A UNIFIED THEORY OF FIRST AND SECOND ORDER CONDITIONS FOR EXTREMA PROBLEMS IN TOPOLOGICAL VECTOR SPACES

by

A. Ben-Tal\textsuperscript{1)} and J. Zöwe\textsuperscript{1)2)3)}

Technical Report #156

April 1979

1) Department of Computer Science, Technion-Israel Institute of Technology, Haifa, Israel

2) Permanent address: Institut für Angewandte Mathematik und Statistik der Universität Würzburg, Am Hubland, D - 8700 Würzburg, W.-Germany.

3) Research supported by DFG (Deutsche Forschungsgemeinschaft).
A UNIFIED THEORY OF FIRST AND SECOND ORDER CONDITIONS FOR EXTREMUM PROBLEMS IN TOPOLOGICAL VECTOR SPACES

by

A. Ben-Tal 1)
J. Zowe 1) 2) 3)

Technical Report #156
April 1979

1) Department of Computer Science
   Technion-Israel Institute of Technology, Haifa, Israel

2) Permanent address: Institut für Angewandte Mathematik
   und Statistik der Universität Würzburg, Am Hubland,
   D - 8700 Würzburg, W. - Germany

3) Research supported by DFG (Deutsche Forschungsgemeinschaft)
ABSTRACT

The paper studies an abstract optimization problem (P) in infinite dimensional spaces. From a general extremality condition, a variety of necessary conditions of first and second order, with or without differentiability assumptions, are derived for special cases of the general problem (P). Classical results are refined and new ones are added. Second order sufficient condition, under differentiability assumptions, are derived as well.

Key words and phrases:
- Optimization in infinite dimensional spaces
- Second order necessary conditions
- Second order sufficient conditions
- Nondifferentiable optimization
- Constraint qualification
Table of contents

1. Introduction
2. The general necessary conditions
3. Support functionals
4. The proof of Theorem 2.1 and some of its consequences
5. Some auxiliary results concerning active constraints
6. Computation of the sets of directions of decrease and feasibility
7. Computation of the set of tangent directions
8. Necessary conditions under differentiability
9. The mathematical programming problem
10. An example
11. Sufficient conditions
12. Summary of necessary and sufficient conditions
13. References
1. Introduction

This paper presents a unified second order theory of optimality conditions for optimization problems in ordered topological vector spaces. The problem studied is given in the form

\[
\begin{align*}
\text{minimize} & \quad f(x) \\
\text{subject to} & \quad -g(x) \in K, \ h(x) = 0, \ x \in X.
\end{align*}
\]

(P)

Here \( f : X \to U, \ g : X \to V \) and \( h : X \to W \) are continuous maps, \( X, \ U, \ V \) and \( W \) are real topological vector spaces, \( K \) is a convex cone in \( V \) with nonempty topological interior \( (\text{int} \ K \neq \emptyset) \) and \( U \) is ordered by a proper convex cone \( C \) with \( \text{int} \ C \neq \emptyset \). Following the usual convention, we write \( u_1 \preceq u_2 \) (or \( u_2 \succeq u_1 \)) if \( u_1 - u_2 \in C \) and \( u_1 \succeq u_2 \) (or \( u_2 \preceq u_1 \)) if \( u_1 - u_2 \in \text{int} \ C \). Then \( \preceq \) is a binary reflexive, antisymmetric and transitive relation on \( U \) invariant under translations and under multiplications with positive scalars. The minimization is to be understood in the \textit{Pareto-sense}, that is, we are looking for points \( x_0 \) in the \textit{feasible set} \( F := \{ x \in X : -g(x) \in K \text{ and } h(x) = 0 \} \) for which there exists a neighborhood \( N(x_0) \) of \( x_0 \) such that for \( x \in N(x_0) \cap F \)

\[ f(x) \preceq f(x_0) \quad \text{implies} \quad f(x) = f(x_0). \tag{1.1} \]

We refer to such points \( x_0 \) as \textbf{optimal} points for \( (P) \). If \( U \) is the real line \( \mathbb{R} \) and \( C \) is the non-negative half line \( \mathbb{R}_+ := \{ \lambda \in \mathbb{R} : \lambda \text{ nonnegative} \} \) then \( (P) \) is a usual minimization problem. Indeed, (1.1) amounts to \( (x \text{ denotes the usual ordering on } \mathbb{R}) \)

\[ f(x) \geq f(x_0) \quad \text{for} \quad x \in N(x_0) \cap F. \tag{1.2} \]
The most important special case of (P) is, of course, the so-called infinite-dimensional mathematical program-
ing problem:

\[ \begin{align*}
U &= \mathbb{R}, \quad C = \mathbb{R}_+, \quad V = \mathbb{R}^n, \quad K = \mathbb{R}_n^+ \ (n \text{ any natural number}). \\
\text{If } g_i, \ i=1,2,\ldots,n, \text{ are the components of } g \text{ then } (P) \text{ becomes:}
\end{align*} \]

\[
\begin{align*}
\text{minimize } & f(x) \\
\text{subject to } & g_i(x) \leq 0 \text{ for } i=1,2,\ldots,n, \quad (MP) \\
& h(x) = 0 \quad \text{and } x \in X.
\end{align*}
\]

As an example for (P), where \( f \) is not a real functional, consider the case \( U = \mathbb{R}^n \) and let \( C \) be the cone defining the lexicographical ordering on \( \mathbb{R}^n \), that is, \( C \) is the set of all vectors in \( \mathbb{R}^n \) whose first non-zero component is positive, together with \( 0_{\mathbb{R}^n} \). Then \( \mathbb{R}^n = C \cup -C \) and thus (1.1) is equivalent to

\[
f(x) \equiv f(x_0) \quad \text{for } x \in N(x_0) \cap F. \quad (1.3)
\]

Necessary and sufficient conditions of first and second order under differentiability assumptions are studied for the finite dimensional case in the books by PIACCO-McCOR-MCK \[8\], LUENBERGER [19], BAZARAA-SHETTY \[1\], HESTENES\[13\] and many others; see also the paper by BEN-TAL \[2\].

In the infinite dimensional case, first order necessary conditions for the problem with and without differentiability assumptions appear in the books by PSHENICHNYI \[25\], GIRSANOV \[30\], HOLMES \[15\], LAURENT \[18\], NEUSTADT \[24\] and e.g. in the papers by DUBOVITSKIĬ-MILYUTIN \[7\], HALKIN \[12\] and BOLTYANSKIĬ \[4\]. For the differentiable case see GUIGNARD \[11\], NASHED \[23\].
Second order conditions in infinite dimensional spaces are rather recent. For the nondifferentiable case see BEN-TAL [3], HOFFMANN-KORNSTAEDT [14] and for the differentiable case BORWEIN [5'] and MAURER-ZOWE [21].

In this paper we treat the very general problem (P) and derive from a general principle a variety of necessary second order (as well as first order) conditions for non-differentiable and for differentiable problems. Sufficient conditions are also treated.

2. The general necessary conditions

We start with the definitions needed for the formulation of our main theorem. A direction \( d \in X \) is called a direction of quasidecrease at \( x \) of the objective function \( f : X \to U \) if for every \( u > 0 \) in \( U \) there is some real \( T > 0 \) such that

\[
f(x+td) \leq f(x)+tu \quad \text{for} \quad 0 < t \leq T.
\]

The element \( d \) is called a quasifeasible direction at \( x \) of the (inequality) constraint \( g : X \to V \) if for every \( v \in \text{int} K \) there is a real \( T > 0 \) such that

\[
g(x+td) \in -K+tv \quad \text{for} \quad 0 < t \leq T.
\]

The cones of all directions of quasidecrease and quasi-feasibility at \( x \) are denoted by \( D_f(x) \) and \( D_g(x) \), respectively.

Further, we call \( z \in X \) a second order direction of decrease of \( f \) at \( x \) with respect to a given \( d \in X \) if there is some \( u > 0 \),
a neighborhood \( N(z) \) of \( z \) and some real \( T > 0 \) such that
\[
f(x+td+t^2z) \leq f(x)-t^2u \quad \text{for all} \quad z \in N(z) \quad \text{and} \quad 0 < t \leq T. \tag{2.1}
\]
The element \( z \in X \) is said to be a second order feasible direction of \( g \) at \( x \) with respect to \( d \in X \) if there is some \( v \in \text{int} \ K \), a neighborhood \( N(z) \) of \( z \) and a real \( T > 0 \) such that
\[
g(x+td+t^2z) \in -K-t^2v \quad \text{for all} \quad z \in N(z) \quad \text{and} \quad 0 < t \leq T. \tag{2.2}
\]
The sets of all \( z \) satisfying (2.1) and (2.2), respectively, are denoted by \( Q_f(x,d) \) and \( Q_g(x,d) \). Obviously, \( Q_f(x,d) \) and \( Q_g(x,d) \) are open. We put
\[
D_f^<(x) := Q_f(x,0) \quad \text{and} \quad D_g^<(x) := Q_g(x,0) \tag{2.3}
\]
and note that these are exactly the cones of first order directions of decrease and feasibility at \( x \), respectively, studied in the theory of first order conditions (see e.g. GIRSANOV [10] or LAURENT [18]). An easy continuity argument shows
\[
Q_f(x,d) = X \quad \text{for} \quad d \in D_f^<(x) \quad \text{and} \quad Q_g(x,d) = X \quad \text{for} \quad d \in D_g^<(x). \tag{2.4}
\]

