0(\bar{z})\}$ is linearly independent (derivatives in $\mathbb{R}^{n+1}$).

Since $m = |I| < n$ throughout the paper, (49) implies that $|J_0(\bar{z})| \geq 2$. After renumbering we assume that $J_0(\bar{z}) = \{1, \ldots, p\}$. From D1, D2, we see that there exist λ_i, μ_j, $i \in I$, $j \in J_0(\bar{z})$, not all vanishing (unique up to a common multiple) such that:

$$\sum_{i=1}^{m} \lambda_i D_x h_i(\bar{z}) + \sum_{j=1}^{p} \mu_j D_x g_j(\bar{z}) = 0. \tag{50}$$

D3.   In (50) we have $\mu_j \neq 0$, $j = 1, \ldots, p$.

From D1, D2, it follows that there exist unique numbers α_i, β_j, $i \in I$, $J \in J_0(\bar{z})$, such that

$$Df(\bar{z}) = \sum_{i=1}^{m} \alpha_i Dh_i(\bar{z}) + \sum_{j=1}^{p} \beta_j Dg_j(\bar{z}). \tag{51}$$

Put

$$\Delta_{ij} = \beta_i - \beta_j \cdot \mu_i / \mu_j, \quad i, j = 1, \ldots, p, \tag{52}$$

and let Δ be the $p \times p$ matrix with Δ_{ij} as its (i, j)th element.

D4.   All off-diagonal elements of Δ are unequal to zero.

Put

$$L(z) = \sum_{i=1}^{m} \lambda_i h_i(z) + \sum_{j=1}^{p} \mu_j g_j(z), \tag{53}$$

where λ_i, μ_j satisfy (50). From D2 we see that $D_t L(\bar{z}) \neq 0$. We define:

$$\gamma_j = \text{sign}(\mu_j \cdot D_t L(\bar{z})), \quad j = 1, \ldots, p. \tag{54}$$

By δ_j we denote the number of negative entries in the jth column of Δ, $j = 1, \ldots, p$.

Characteristic numbers: γ_j, δ_j, $j = 1, \ldots, p$.

We proceed with an analysis of the set Σ. From conditions D1–D3 it follows that for every $q \in \{1, \ldots, p\}$ the following set is linearly independent:

$$\{D_x h_i(\bar{z}), D_x g_j(\bar{z}), i \in I, j \in J_0(\bar{z}) \setminus \{q\}\}. \tag{55}$$

Next, conditions D1, D2 and the linear independence of (55) for all $q \in J_0(\bar{z}) = \{1, \ldots, p\}$ imply that there exists a neighbourhood $\mathcal{O}$ of $\bar{z}$ such that $\{D_x h_i(z), D_x g_j(z), i \in I, j \in J_0(z)\}$ is linearly independent for all $z \in \mathcal{O} \setminus \{\bar{z}\}$. But then,

apart from $\bar{z}$, all generalized critical points in some neighbourhood of $\bar{z}$ are critical points. A combination of (50) and (51) implies at $z = \bar{z}$:

$$D_x f = \sum_{i=1}^{m} \left(\alpha_i - \beta_q \cdot \frac{\lambda_i}{\mu_q} \right) D_x h_i + \sum_{j=1}^{p} \left(\beta_j - \beta_q \frac{\mu_j}{\mu_q} \right) D_x g_j, \quad q = 1, \ldots, p. \tag{56}$$

Consequently, for $q = 1, \ldots, p$, we obtain that $\bar{z}$ is a nondegenerate critical point *if we delete g_q as a constraint* (combine (56) and condition D4). We put

$$M_q = \{z \mid h_i(z) = 0, \, g_j(z) = 0, \, i \in I, j \in J_0(\bar{z}) \backslash \{q\}\}, \quad q = 1, \ldots, p, \tag{57}$$