Typically, in a second order theory one will consider second order sets \( Q_f(x,d) \) (and \( Q_g(x,d) \)) for \( d \) belonging to \( D_f(x) \) but not to \( D_f^<(x) \). The set of all such directions, denoted by \( D_f(x) \), is called the critical cone of \( f \) at \( x \)
\[
D_f^<(x) = D_f(x) \setminus D_f^<(x). \tag{2.5}
\]
Analogously, we define
\[
D_g^<(x) = D_g(x) \setminus D_g^<(x).
\]
Next, the first and second order sets corresponding to the equality constraint are considered. We call a function
\( r(\cdot) : (0,\infty) \to X \) a curve of order \( o(t^k)(r(t) \sim o(t^k)) \) if for every neighborhood \( N \) of the origin in \( X \) there is a real \( T_1 > 0 \) such that \( t^{-k}r(t) \in N \) for \( 0 < t \leq T_1 \). A vector \( z \in X \) is said to be a second order tangent direction of \( h : X \to W \) at \( x \) with respect to a given \( d \in X \) if there is a real \( T > 0 \) and a curve \( r(t) \sim o(t^2) \) such that
\[
h(x+td+t^2z+r(t)) = 0, \quad \text{for} \quad 0 < t \leq T.
\]
The set of all such \( z \) is denoted by \( V_h(x,d) \). We put
\[
T_h(x) := V_h(x,0)
\]  
(2.6)
and note that this is exactly the cone of all first order tangent directions studied in the classical first order theory.

For our main theorem we need a certain well-behavior of \( Q_f(x,d), Q_g(x,d) \) and \( V_h(x,d) \). Let \( d \in X \) be given. Then we call \( f \) \( d \)-regular at \( x \) if \( Q_f(x,d) \) is nonempty and convex. Analogously, we say \( g \) is \( d \)-regular if \( Q_g(x,d) \) is nonempty and convex; \( h \) is \( d \)-regular if \( V_h(x,d) \) is nonempty and convex. If \( f \) is \( d \)-regular at \( x \) for all \( d \in D_f(x) \) then \( f \) is called regular at \( x \) (because of (2.4) it is sufficient to demand this for \( d \in D_f^W(x) \)). Correspondingly \( g \) is called regular at \( x \) whenever \( g \) is \( d \)-regular for all \( d \in D_g(x) \), and \( h \) is said to be regular at \( x \) if \( h \) is \( d \)-regular for all \( d \in T_h(x) \).

Finally we associate with a subset \( S \) of \( X \) its so-called support functional \( \delta^*(\cdot|S) \) defined on the topological dual.
of $X$ and with values in the extended real line $\mathbb{R} \cup \{+\infty\}$:

$$\delta^*(x^*\mid S) = \sup_{x^* \in S} x^* x$$

(If $S = \emptyset$ then by convention $\delta^*(\cdot\mid S) = -\infty$). The effective domain of $\delta^*(\cdot\mid S)$ is denoted by $\Lambda(S)$, i.e.

$$\Lambda(S) = \{x^* \in X^* : \delta^*(x^*\mid S) < \infty\}.$$  

It is easily checked that $\delta^*(\cdot\mid S)$ is a positively homogenous closed convex function and that $\Lambda(S)$ is a convex cone, closed in the weak topology (i.e. the topology under which $X^*$ is the dual of $X$).

If $S^+$ stands for the polar of $S$,

$$S^+ = \{x^* \in X^* : x^* x \geq 0 \text{ for all } x \in S\},$$

then, as is easily seen,

$$\Lambda(S) = -S^+ \quad \text{whenever } S \text{ is a cone}.$$  

$$\delta^*(x^*\mid S) = \begin{cases} 0 & \text{if } x^* \notin \Lambda(S) \\ \infty & \text{otherwise} \end{cases} \quad (2.8)$$

We are now in a position to formulate our main theorem.

**Theorem 2.1** Let $x_0$ be an optimal point for problem (P). Then for every

$$d \in D_f(x_0) \cap D_g(x_0) \cap T_h(x_0), \quad (2.9)$$

for which $f$, $g$, and $h$ are $d$-regular, there correspond continuous linear functionals on $X$

$$l_f \in \Lambda(Q_f(x_0,d)), \ l_g \in \Lambda(Q_g(x_0,d)), \ l_h \in \Lambda(V_h(x_0,d)), \quad (2.10)$$

not all zero, which satisfy the Euler-Lagrange equation

$$l_f + l_g + l_h = 0 \quad (2.11)$$

and the Legendre inequality

$$\delta^*(l_f\mid Q_f(x_0,d)) + \delta^*(l_g\mid Q_g(x_0,d)) + \delta^*(l_h\mid V_h(x_0,d)) \leq 0 \quad (2.12)$$
The proof of the above necessary conditions will be postponed to section 4. Following the proof we will give some additional remarks. In section 8 and 9 we will show how the above conditions specialize for (MP), in particular for the differentiable case. In section 10 we will apply Theorem 2.1 to a non-differentiable problem. Sufficient conditions are considered in section 11.

3. Support functionals

Lemmas 3.2 and 3.3 below are the main auxiliary results needed for the proof of Theorem 2.1. Lemma 3.2 characterizes the support functional of the intersection of finitely many convex sets. Lemma 3.3 expresses in a dual way the property that finitely many convex sets have an empty intersection. The proof of both theorems can be most easily obtained by using the following extension theorem which is of interest in itself. For $S$ being a convex cone and $u = 0$ this result reduces to a well-known extension theorem for monotone linear functionals by KREIN-RUTMAN [17].

**Lemma 3.1** Let $M$ be a linear subspace of the real topological vector space $X$, $x^*_M$ a real linear functional on $M$ and $S$ a convex subset of $X$. Suppose $u \in \mathbb{R}$ is such that $x^*_M x \leq u$ for all $x \in M \cap S$. If $M \cap \text{int } S \neq \emptyset$ then there is some $x^* \in X^*$ such that $x^* x = x^*_M x$ for all $x \in M$ and $x^* x \leq u$ for all $x \in S$.

**Proof:** Consider in the topological product space $X \times \mathbb{R}$ the
convex set

\[ A := \{ (m-s) : m \in M, s \in S, \alpha \leq x^*_M \} \]

By assumption there is some \( \bar{x} \in M \cap S \) and a 0-neighborhood \( N \) in \( X \) such that \( \bar{x} + N \subset S \). Thus

\[ \{ \left( \frac{\bar{x}-(\bar{x}+x)}{\alpha} \right) : x \in N, \alpha \leq x^*_M \} \]

is a subset of \( A \) showing that \( \text{int} \ A \neq \emptyset \). Since

\[ A \cap \{ \left( \frac{\gamma}{\beta} \right) : \beta > \mu \} = \emptyset \]

there exist a closed hyperplane in \( X \times R \) separating these two convex sets, i.e., with some \( x^* \in X^* \) and \( \gamma \in R ((x^*,\gamma) \neq (0,0)) : \)

\[ x^*(m-s)+\gamma \alpha \geq \gamma \beta \quad \text{for all} \quad m \in M, \quad s \in S, \quad \alpha \leq x^*_M, \quad \beta > \mu. \]

It follows \( \gamma \leq 0 \) and without loss of generality \( \gamma = -1 \).

If \( \gamma = 0 \) then, \( x^*(\bar{x}-(\bar{x}+N)) \geq 0 \) i.e. \( x^* = 0 \) and thus \( (x^*,\gamma) = (0,0) \). We get

\[ (x^*-x^*_M)m - x^*s \geq -\mu \quad \text{for all} \quad m \in M \quad \text{and} \quad s \in S. \]

Now, this can hold only if \( x^*_m = x^*_M \) for all \( m \in M \). But then also \( x^*s \leq \mu \) for \( s \in S \).

\[ \text{qed} \]

Recall the definition of the support functional of a subset \( S \) of \( X \)

\[ \delta^*(x^*|S) = \sup\{x^*x : x \in S\}, \quad x^* \in X^*. \]

Moreover, recall that \( \Lambda(S) \) denotes the effective domain of \( \delta^*(\cdot|S) \). The appropriateness of the \( \delta^* \)-notation follows from

the observation that \( \delta^*(\cdot|S) \) is the conjugate of the indicator function of \( S \), \( \delta(x|S) = \begin{cases} 0 & \text{if} \quad x \in S, \\ \infty & \text{otherwise}. \end{cases} \)

Actually, our
next result can be proved by applying to $\delta$ and $\delta^*$ the so-called infimum-convolution formula of the theory of conjugate functions (see MOREAU [22]). We give here a more direct proof using Lemma 3.1.

**Lemma 3.2** Let $S_1, \ldots, S_n$ be convex subsets of $X$ and let $x^* \in \Lambda(\bigcap_{i=1}^n S_i)$. If $\bigcap_{i=1}^n \text{int} S_i \neq \emptyset$ then

$$\delta^*(x^*|\bigcap_{i=1}^n S_i) = \min \left\{ \sum_{i=1}^n \delta^*(x^*|S_i) : x^* = x^1 + \ldots + x^n, x^i \in \Lambda(S_i) \right\}.$$  

**Proof:** Let $x^*_i \in \Lambda(S_i)$ for $i=1, \ldots, n$ and $x^* = x^*_1 + \ldots + x^*_n$. Then

$$\delta^*(x^*|\bigcap_{i=1}^n S_i) \leq \sum_{i=1}^n \delta^*(x^*_i|S_i) \leq \sum_{i=1}^n \delta^*(x^*_i|S_i). \quad (3.1)$$

It remains to show that equality holds in (3.1) for suitable $x^*_i \in \Lambda(S_i)$ with $x^*_1 + \ldots + x^*_n = x^*$. To this end we consider in the topological product space $\tilde{X} := \prod_i X$ the convex set

$$\tilde{S} := \bigcap_{i=1}^n S_i$$

and the linear subspace $M := \{x=(x_1, \ldots, x_n) \in \tilde{X} : x_i - \ldots - x_n \forall i \}$. In terms of $\tilde{X}$, $\tilde{S}$ and $M$ the assumption

$$\bigcap_{i=1}^n \text{int} S_i \neq \emptyset$$

is equivalent to $M \cap \text{int} \tilde{S} \neq \emptyset$.

On $M$ we define a linear functional $\tilde{x}_M$ by

$$\tilde{x}_M := x^* x \text{ for } x=(x_1, \ldots, x_n) \in M.$$  

Then

$$\tilde{x}_M \tilde{x} = \delta^*(x^*|\bigcap_{i=1}^n S_i) \text{ for } \tilde{x} \in M \cap \tilde{S}.$$  

We apply Lemma 3.1 to see that there is some $\tilde{x}^* \in \tilde{x}_M$ such that
\[ x^* \mathcal{X} = x^*_\mathcal{X}_M \quad \text{for } \mathcal{X} \in \mathcal{M} \quad (3.2) \]
\[ x^* \mathcal{X} \leq u \quad \text{for } \mathcal{X} \in \mathcal{S} \quad (3.3) \]

Now \( \mathcal{X}^* = (\prod_{i=1}^n X)_\mathcal{X}_i^{\mathcal{X}} \), i.e. \( \mathcal{X}^* = x_1^* + \ldots + x_n^* \) with suitable \( x_i \in X^* \). Furthermore (3.2) and (3.3) become

\[ x_1^* x_1 + \ldots + x_n^* x_n = x_\mathcal{X}_M(x, \ldots, x) = x^* x \quad \text{for all } x \in X \]

\[ x_1^* x_1 + \ldots + x_n^* x_n \leq \mu \quad \text{for all } x_i \in S_i \text{ and } i=1, \ldots, n. \]

Hence

\[ x^* = x_1^* + \ldots + x_n^* \quad \text{with } x_1^* \in \Lambda(S_1) \text{ and } \sum_{i=1}^n \delta^*(x_i^*|S_i) \leq \delta^*(x^*|S_{n+1}) \]

Together with (3.1) this proves the assertion. \( \text{qed} \).

The next result is given in a different but equivalent formulation e.g. in TIKHOMIROV [16]. Our proof, however, is different and relies on the two preceding lemmas.

**Lemma 3.3** Let \( S_1, \ldots, S_n, S_{n+1} \) be non-empty convex subsets of \( X \), where \( S_1, \ldots, S_n \) are open. Then

\[ \bigcap_{i=1}^{n+1} S_i = \emptyset \quad (3.4) \]

if and only if there are \( x_i^* \in \Lambda(S_i) \), \( i=1, \ldots, n+1 \), not all zero, such that

\[ x_1^* + x_2^* + \ldots + x_{n+1}^* = 0 \quad (3.5) \]

and

\[ \delta^*(x_1^*|S_1) + \delta^*(x_2^*|S_2) + \ldots + \delta^*(x_{n+1}^*|S_{n+1}) \leq 0. \quad (3.6) \]

**Proof.** Sufficiency: Suppose (3.5) and (3.6) hold but not (3.4), e.g., \( x_0 \in \bigcap_{i=1}^{n+1} S_i \). Put \( \mathcal{S}_i := S_i - x_0 \) for \( i=1, \ldots, n+1 \).
Then $0 \in \tilde{S}_i$ and thus $\delta^*(x_j^*|\tilde{S}_j) > 0$ for all $i$. Now $x_j^* \neq 0$
for at least one $j < n$ (otherwise $-x_{n+1}^* = x_1^* + \ldots + x_n^* = 0$) and thus $x_1^* = \ldots = x_{n+1}^* = 0$). But $0 \in \text{int } \tilde{S}_j$, so $\delta^*(x_j^*|\tilde{S}_j) > 0$ for this special $j$. Thus

$$\sum_{i=1}^{n+1} \delta^*(x_i^*|S_i) = \sum_{i=1}^{n+1} \delta^*(x_i^*|\tilde{S}_i) + \sum_{i=1}^{n+1} x_i^* x_0$$

$$= \sum_{i=1}^{n+1} \delta^*(x_i^*|\tilde{S}_i) > 0.$$ We reach a contradiction to (3.6).

Necessity: Put $S := \bigcap_{i=1}^{n} S_i$ and suppose for the moment that $S \neq \emptyset$. Then $S$ and $S_{n+1}$ are non-empty convex sets, $S$ is open and by (3.4), $S \cap S_{n+1} = \emptyset$. Hence by a standard separation theorem we have with a suitable $x^* \in X^*$, $x^* \neq 0$,

$$x^* x < x^* \bar{x}$$

for all $x \in S$, $\bar{x} \in S_{n+1}$

i.e. $x^* \in \Lambda(S)$, $-x^* \in \Lambda(S_{n+1})$ and

$$\delta^*(x^*|S_i) + \delta^*(-x^*|S_{n+1}) < 0.$$ We apply Lemma 3.2 to see that with suitable $x_i^* \in \Lambda(S_i)$, $i=1,2,\ldots, n$,

$$x^* = x_1^* + \ldots + x_n^*$$

$$\delta^*(x^*|S_i) = \sum_{i=1}^{n} \delta^*(x_i^*|S_i).$$

(3.5) and (3.6) follow if we put $x_{n+1}^* := -x^*$.

Note that, if $S = \bigcap_{i=1}^{m} S_i = \emptyset$, then there exists $1 < m < n$ such that $\bigcap_{i=1}^{m} S_i \neq \emptyset$ and $\bigcap_{i=1}^{m+1} S_j = \emptyset$. We apply the result just proved with $n$ replaced by $m$ and put $x_{m+1}^* = \ldots = x_{n+1}^* = 0$.

qed
For later use we add the following

\[ n+1 \]\nRemark. If \( \bigcap_{i=2}^{n+1} S_i = \emptyset \) then \( x_1^* \neq 0 \) in Lemma 3.3. To prove this let, \( x_0 \in \bigcap_{i=2}^{n+1} S_i \) and suppose (3.5) and (3.6) hold with \( x_1^* = 0 \). Copying the first part of the above proof one gets a contradiction to (3.6).

Suppose the sets \( S_1, \ldots, S_{n+1} \) in Lemma 3.3 are (in addition) cones. Then (3.6) becomes the trivial inequality 0 \( \leq 0 \) (see (2.8)) and Lemma 3.3 reduces to a well-known theorem on cones (see e.g. GIRSANOV [10, p. 37]).

Corollary 3.4 Let \( S_1, \ldots, S_n, S_{n+1} \) be non-empty convex cones in \( X \) and suppose \( S_1, \ldots, S_n \) are open. Then

\[ \bigcap_{i=1}^{n+1} S_i = \emptyset \]

if and only if there are \( x_1^* \in S_1^+, \ldots, x_n^* \in S_n^+, \) not all zero such that

\[ x_1^* + \cdots + x_{n+1}^* = 0 \]

We close this section with a result which will be needed in sections 6 and 7 to compute the second order sets in the differentiable case. It reduces (for \( S \) being a cone) to the so-called MINKOWSKI-FARKAS-Lemma (see e.g. GIRSANOV [10, p. 70]).

Lemma 3.5 Let \( A : X \to U \) be a continuous linear operator with range \( R(A) \) and let \( S \) be a non-empty convex subset of \( U \). Put

\[ A^{-1}S := \{ x \in X : Ax \in S \} \]
and suppose $x^* \in \Lambda(A^{-1}S)$. If either of the following conditions holds

(i) $R(A) \cap \text{Int } S \neq \emptyset$ (SLATER condition)

(ii) $X$ and $U$ are Banach-spaces and $A$ is onto (range-condition), then

$$\delta^*(x^*|A^{-1}S) = \min\{\delta^*(u^*|S) : x^* = u^* \cdot A, u^* \in \Lambda(S)\},$$

(here $\delta^*(\cdot|A^{-1}S)$ is defined on $X^*$ and $\delta^*(\cdot|S)$ on $U^*$).