$$M_q^+ = \{z \in M_q \mid g_q(z) \geq 0\}. \tag{58}$$

From conditions D1, D2, and the fact that h_i, g_j are C^3-functions, it follows that, locally around $\bar{z}$, the set M_q is a one-dimensional C^3-manifold, $q = 1, \ldots, p$. Furthermore, *locally around $\bar{z}$, the set Σ is equal to the union* $\bigcup_{q=1}^{p} M_q^+$. A moment of reflection shows that there exists a local C^3-coordinate transformation of a neighbourhood $\mathcal{V}$ of $\bar{z}$ which maps the set $\Sigma \cap \mathcal{V}$ to the union of the coordinate-axes of the orthant $\mathbb{H}^p = \{y \in \mathbb{R}^p \mid y_i \geq 0, \, i = 1, \ldots, p\}$. The indices (LI, LCI, QI, QCI) along $M_q \backslash \{\bar{z}\}$ are equal to $(\delta_q, p - 1 - \delta_q, 0, 0)$, where δ_q is a characteristic number as introduced above (combine the definition of δ_q with (52) and (56)).

It is easily seen that the t-component of the tangent space of M_q at $\bar{z}$ is unequal to zero, $q = 1, \ldots, p$. Furthermore, a short calculation shows that, as t increases and passes the value $\bar{t}$, the set M_q^+ emanates from $(\bar{x}, \bar{t})$, resp. ends at $(\bar{x}, \bar{t})$ according to $\gamma_q = +1$, resp. $\gamma_q = -1$, where γ_q is one of the characteristic numbers (cf. (54)).

4. The Kuhn–Tucker subset

In this section we will look more closely to the subset of Σ consisting of Kuhn-Tucker points. This set has been studied extensively in the important paper [12] under the additional assumption of the so-called Mangasarian–Fromowitz Constraint Qualification (shortly: MFCQ). However, the MFCQ is not a generic condition in optimization problems depending on parameters. With the aid of the foregoing analysis we are able to describe generically the Kuhn-Tucker subset (without explicitly appealing to the MFCQ) and we will discuss those situations where the MFCQ fails to hold.

We will adopt the notations used in the foregoing sections. In particular, recall that $\mathcal{F}$ denotes the set of generic one-parametric programs (cf. Definition 2.1 and Theorem 2.1) and note that Σ^1 stands for that subset of Σ, consisting entirely of nondegenerate critical points.

Definition 4.1. The subset Σ_{KT} (the 'Kuhn–Tucker subset') of Σ is defined to be the *closure* of the set $\{z \in \Sigma^1 \mid LI = 0 \text{ at } z\}$.

Definition 4.2. Let $\bar{z} = (\bar{x}, \bar{t})$ and let $\bar{x}$ be a feasible point for $\mathcal{P}(\bar{t})$, i.e. $\bar{x} \in M(\bar{t})$. The Mangasarian–Fromowitz Constraint Qualification (MFCQ) is said to be satisfied at $\bar{z}$ if both of the following conditions hold:

  a. $\{D_x h_i(\bar{z}), i \in I\}$ is a linearly independent set,
  b. There exists a vector $\xi \in \mathbb{R}^n$ such that

$$D_x h_i(\bar{z})\xi = 0, \quad i \in I, \qquad D_x g_j(\bar{z})\xi > 0, \quad j \in J_0(\bar{z}). \tag{59}$$

Lemma 4.1. Let $(f, h_1, \ldots, h_m, g_1, \ldots, g_s) \in \mathscr{F}$ and $\bar{z} \in \Sigma_{KT}$. Then, at $\bar{z}$, the MFCQ is not satisfied iff

either: $\bar{z}$ is of Type 4, or: $\bar{z}$ is of Type 5, and all μ_j in (50) have the same sign.

Proof. If $\bar{z}$ is of Type 1, 2 or 3, then the set $\{D_x h_i(\bar{z}), D_x g_j(\bar{z}), i \in I, j \in J_0(\bar{z})\}$ is linearly independent, and this obviously implies the validity of the MFCQ. Next, suppose that $\bar{z}$ is of Type 4. If $J_0(\bar{z}) = \emptyset$, then condition C2 implies that $\{D_x h_i(\bar{z}), i \in I\}$ is a linearly dependent set and hence the MFCQ is not satisfied in view of condition a in Definition 4.2. Now, let $J_0(\bar{z}) \neq \emptyset$. Since $\bar{z} \in \Sigma_{KT}$ it follows from Fig. 4 that $\gamma = 0$ (recall: $\gamma < p$). Hence, all μ_j in (37) have the *same* sign (unequal zero, by taking C4 into account). But then it is easily seen that a solution ξ of (59) contradicts (37). So, also in this case the MFCQ is not satisfied. Next, let $\bar{z}$ be of Type 5. Since $J_0(\bar{z}) \neq \emptyset$ in this case, it follows from condition D3 that the linear independence condition a in Definition 4.2 is satisfied. From the well-known alternative theorems of Farkas' type (e.g. [6]) and condition D3 it follows that (59) is *not* solvable iff in (50) all μ_j have the same sign. This proves the lemma. $\square$