**Proof** Let $\mu := \delta^*(x^*|A^{-1}S)$. Then the system

$$\begin{align*}
Ax & \in S \\
x^*x & > \mu
\end{align*}$$

has no solution $x \in X$. This is equivalent to

$$\left\{(\frac{Ax-s}{a}) : x \in X, s \in S, a < x^*x\right\} \cap \left\{\left(\begin{array}{c} 0 \\ u \end{array}\right)\right\} = \emptyset.$$  

Let us call the first set $B$. We want to apply a separation argument. To this we have to guarantee that $B$ has a non-empty interior in the topological product $U \times \mathbb{R}$. If (i) holds then there is some $\bar{x} \in X$ and a $0$-neighborhood $N$ in $U$ such that $A\bar{x} - N \subset S$, i.e., $N \subset A\bar{x} - S$. This implies

$$\left\{(\frac{Ax-x^*}[\varepsilon, 2\varepsilon]) \right\} \subset B \text{ for } \varepsilon > 0$$

and thus $\text{int } B \neq \emptyset$. Suppose (ii) holds. Then the open mapping theorem (see e.g. SCHAEFER [26 p. 77]) states that

$$\{Ax : \|x\| < 1\}$$

is a non-empty open set in $U$. If we fix some $\bar{s} \in S$ then

$$\left\{(\frac{Ax-x}{a}) : \|x\| < 1, a < \inf x^*x\right\}$$

is an open set contained in $B$. Again we see $\text{int } B \neq \emptyset$. Hence (i) as well as (ii) guarantees that there is a closed hyper-
plane in $U \times \mathbb{R}$ separating $B$ and $(0)$, i.e., with some $u^* \in U^*$ and $\beta \in \mathbb{R}$ $((u^*, \beta) \neq (0,0))$:

$$u^*(Ax - s) + \beta x \geq \delta \mu \quad \text{for} \quad x \in X, \ s \in S, \ \alpha < x^*x.$$  \hfill (3.7)

Obviously, $\beta \leq 0$. Suppose $\beta = 0$. Then

$$u^*u \geq 0 \quad \text{for all} \quad u \in R(A)-S. \hfill (3.8)$$

If (ii) holds then $R(A)-S = Y$. If (i) holds, then $R(A)-S$ contains a $0$-neighborhood (see above). In both cases we get from (3.8), $u^* = 0$ which contradicts $u^* = 0$. Hence we may assume $\beta = -1$. We get from (3.7)

$$u^*Ax - u^*s - x^*x \geq -\mu \quad \text{for} \quad x \in X, \ s \in S.$$  

This shows $x^* = u^*A$ and

$$\delta^* (u^*|S) \leq \mu.$$  

The assertion follows since the converse inequality

$$\mu \leq \delta^* (u^*_o|S)$$

is valid for any $u^*_o$ such that $x^* = u^*_oA$. \hfill qed

4. The proof of Theorem 2.1 and some of its consequences

We are now in a position to give the Proof of Theorem 2.1. Let $d \in X$ be a fixed vector satisfying (2.9) and suppose $f$, $g$ and $h$ are $d$-regular at $x_0$. Then the second-order sets $Q_f(x_0,d)$, $Q_g(x_0,d)$ and $V_h(x_0,d)$ are convex and non-empty. Going back to the definitions it is easy to see that, moreover, $Q_f(x_0,d)$ and $Q_g(x_0,d)$ are open. The assertion follows from Lemma 3.3 if we can show that

$$Q_f(x_0,d) \cap Q_g(x_0,d) \cap V_h(x_0,d) = \emptyset.$$  \hfill (4.1)
Assume to the contrary that $z$ belongs to the intersection (4.1). Then there exists a curve $r(t) \sim o(t^2)$, some $u > 0$ in $U$, $v \in \text{int } K$, neighborhoods $N_1$ and $N_2$ of $z$ and real positive $T_1$, $T_2$, $T_3$ such that

\begin{align*}
&f(x_0 + td + t^2z) \leq f(x_0) - t^2u \text{ for } \overline{z} \in N_1 \text{ and } 0 < t \leq T_1, \\
g(x_0 + td + t^2z) \in - K - t^2v \text{ for } \overline{z} \in N_2 \text{ and } 0 < t \leq T_2, \\
h(x_0 + td + t^2z + r(t)) = 0 \text{ for } 0 < t \leq T_3.
\end{align*}

Since $r(t) \sim o(t^2)$ we can choose a positive $T \leq \min\{T_1, T_2, T_3\}$ small enough such that

\[
\overline{z}(t) := z + t^2r(t) \in N_1 \cap N_2 \text{ for } 0 < t \leq T.
\]

It follows (for the first point see Proposition 5.1 (ii))

\begin{align*}
f(x_0 + td + t^2z(t)) - f(x_0) \in & - C \setminus \{0\} \text{ for } 0 < t \leq T, \\
g(x_0 + td + t^2z(t)) \in & - K \text{ for } 0 < t \leq T, \\
h(x_0 + td + t^2z(t)) = & 0 \text{ for } 0 < t \leq T.
\end{align*}

Obviously, this contradicts the optimality of $x_0$. q.e.d.

Theorem 2.1 contains a first order condition as a special case. We state this as

**Corollary 4.1** Let $x_0$ be an optimal point for problem (P) and suppose $f$, $g$ and $h$ are $O$-regular at $x_0$. Then there are continuous linear functionals on $X$

\[ l_f \in D_f^c(x_0)^+, \quad l_g \in D_g^c(x_0)^+, \quad l_h \in D_h^c(x_0)^+ , \quad (4.2) \]

not all zero, which satisfy

\[ l_f + l_g + l_h = 0. \quad (4.3) \]
Proof. Choose in Theorem 2.1, \( d = 0 \). Then (2.9) holds. It follows from (2.3) and (2.8) that in (2.10) the set \( \Lambda(Q_f(x_0,0)) \) becomes \( D_f^<(x_0)^+ \); analogously for \( g \) and \( h \). (2.12) reduces to the trivial inequality \( 0 \leq 0 \).

qed

Remark 1. The difference between the 1st and 2nd order-conditions reflects the difference between Lemma 3.3 and Corollary 3.4. (4.2) and (4.3) express in a dual way that the intersection of the 1st order cones \( D_f^<(x_0), D_g^<(x_0) \) and \( T_h(x_0) \) is empty. (2.10) - (2.12) do the same for the 2nd order sets \( Q_f(x_0,d), Q_g(x,d) \) and \( V_h(x_0,d) \). The additional inequality (2.12) stems from the fact that, in general, these sets will not be cones (see the example of Section 10).

Remark 2. It should be noted that the multipliers \( l_f, l_g \) and \( l_h \) in Theorem 2.1 depend on the direction \( d \). In Ben-Tal [2] an example is given where there are no fixed multipliers satisfying (2.10) - (2.12). Second order conditions, where the multipliers depend on \( d \) seem to have been first given by COX [6] (for problem (MF) with \( \dim X < \infty \)).

Remark 3. Suppose \( d \) satisfies the assumptions of Theorem 2.1. If \( CQ(d); Q_g(x_0,d) \cap V_h(x_0,d) \neq \emptyset \), then the corresponding \( l_f \) in (2.10) and (2.11) is not the zero functional. This follows from the above proof and the Remark following Lemma 3.5.

Under some differentiability assumptions Theorem 2.1 can
be given a much more explicit form. To this, one has to compute the first and second order sets appearing in Theorem 2.1. This will be done in sections 6 and 7.

5. Some auxiliary results concerning active constraints

In the mathematical programming problems, where one has only a finite number of real inequality constraints $g_i(x) \leq 0$ for $i=1,2,\ldots,n$, the optimality conditions for a point $x_0$ are expressed in terms of the index set $I(x_0)$ of active constraints at $x_0$,

$$I(x_0) := \{ i \in \{1,\ldots,n\} : g_i(x_0) = 0 \}.$$  

(5.1)

In second order optimality conditions another (smaller) index set $J(x_0,d)$ is used for $d \in X$ such that $f'(x_0)d \leq 0$ and $g_i'(x_0)d \leq 0$ for $i \in I(x_0)$ (see Theorem 9.1)

$$J(x_0,d) := \{ i \in I(x_0) : g_i'(x_0)d = 0 \}.$$  

(5.2)

In this section we introduce a corresponding concept of "activity" for the more general constraint $g(x) \in -K$ (compare ZOWE [28]).

We start with some simple observations. Suppose, $v$ is an interior point of the convex cone $K$. Then for a suitable 0-neighborhood $N$ in $V$

$$v + N \subset K.$$  

(5.3)

It follows $\lambda v + \lambda N \subset \lambda K = K$ for $\lambda > 0$, i.e., $\lambda v \in \text{int} K$. Furthermore, (5.3) shows

$$v + K \subset (v+N) + K \subset K + K = K.$$
We may assume \( N \) to be open. Then also \( v + N + K = \bigcup_{k \in K} (v+k) + N \) is open and thus
\[
v + K \subseteq \text{int} K.
\]
Moreover, we can also choose \( N \) to be symmetric, \( N = -N \).

Then (5.3) implies
\[
N = N \cap -N \subseteq (-v+K) \cap (v-K).
\]
Hence \( N \) is contained in the order interval
\[
[-v,v] := (-v+K) \cap (v-K).
\]

We summarize these points in

**Proposition 5.1:** Let \( v \notin \text{int} K \). Then
(i) \( \lambda v \notin \text{int} K \) for all real \( \lambda > 0 \),
(ii) \( v + K \subseteq \text{int} K \),
(iii) \( [-v,v] \) is a neighborhood of the origin in \( V \).

When computing \( Q_g(x,d) \) we will work with a cone generated by \( K + a \) where \( a \) is some fixed element in \( V \). This shifting will allow us to give the concept of 'activity' a meaning even if \( K \) is not the positive orthant \( \mathbb{R}^n_+ \). For convenience let us write \( K_a \) for the closure of the conical hull of \( K+a \), that is (let \( \text{cone} A \) denote the conical hull of \( A \)):
\[
K_a = \text{cl} \text{cone}(K+a) = \text{cl}\{\lambda(k+a) : k \in K, \lambda \geq 0\}.
\]
Obviously,
\[
\text{cone}(K+a) \subseteq K + \{\lambda a : \lambda \geq 0\}.
\]
If \( a \in -K \), then even equality holds. Hence
\[
K_a = \text{cl}(K+\{\lambda a : \lambda \geq 0\}) \quad \text{whenever} \quad a \notin -K.
\]

We have to repeat the above construction. With another element \( b \) in \( V \) we write \( K_{a,b} \) for the closure of the
conical hull of $K_a + b$, i.e.,

$$K_{a,b} = \text{cl} \{(k_a + b) : k_a \in K_a, \lambda \geq 0\}.$$

As an extension of (5.5) one has

**Lemma 5.2** Let $a \in -K$ and $b \in -K_a$. Then

$$K_{a,b} = \text{cl} (K + \{\lambda a + \mu b : \lambda \geq 0, \mu \geq 0\}).$$

**Proof** Put $\{a\} := \{\lambda a : \lambda \geq 0\}$ and $\{b\} := \{\mu b : \mu \geq 0\}$. We apply (5.5) twice to see

$$K_{a,b} = \text{cl} (K_a + \{b\})$$

$$= \text{cl} (\text{cl}(K + \{a\}) + \{b\}).$$

(5.6)

Now one has with $A := K + \{a\} + \{b\}$,

$$\text{int} A \subset \text{cl}(K + \{a\}) + \{b\} \subset \text{cl} A.$$

Here $A$ is convex and, by our general assumption on $K$, $\text{int} A \neq \emptyset$ (int $A \supset \text{int} K + \{a\} + \{b\} \neq \emptyset$). But then (see e.g. SCHAEFER [26p. 38])

$$\text{cl} \text{int} A = \text{cl} A.$$

We get $\text{cl} (\text{cl}(K + \{a\}) + \{b\}) = \text{cl}(K + \{a\} + \{b\})$. Together with (5.6) this proves the assertion.

qed

Typically, we will consider $K_{a,b}$ for $a := g(x)$ and $b := g'(x)d$ ($g'(x)$ denotes the Fréchet-derivative) where $x$ and $d$ are given elements in $X$. Let us shortly discuss the special case $V = \mathbb{R}^n$ and $K = \mathbb{R}_+^n$. Going back to the definitions of $K_a$ and $K_{a,b}$ one easily verifies (recall that $I(x)$ and $J(x,d)$ are defined in (5.1) and (5.2)):

**Proposition 5.3** Let $V = \mathbb{R}^n$, $K = \mathbb{R}_+^n$ and let $x$, $d \in X$ be
given. Suppose $g'(x)$ exists.

(i) If $g(x) \in -K$ then

$$K_g(x) = \prod_{i=1}^{n} S_i \quad \text{where} \quad S_i = \begin{cases} \mathbb{R}^+ & \text{if } i \in I(x) \\ \mathbb{R} & \text{otherwise} \end{cases}$$

(ii) If $g(x) \in -K$ and $g'(x) d \in -K g(x)$ then

$$K_g(x), g'(x) d = \prod_{i=1}^{n} S_i \quad \text{where} \quad S_i = \begin{cases} \mathbb{R}^+ & \text{if } i \in J(x,d) \\ \mathbb{R} & \text{otherwise} \end{cases}$$

The first part of Proposition 5.3 says:

$$g'(x) d \in -K g(x) \quad \text{if and only if } g_i'(x) d \leq 0 \quad \text{for } i \in I(x).$$

Hence the left-hand side is actually a condition only for the active components of $g$. But note that the left-hand side makes sense for arbitrary $K$, and thus can be interpreted as the attempt to give the concept of activity a meaning for arbitrary cones $K$. We illustrate this by another example where $K$ is not the positive orthant but can be expressed by finitely many real inequalities.

**Example** Let $V = \mathbb{R}^3$ and let $K$ be the 'icecream-cone',

$$K = \{ v \in \mathbb{R}^3 : v_1^2 + v_2^2 \leq v_3^2, \; v_3 \geq 0 \}.$$ 

Let $g : X \rightarrow \mathbb{R}^3$ be a differentiable function and consider the constraint $g(x) \in -K$. If we introduce another function $q : X \rightarrow \mathbb{R}^2$ by

$$q(x) := \begin{pmatrix} g_1^2(x) + g_2^2(x) - g_3^2(x) \\ -g_3(x) \end{pmatrix}$$

then

$$g(x) \in -K \quad \text{if and only if } q_i(x) \leq 0 \quad \text{for } i = 1, 2.$$  \hspace{1cm} (5.7)
Suppose \( g(x_0) = (-1, 0, -1)^T \). Some calculation shows that 
\[
K_g(x_0) = \{ v \in \mathbb{R}^3 : v_3 \geq v_1 \}.
\]
Hence 
\[
g'(x_0) d \in -K_g(x_0) \text{ if and only if } d^T v_2(x_0) \leq d^T v_1(x_0). \tag{5.8}
\]

Now \( q(x_0) = (0, -1)^T \), that is, \( I_q(x_0) = \{1\} \) and thus 
\[
d^T v_{q_1}(x_0) \leq 0 \text{ for } i \in I_q(x_0)
\]
if and only if 
\[
d^T v_{q_1}(x_0) = d^T (-2v_{g_1}(x_0) + 2v_{g_3}(x_0)) < 0. \tag{5.9}
\]

Comparing (5.8) and (5.9) we see, analogously to (5.7), 
\[
g'(x_0) d \in -K_g(x_0) \text{ if and only if } d^T v_{q_i}(x_0) \leq 0
\]
for \( i \in I_q(x_0) \).

Later on we will need the following statement on the interior of \( K_a \) and \( K_{a,b} \). Recall that by our general assumption \( \text{int } K \not= \emptyset \).

**Lemma 5.4:** Let \( a \in -K \) and \( b \in -K_a \). Then 
\[
\text{int } K_a = \text{int } K + \{ \lambda a : \lambda \geq 0 \}
\]
\[
\text{int } K_{a,b} = \text{int } K + \{ \lambda a + \mu b : \lambda \geq 0, \mu \geq 0 \}. \tag{5.10}
\]

**Proof:** By (5.5), cone \( (K+a) = K + \{ \lambda a : \lambda \geq 0 \} \). In ZOWE [28, Th. 22] it is shown that 
\[
\text{int } (K+\{ \lambda a : \lambda \geq 0 \}) = \text{int } K + \{ \lambda a : \lambda \geq 0 \} .
\]
(5.10) follows, since \( \text{int } \text{cl } A = \text{int } A \) for a convex set \( A \) with \( \text{int } A \not= \emptyset \) (see e.g. SCHAEFER [26 p. 38]). (5.11) follows using (5.10). \( \text{qed} \)
6. Computation of the sets of directions of decrease and feasibility

We start with an analytical description of \( D_f^<(x) \) and \( D_g^<(x) \) in terms of the (first order) directional derivative

\[
f'(x;d) := \lim_{t \to 0^+} \frac{f(x+t\delta) - f(x)}{t}.
\]

Proposition 6.1: Let \( x \in X \) be given. Suppose \( f'(x;d) \) and \( g'(x;d) \) exist for all \( d \in X \). Then

(i) \( D_f^<(x) = \{ d \in X : f'(x;d) \leq 0 \} \).

(ii) If \( g(x) \in -K \):

\[
D_g^<(x) = \{ d \in X : g'(x;d) \in -K g(x) \}.
\]

Proof: We only prove (ii). The (more easy) proof of (i) can be copied from this.

\( \Rightarrow \): Suppose \( g'(x;d) \in -K g(x) \). Let \( v \in \text{int } K \) be given. By Proposition (5.1) (iii) and (5.5) there is some \( \bar{v} \in \frac{1}{2}[-v,v] \), \( k \in K \) and \( \lambda > 0 \) such that

\[
g'(x;d) + \bar{v} = -k - \lambda g(x).
\]

Now put

\[
r(t) := t^{-1}(g(x+t\delta) - g(x)) - g'(x;d), \quad t > 0.
\]

Then \( r(t) \to 0 \) for \( t \to 0^+ \) and, once more, by Proposition 5.1 (iii) we can choose \( T_1 > 0 \) so small such that

\[
x(t) := \frac{1}{2}[-v,v] \quad \text{for } 0 < t < T_1.
\]

We get for \( 0 < t \leq T := \min \{ T_1, \lambda^{-1} \} \) (\( \lambda^{-1} := \infty \) if \( \lambda = 0 \)):

\[
t^{-1}g(x+t\delta) = t^{-1}g(x) + g'(x;d) + r(t)
\]

\[
\in \left( t^{-1}-\lambda \right)g(x) - k + \frac{1}{2}v - K + \frac{1}{2}v - K
\]

\[
\subset K - k + v - K - K
\]

\[
\subset v - K.
\]
This proves $d \in D_g(x)$.