Theorem 4.1. Let $(f, h_1, \ldots, h_m, g_1, \ldots, g_s) \in \mathscr{F}$. Then Σ_{KT} is a one-dimensional (piecewise C^2-)manifold with boundary. In particular, $\bar{z} \in \Sigma_{KT}$ is a boundary point iff at $\bar{z}$ we have: $J_0(\bar{z}) \neq \emptyset$ and the MFCQ fails to hold.

Proof. It suffices to consider the structure of Σ_{KT} in a neighbourhood of each of the points of Type 1–5. Locally around a point of Type 1, the set Σ_{KT} is a C^2-manifold (cf. Section 3.1). If $\bar{z}$ is of Type 2, it follows from Fig. 2 that, locally, the structure of Σ_{KT} is one of those depicted in Fig. 5. Now, let $\bar{z}$ be of Type 3. Then, the local structure of Σ_{KT} coincides with Fig. 3. If $\bar{z}$ is of Type 4, then, in view of Fig. 4,
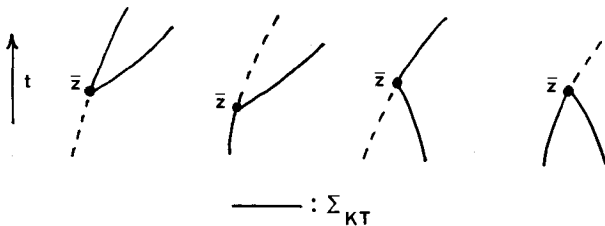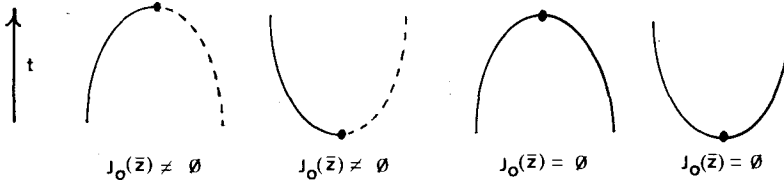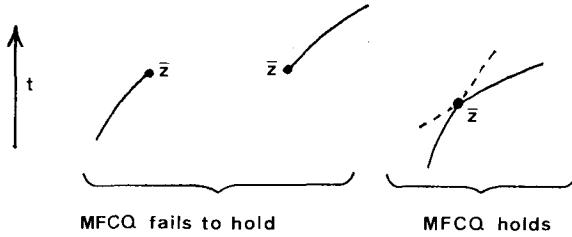

Fig. 5.

$J_0(\bar{z}) \neq \emptyset$ $J_0(\bar{z}) \neq \emptyset$ $J_0(\bar{z}) = \emptyset$ $J_0(\bar{z}) = \emptyset$

Fig. 6.

MFCQ fails to hold   MFCQ holds

Fig. 7.

the set Σ_{KT} has a local structure as depicted in Fig. 6. Note: $\bar{z}$ is a boundary point of Σ_{KT} if $J_0(\bar{z}) \neq \emptyset$, and recall Lemma 4.1.