"$\subset$": Let $d \in D_g(x)$ be given. Then for every $v \in \text{int} K$ there is some real $T > 0$ such that

$$g(x+td) \in -K + tv \quad \text{for} \quad 0 < t \leq T.$$ 

Hence

$$g(x+td) - g(x) \in -(K+g(x)) + tv$$

$$\subset -Kg(x) + tv$$

and, since $Kg(x)$ is closed,

$$g'(x;d) = \lim_{t \to 0^+} \frac{1}{t} (g(x+td)-g(x)) \in -Kg(x) + v.$$ 

This last relation holds for every $v \in \text{int} K$. Using once more that $Kg(x)$ is closed we get $g'(x;d) \subset -Kg(x)$ which was to be shown.

qed

Now let $x,d \in X$ be given and assume $f'(x;d)$ exists. Then we define a second order directional derivative by (if the limit exists)

$$f''(x,d;z) = \lim_{t \to 0^+} \lim_{t' \to 0^+} \frac{1}{t} \left( f(x+td+t'z) - f(x+t'z) \right) - f'(x;d)$$

for $z \in X$. In the same way we introduce $g''(x,d;z)$. If $X$ and $U$ are normed spaces then we say that $f$ satisifies a \textit{local Lipschitz condition} at $x$ if there exist real $\varepsilon > 0$ and $\alpha > 0$ such that

$$\|f(x_1) - f(x_2)\| \leq \alpha \|x_1 - x_2\| \quad \text{for all} \quad \|x_i - x\| \leq \varepsilon, \quad i=1,2.$$ 

\textbf{Proposition 6.2} Suppose $X$, $U$ and $V$ are normed spaces. Let $x,d \in X$ be given and suppose $f''(x,d;z)$ and $g''(x,d;z)$ exist for all $z \in X$. Suppose furthermore that $f$ and $g$ satisfy a local Lipschitz condition at $x$. Then

(i) If $f'(x;d) \neq 0$:

$$Q_f(x;d) = \{ z \in X : f''(x,d;z) \in -\text{int} C_{f'}(x;d) \}$$
(ii) If \( g(x) \in -K \) and \( g'(x; d) \in -\text{cone} (K + g(x)) \) \((-K \) whenever \( \text{cone} (K + g(x)) \) is closed):
\[
Q_g(x, d) = \{ z \in X : g''(x, d; z) \in -\text{int} \, K_g(x), g'(x) d \}.
\]

(iii) If \( g(x) \in -K \) and \( g'(x; d) \in -K_g(x) \):
\[
Q_g(x, d) = \{ z \in X : g''(x, d; z) \in -\text{int} \, K_g(x), g'(x) d \}.
\]

**Proof:** We only prove (ii) and (iii). The easier proof of (i) can be copied from this. We start with

(iii): Suppose \( x \) and \( d \) are such that \( g(x) \in -K \) and \( g'(x; d) \in -K_g(x) \). For simplicity let us write \( K_1 \) for \( K_g(x) \) and \( g'(x) d \). Let some \( z \in Q_g(x, d) \) be given and put
\[
x_t := x + td + t^2 z \quad \text{for} \quad t \in \mathbb{R}.
\]
By definition there exists some \( v \in \text{int} \, K \) and \( T > 0 \) such that
\[
t^{-2} g(x_t) \in -K - v \quad \text{for} \quad 0 < t < T.
\]
It follows for \( 0 < t < T \):
\[
t^{-2} (g(x_t) - g(x)) - t^{-1} g'(x; d) - v = -(K + t^{-2} g(x) + t^{-1} g'(x; d)) - v \subseteq -K_1 - v.
\]
Since, by definition, \( K_1 \) is closed we get for \( t \to 0^+ \)
\[
g''(x, d; z) \in -K_1 - v.
\]
Now \( v \in \text{int} \, K \subseteq \text{int} \, K_1 \) and thus (see Proposition 5.1 (ii)) \( v + K_1 \subseteq \text{int} \, K_1 \). This proves (iii).

(ii): Suppose \( g(x) \in -K \) and \( g'(x; d) \in -\text{cone}(K + g(x)) \). Only the inclusion \( \subseteq \) has to be shown. Let \( z \in X \) be such that
\[
g''(x, d; z) \in -\text{int} \, K_1.
\]
Then by Lemma 5.4
\[
g''(x, d; z) = -\lambda g(x) - \mu g'(x; d)
\]
with some \( \lambda, \mu \geq 0 \). We put again \( x_t := x + td + t^2 z \) and choose \( T_1 > 0 \) small enough such that
for $0 < t \leq T$. Rearranging and using that, by assumption,
\[ g'(x; d) = -v_1 - \alpha g(x) \] with some $v_1 \in K$ and $\alpha > 0$ we get
\[ t^{-2} g(x_t) \in -2v + (t^{-2} - \alpha t^{-1} + \alpha u - \lambda) g(x) - (t^{-1} - u) v_1 - K \]
for $0 < t \leq T$. Now $v_1 \in K$ and $g(x) \in -K$. Hence with
\[ 0 < T_2 < T \] small enough
\[ t^{-2} g(x_t) \in -2v - K - K = -2v - K \] for $0 < t \leq T_2$.
that is
\[ g(x + td + t^2z) \in -K - 2t^2v, \quad 0 < t \leq T_2. \tag{6.1} \]

To end the proof choose a real $\beta > 0$ and $0 < T < T_2$ so small such that
\[ \|x + td + t^2z\| < \epsilon \] for $\|z - \overline{z}\| \leq \beta$ and $0 < t \leq T$.
Then by the assumed Lipschitz continuity of $g$ at $x$,
\[ \|g(x + td + t^2z) - g(x + td + t^2\overline{z})\| \leq t^2\alpha \beta \]
for $\|z - \overline{z}\| \leq \beta$ and $0 < t \leq T$. By Proposition 5.1 (iii) we can choose $\beta > 0$ so small such that the ball around zero with
radius $\alpha \beta$ is contained in $[-v, v]$. Together with (6.1) we get
for such $\beta$:
\[ g(x + td + t^2z) = g(x + td + t^2\overline{z}) + (g(x + td + t^2z) - g(x + td + t^2\overline{z})) \]
\[ \in -K - 2t^2v + t^2[-v, v] \]
\[ \in -K - t^2v \]
for $0 < t \leq T$ and $\|z - \overline{z}\| \leq \beta$. This shows $z \in Q_g(x, d)$.
\[ \text{Q.E.D.} \]
For the first order sets we get from (2.3) and the above result

**Corollary 6.3** Let $X$, $U$ and $V$ be normed spaces and suppose $f$ and $g$ satisfy a local Lipschitz condition at a given $x$. Suppose $f'(x;d)$ and $g'(x;d)$ exist for all $d \in X$. Then

(i) $D^<_f(x) = \{d \in X : f'(x;d) < 0\}.$

(ii) If $g(x) \in -X$:

$$D^<_g(x) = \{d \in X : g'(x;d) \in -\text{int} K_g(x)\}.$$ 

The directional derivatives used in the above computation exist in particular if $X$, $U$ and $V$ are normed and if the functions $f$ and $g$ are Fréchet-differentiable (F-differentiable). We write $f'(x)$, $g'(x)$ and $f''(x)$, $g''(x)$ for the first and second F-derivatives of $f$ and $g$, respectively, at a given point $x$. For fixed $x$, $f''(x)$ and $g''(x)$ are interpreted as bilinear forms on $X \times X$ with values in $U$ and $V$, respectively. It is easy to see that for F-differentiable $f$ and $g$ the directional derivatives $f'(x;d)$ and $g'(x;d)$ exist and

$$f'(x;d) = f'(x)d, \quad g'(x;d) = g''(x)d \quad \text{for all } d \in X. \quad (6.2)$$

Now suppose $f''(x)$ exist. Then by definition of the second derivative

$$f(x+td+t^2z) - f(x)$$

$$= f'(x)(td+t^2z) + \frac{1}{2}f''(x)(td+t^2z,td+t^2z) + o(||td+t^2z||^2)$$

$$= tf'(x)d + t^2f'(x)z + \frac{1}{2}t^2f''(x)(d,d) + o(t^2).$$

It follows that $f''(x;d,z)$ exist and

$$f''(x,d;z) = f'(x)z + \frac{1}{2}f''(x)(d,d) \quad \text{for all } d,z \in X. \quad (6.3)$$
Analogously,

\[ g'(x,d;z) = g'(x)z + \frac{1}{2} g''(x)(d,d) \text{ for all } d,z \in X. \]  

(6.4)

Note, furthermore, that twice F-differentiable functions (hence, continuously F-differentiable) are always locally Lipschitzian; this follows by an easy argument using the mean value theorem. (6.3) and (6.4) show that in the differentiable case the sets \( Q_f(x,d) \) and \( Q_g(x,d) \) in Proposition 6.2 (i), (ii) are convex.

7. Computation of the set of tangent directions

The proof of the Lemma below closely follows the proof of Theorem 2 in Ljusternik and Sobolev [20 chapter VIII section 10].

**Lemma 7.1:** Let \( X \) and \( W \) be Banach spaces, \( h : X \rightarrow W \) twice F-differentiable in a neighborhood of \( x \in X \), where \( h(x) = 0 \). Assume that \( h'(x) \) maps \( X \) onto \( W \). If \( d \in X \) and \( z \in X \), satisfy

\[ h'(x)d = 0 \]

(7.1)

\[ h'(x)z + \frac{1}{2} h''(x)(d,d) = 0, \]

(7.2)

then,

\[ z \in V_h(x,d). \]

**Proof:** Let \( T_0 \) be the set of \( d \)'s satisfying (7.1), and denote by \( X/T_0 \) the quotient space. We use the same symbol for the norm on \( X \) and \( W \), for the norm defining the quotient topology of \( X/T_0 \) and for the operator norm on the space of all
bounded linear operators from \( W \) into \( X/T_o \). The map \( h'(x) \) gives rise to a linear operator \( A : X/T_o \to W \) in the following sense: if \( T \in X/T_o \) then \( AT := h'(x_0)g \) where \( g \) is an arbitrary element in the equivalence class \( T \). By the fact that \( h'(x) \) is onto it follows that \( A \) has an inverse, \( A^{-1} \), which is a continuous linear operator (Banach's Homomorphism, see e.g. SCHAEFER [26]).

Now let \( y \in X \) be given; we construct a sequence \( \{T_n\} \) of elements of \( X/T_o \) and choose simultaneously a corresponding sequence of elements \( r_n \in T_n \) by the recurrence relation

\[
T_n := T_{n-1} - A^{-1}h(x+y+r_{n-1}),
\]

where \( r_0 = 0(\in T_o) \). By definition of the quotient topology we may assume that \( r_n \in T_n \) is chosen such that

\[
\|r_n - r_{n-1}\| \leq 2\|T_n - T_{n-1}\|.
\]

In the above mentioned proof [20] it is shown that, for arbitrary \( y \in X \):

\[
T_n - T_{n-1} = -A^{-1} \int_0^1 [h'(x+y+\alpha r) - h'(x)] \mathrm{d}\alpha \cdot (r_{n-1} - r_{n-2})
\]

where

\[
r_\alpha := \alpha r_{n-1} + (1-\alpha)r_{n-2}.
\]

By a Taylor expansion

\[
h(x+y) = h(x)^\gamma + h'(x)y + \frac{1}{2}h''(x)(y, y) + o(\|y\|^2).
\]

In particular, for

\[
y = td + t^2z,
\]

where \( d, z \) satisfy (7.1) and (7.2), it follows from (7.6) that

\[
h(x+y) = o(\|y\|^2).
\]

By (7.3), (7.4), (7.8)

\[
\|r_1\| \leq 2\|T_1\| \leq 2\|A^{-1}\| \|h(x+y)\| \leq 2\|A^{-1}\| \cdot o(\|y\|^2)
\]
and for $\|y\|$ small enough so that $o(\|y\|^2) \leq \frac{1}{4\|A^{-1}\|} \|y\|^2$:

$$\|r_1\| \leq \frac{1}{2} \|y\|^2.$$  \hspace{1cm} (7.9)

We proceed now to show by induction that

$$\|r_n\| \leq \|y\|^2$$ for $n=0,1,2,\ldots$ \hspace{1cm} (7.10)

This is obvious for $n = 0$ ($r_0 = 0$) and for $n = 1$ (by (7.9)). Assume then that

$$\|\tilde{f}_i\| \leq \|y\|^2, \quad i = 0,1,2,\ldots, n-1.$$ By the continuity of $h'$ at $x$, there exists for every $\varepsilon > 0$, a real $\rho > 0$ such that

$$\|h'(z) - h'(x)\| \leq \varepsilon \quad \text{for} \quad \|z-x\| \leq 2\rho.$$ Therefore for $\|y\|^2 \leq \rho$

$$\|h'(x + y + r_0) - h'(x)\| \leq \varepsilon$$

and so, from (7.5)

$$\|T_n - T_{n-1}\| \leq \varepsilon \|A^{-1}\| \|r_{n-1} - r_{n-2}\|.$$ The latter and (7.4) imply

$$\|r_n - r_{n-1}\| \leq 2\varepsilon \|A^{-1}\| \|r_{n-1} - r_{n-2}\|$$

and for $\varepsilon$ small enough so that $2\|A^{-1}\|\varepsilon \leq \frac{1}{2}$:

$$\|r_n - r_{n-1}\| \leq \frac{1}{2} \|r_{n-1} - r_{n-2}\|$$

$$\leq \ldots \leq \frac{1}{2^{n-1}} \|r_1 - r_0\| \leq \frac{1}{2^{n-1}} \|r_1\|.$$ \hspace{1cm} (7.11)

Moreover,

$$\|r_n\| = \|r_1 + (r_2-r_1) + \cdots + (r_n - r_{n-1})\|$$

$$\leq \|r_1\| + \|r_2-r_1\| + \cdots + \|r_n - r_{n-1}\|$$

$$\leq \|r_1\| (1 + \frac{1}{2} + \cdots + \frac{1}{2^{n-1}})$$

$$\leq 2\|r_1\|.$$ \hspace{1cm} (7.12)

So by (7.9) and the fact that $\|y\|^2 \leq r$:

$$\|r_n\| \leq \|y\|^2.$$
Together with (7.11) this shows that \( r_n \) converges to some \( r \in X \) and
\[
\|r\| \leq \|y\|^2 \text{ where } y = td + \frac{1}{2}t^2d, \quad t > 0 \text{ small enough.}
\]

Using (7.12) and the inequality following (7.8) we get
\[
\|r\| \leq 2\|r_1\| \leq 4\|A^{-1}\| \circ (\|y\|^2)
\]

which shows that for \( y = td + t^2z \)
\[
r = r(t) \sim o(t^2). \tag{7.13}
\]

If \( T \) is the limit of the sequence \( \{T_n\} \) then, letting \( n \to \infty \), relation (7.3) yields
\[
T = T - A^{-1}h(x+y+r)
\]
or, in other words,
\[
h(x+td+t^2z+r(t)) = 0, \quad t > 0 \text{ small enough.}
\]

This together with (7.13) shows that \( z \in V_h(x_0,d) \).
\[
\text{qed}
\]

**Proposition 7.2:** Let \( h \) be a mapping from the Banach space \( X \) to the Banach space \( W \) which is twice \( F \)-differentiable in a neighborhood of a given point \( x \) with \( h(x) = 0 \). Assume that \( h'(x) \) maps \( X \) onto \( W \). Then, \( h \) is regular at \( x \) and for every \( d \in T_h(x) \)
\[
V_h(x,d) = \{z \in X : h'(x)z + \frac{1}{2}h''(x)(d,d) = 0\}. \tag{7.14}
\]

In particular,
\[
T_h(x) = V_h(x,0) = \{z \in X : h'(x)z = 0\}. \tag{7.15}
\]

**Proof:** We first prove that (7.14) holds for every \( d \) such that
\[
h'(x)d = 0. \tag{7.16}
\]
The inclusion $\subset$ in (7.14) follows from Lemma 7.1. For the reverse inclusion assume that $z \in \mathcal{V}_h(x,d)$, i.e., for some $T > 0$ and a curve $r(t) \sim o(t^2)$

$$h(x+td+t^2z+r(t)) = 0 \quad \text{for} \quad 0 < t \leq T.$$  \hspace{1cm} (7.17)

A Taylor expansion gives (using $h(x) = 0$ and $h'(x)d = 0$)

$$h(x+td+t^2z+r(t))$$

$$= h'(x)(t^2z+r(t)) + \frac{1}{2}h''(x)(td+t^2z+x(t),td+t^2z+r(t))$$

$$+ o(t^2).$$

Since $r(t) \sim o(t^2)$ the latter implies

$$t^2[h'(x)z + \frac{1}{2}h''(x)(d,d)] + o(t^2) = 0, \quad 0 < t \leq T.$$  \hspace{1cm} (7.18)

Dividing by $t^2$ and letting $t \to 0$ we get

$$h'(x)z + \frac{1}{2}h''(x)(d,d) = 0.$$  \hspace{1cm} (7.19)

This completes the proof of the inclusion $\subset$ in (7.14).

Now, since $d = 0$ satisfies (7.16), the result (7.15) follows immediately from what we know already. Moreover, by (7.15) the set of $d$'s satisfying (7.15) coincides with $T_h(x)$. Thus (7.14) holds in fact for every $d \in T_h(x)$.  \hspace{1cm} Qed
8. Necessary conditions under differentiability

For the differentiable version of Theorem 2.1 we have to compute the support functionals appearing in Theorem 2.1. This is done in Proposition 8.1.

Proposition 8.1: Let $X$, $U$, $V$ and $W$ be normed spaces and let $f$, $g$ and $h$ be twice $F$-differentiable. Let $x$, $d \in X$ be given.

(i) Suppose $f'(x)d \leq 0$ and $Q_f(x,d) \neq \emptyset$. Then for $l_f \in \Lambda(Q_f(x,d))$ there is some $u^* \in C^+$ such that

$$l_f = u^* f'(x), \quad u^* f''(x) d = 0$$

$$\delta*(l_f | Q_f(x,d)) = - \frac{1}{2} u^* f''(x)(d,d).$$

(ii) Suppose $g'(x) \in -K$, $g'(x)d \notin -\text{cohe}(K+g(x))$ and $Q_g(x,d) \neq \emptyset$. Then for $l_g \in \Lambda(Q_g(x,d))$ there is some $v^* \in K^+$ such that

$$l_g = v^* g'(x), \quad v^* g''(x) d = 0, \quad v^* g(x) = 0,$$

$$\delta*(l_g | Q_g(x,d)) = - \frac{1}{2} v^* g''(x)(d,d).$$

(iii) Suppose $X$ and $W$ are Banach spaces, $h'(x)$ is onto, $h(x) = 0$ and $h'(x)d = 0$. Then for $l_h \in V_h(x,d)$ there is some $w^* \in W^*$ such that

$$l_h = w^* h'(x),$$

$$\delta*(l_h | V_h(x,d)) = - \frac{1}{2} w^* h''(x)(d,d).$$

Proof (i): The proof can be copied from (ii).