Finally, let $\bar{z}$ be of Type 5. A short calculation shows that the numbers Δ_{ij} in (52) satisfy the following relation:

$$\Delta_{ij} = -\frac{\mu_i}{\mu_j}\Delta_{ji}, \quad i,j = 1, \ldots, p. \tag{60}$$

In case that all μ_i, $i = 1, \ldots, p$, have the same sign (and thus, in view of Lemma 4.1, the MFCQ is not satisfied), (60) implies: $\text{sign}(\Delta_{ij}) = -\text{sign}(\Delta_{ji})$. But then, if for some $q \in \{1, \ldots, p\}$ we have $\delta_q = 0$, it follows that $\delta_j > 0$ for all $j \in \{1, \ldots, p\}\backslash\{q\}$. Hence, $\bar{z}$ is a boundary point of Σ_{KT}. Now suppose that μ_j, $j = 1, \ldots, p$, do not all have the same sign. Then, the MFCQ is satisfied. From (60) it is easily seen that we have for at most one of the $M_q^+\backslash\{\bar{z}\}$ (cf. (58)), emanating from $\bar{z}$ (resp. ending at $\bar{z}$) as t increases and passes the value $\bar{t}$, that $LI = 0$. A moment of reflection shows that in case $LI = 0$ on $M_q^+\backslash\{\bar{z}\}$, M_q^+ ending at $\bar{z}$, there *exists* exactly one $r \in \{1, \ldots, p\}\backslash\{q\}$ such that M_r^+ emanates from $\bar{z}$ and $LI = 0$ on $M_r^+\backslash\{\bar{z}\}$. Altogether, the local structure of Σ_{KT} around $\bar{z}$ has one of the forms as depicted in Fig. 7. This completes the proof of our theorem. □

Acknowledgement

References

[1] E.L. Allgower and K. Georg, "Predictor-corrector and simplicial methods for approximating fixed points and zero points of nonlinear mappings", in: A. Bachem, M. Grötschel and B. Korte, eds., *Mathematical programming, the state of the art* (Springer-Verlag, Berlin, 1983) pp. 15–56.

[2] V.I. Arnol'd, A.N. Varchenko and S.M. Gusein-Zade, *Singularities of differentiable maps* (Nayka, Moscow, 1982) (in Russian).

[3] A.V. Fiacco, *Introduction to sensitivity and stability analysis in nonlinear programming* (Academic Press, New York, 1983).

[4] A.V. Fiacco, ed., "Sensitivity, stability and parametric analysis", *Mathematical Programming Study* 21 (1984).

[5] J. Guddat and K. Wendler, "On dialogue-algorithms for linear and nonlinear vector optimization from the point of view of parametric programming", in: M. Grauer and A.P. Wierzbicki, eds., *Interactive decision analysis proceedings, Laxenburg, Austria 1983*, Lecture Notes in Economics and Mathematical Systems 229 (Springer-Verlag, Berlin, 1984).

[6] R. Hettich and P. Zencke, *Numerische Methoden der Approximation und semi-infiniten Optimierung* (Teubner Studienbücher, Stuttgart, 1982).

[7] M.W. Hirsch, *Differential topology* (Springer-Verlag, Berlin, 1976).

[8] H.Th. Jongen, P. Jonker and F. Twilt, "On deformation in optimization", *Methods of Operations Research* 37 (1980) 171–184.

[9] H.Th. Jongen, P. Jonker and F. Twilt, "On one-parameter families of sets defined by (in)equality constraints", *Nieuw Archief voor Wiskunde* (3) 30 (1982) 307–322.

[10] H.Th. Jongen, P. Jonker and F. Twilt, *Nonlinear optimization in $\mathbb{R}^n$, I. Morse theory, Chebyshev approximation*, Methoden und Verfahren der mathematischen Physik 29 (Peter Lang Verlag, Frankfurt a.M., 1983).

[11] H.Th. Jongen, P. Jonker and F. Twilt, "One-parameter families of optimization problems: equality constraints", *Journal of Optimization Theory and Applications* 48 (1986) 141–161.

[12] M. Kojima and R. Hirabayashi, "Continuous deformation of nonlinear programs", *Mathematical Programming Study* 21 (1984) 150–198.

[13] M. Marcus and H. Minc, *A survey of matrix theory and matrix inequalities* (Allyn and Bacon Inc., Boston, 1964).

[14] J. Milnor, *Morse theory*, Annals of Mathematics Studies 51 (Princeton University Press, Princeton, NJ, 1963).

[15] D. Siersma, "Singularities of functions on boundaries, corners etc.", *The Quarterly Journal of Mathematics, Oxford Series* (2) 32 (1981) 119–127.

[16] J.E. Spingarn, "On optimality conditions for structured families of nonlinear programming problems", *Mathematical Programming* 22 (1982) 82–92.