(ii): Let $l_g \in \Lambda(Q_g(x,d))$ be given. We put in Lemma 3.5

$$A := g'(x) \quad \text{and} \quad S := \frac{1}{2} g''(x)(d,d) - \text{int} \, K_g(x), g''(x)d,$$

By Proposition 6.2 (ii) and (6.4), $l_g \in \Lambda(A^{-1}S)$. Now condition (i) in Lemma 3.5 is satisfied since, by assumption,
Q_{g}(x,d) \neq \emptyset. We get from 3.5 that with a suitable \( v^* \in \Lambda(S) \),

\[ l_g = v^* \cdot g'(x) \quad \text{and} \quad \delta^*(v^*|S) = \delta^*(l_g|Q_g(x,d)). \]

Taking into account that \( \int K_g(x), g'(x)d = \int K^+(\lambda g(x)+\mu g'(x)d: \lambda, \mu \geq 0 \) (see Lemma 5.4) we get for \( v^* \in \Lambda(S) \):

\[ v^* \in K^+, v^*g(x) = 0 \quad \text{and} \quad v^*g'(x)d = 0 \]

and, finally,

\[ \delta^*(v^*|S) = -\frac{1}{2} v^* g''(x)(d,d). \]

(iii) The assertion follows in a similar way from Proposition 7.2 and Lemma 3.5 if one puts

\[ A := h'(x) \quad \text{and} \quad S := -\frac{1}{2} h''(x)(d,d). \]

\[ \text{qed} \]

We get

**Theorem 6.2:** Let \( x_0 \) be an optimal point for problem (P). Let \( X \) and \( W \) be Banach spaces and \( U \) and \( V \) normed. Suppose \( f, g \) and \( h \) are twice \( F \)-differentiable and the range \( R(h'(x_0)) \) of \( h'(x_0) \) is closed. Then for every \( d \) satisfying

\[ f'(x_0)d \leq 0, \quad g'(x_0)d \in -\text{cone}(K+g(x_0)), \quad h'(x_0)d = 0 \quad (8.1) \]

there correspond continuous linear functionals \( u^* \in C^+, v^* \in K^+ \) and \( w^* \in W^*, \) not all zero, such that

\[ u^* f'(x_0)d = 0, \quad v^* g(x_0) = 0, \quad v^* g'(x_0)d = 0 \quad (8.2) \]

\[ u^* f''(x_0) + v^* g''(x_0) + w^* h''(x_0)(d,d) \geq 0. \quad (8.4) \]
Proof: Suppose first, \( h'(x_0) \) is not onto. Then \( R(h'(x_0)) \) and an arbitrary element in \( W \) not belonging to \( R(h'(x_0)) \) can be separated, that is, there is some \( \overline{w}^* \in W^*, \overline{w}^* \not= 0 \), such that

\[
\overline{w}^* \cdot h'(x_0)x > 0 \quad \text{for all } x \in X.
\]

It follows \( \overline{w}^* \cdot h'(x_0)x = 0 \). (8.2) - (8.4) hold with \( u^* = 0 \), \( v^* = 0 \) and \( w^* = \overline{w}^* \) or \( w^* = -\overline{w}^* \). Hence we may assume throughout the following that \( h'(x_0) \) is onto.

Now, suppose (8.1) holds with a given \( d \). By, Proposition 6.1 and Proposition 7.1, \( d \) satisfies (2.9) in Theorem 2.1. Proposition 7.2 and 6.2 together with (6.3) and (6.4) imply that \( Q_f(x_0,d) \), \( Q_g(x_0,d) \) and \( V_h(x_0,d) \) are convex; moreover, \( V_h(x_0,d) \) is non-empty. If also \( Q_f(x_0,d) \) and \( Q_g(x_0,d) \) are non-empty then \( f, g \) and \( h \) are \( d \)-regular and Theorem 2.1 applies. The assertion follows from 2.1 and Proposition 8.1.

Now suppose \( Q_g(x_0,d) = \emptyset \) (if \( Q_f(x_0,d) \not= \emptyset \), then a similar argument applies). Then the non-empty convex sets

\[
\{ g'(x_0)x + \frac{1}{2} g''(x_0)(d,d) : z \in X \} \text{ and } \text{int } K_g(x_0), g'_*(x_0) d
\]

be separated and thus there exists some \( v^* \not= 0 \) in \( V^* \) such that for all \( z \in X \), \( v \in K \) and \( \lambda, \mu \geq 0 \):

\[
v^*(g'(x_0)x + \frac{1}{2} g''(x_0)(d,d)) \geq v^*(-k-\lambda g(x_0)-\mu g''(x_0)d).
\]

It follows

\[
v^* \in K^+, v^*g(x_0) = 0, v^*g'(x_0)d = 0
\]

and, finally, \( v^*g''(x_0)(d,d) \geq 0 \). (8.3), (8.4) follow if we put \( u^* = 0 \) and \( w^* = 0 \).

\( \text{qed} \)
The constraint qualification \( CQ(\hat{d}) \) introduced in section 4 Remark 3 becomes:

There exists \( z \in X \) such that
\[
CQ(\hat{d}): \quad g'(x_0)z + g''(x_0)(\hat{d}, \hat{d}) \in -\text{int} K_{g(x_0)}, g'(x_0)\hat{d} \\
h'(x_0)z + h''(x_0)(\hat{d}, \hat{d}) = 0.
\]

If \( CQ(\hat{d}) \) holds for \( \hat{d} \) satisfying (8.1), then
\[
Q_{g}(x_0, \hat{d}) \cap V_{h}(x_0, \hat{d}) \neq \emptyset \quad \text{(Proposition 6.2 (ii) and 7.2) and thus \( l_\hat{f} \neq 0 \) in Theorem 2.1 (see Remark 3 in Section 4)};
\]
this shows \( u^*_\hat{f} \neq 0 \).

Corollary 8.3: If for the vector \( \hat{d} \) in (8.1) condition \( CQ(\hat{d}) \) is satisfied, then, the corresponding \( u^* \) in (8.2) - (8.4) is not the zero-functional.

9. The mathematical programming problem

We specialize Theorem 2.1 to problem (MP). To this denote by
\[
D_{g_1}(x) \quad \text{the set of quasifeasible directions of the component} \quad g_1 \quad \text{of} \quad g \quad \text{(replace in the definition} \quad K \quad \text{by} \quad \mathbb{R}_+); \quad \text{the same for} \quad Q_{g_1}(x, \hat{d}) \quad \text{and} \quad D_{g_1}(x).
\]
We call \( g_1 \) regular at \( x \) if \( Q_{g_1}(x, \hat{d}) \) is non-empty and convex for all \( \hat{d} \in D_{g_1}(x) \). Further recall the definition of \( I(x) \),
\[
I(x) = \{ i \in \{1, \ldots, n\} : g_i(x) = 0 \},
\]
and put for \( d \in \mathcal{D}_g(x) \) (compare (5.2))
\[
J(x, d) = \{ i \in I(x) : d \in \mathcal{D}_g_i(x) \}.
\]

**Theorem 9.1** Let \( x_0 \) be an optimal point for problem (MP) and suppose \( f, h \) and \( g_i, i \in I(x_0), \) are regular at \( x_0. \) Then for every
\[
d \in \mathcal{D}_f(x_0) \cap \bigcap_{i \in J(x_0)} \mathcal{D}_g_i(x_0) \cap T_h(x_0)
\]
(9.1)
there correspond continuous linear functionals,
\[
l_f \in \Lambda(Q_f(x_0, d)), l_h \in \Lambda(V_h(x_0, d))
\]
(9.2)
\[l_{g_i} \in \Lambda(Q_{g_i}(x_0, d)) \text{ for } i \in J(x_0, d)
\]
not all zero, satisfying
\[
l_f + \sum_{i \in J(x_0, d)} l_{g_i} + l_h = 0
\]
(9.3)
and
\[
\delta^*(l_f|Q_f(x_0, d)) + \sum_{i \in J(x_0, d)} \delta^*(l_{g_i}|Q_{g_i}(x_0, d))
\]
(9.4)
\[+ \delta^*(l_h|V_h(x_0, d)) \leq 0.
\]

**Proof** Let \( d \) be given and suppose (9.1) holds. A look at the definition of \( D_g(x_0) \) shows
\[
D_g(x_0) = \bigcap_{i \in I(x_0)} D_{g_i}(x_0).
\]
Hence, (2.9) is satisfied by \( d. \) Furthermore, one has
\[
Q_g(x_0, d) = \bigcap_{i \in I(x_0)} Q_{g_i}(x_0, d)
\]
and since, by assumption, \( Q_{g_i}(x_0, d) \) is convex and non-empty for \( i \in I(x_0) \) the set \( Q_g(x_0, d) \) is also convex. If \( Q_g(x_0, d) \)
is non-empty then \( f, g \) and \( h \) are \( d \)-regular at \( x_0 \) and
Theorem 2.1 applies. From Lemma 3.2 we get that $l_g$ in 2.1 can be written as a sum:

$$l_g = \sum_{i \in I(x_0)} l_{g_i}$$

where $l_{g_i} \in \Lambda(Q_{g_i}(x_0, d))$

and that

$$\delta^*(l_g | Q_g(x_0, d)) = \sum_{i \in I(x_0)} \delta^*(l_{g_i} | Q_{g_i}(x_0, d)).$$

(2.4) implies $l_{g_i} = 0$ for $i \in I(x_0) \setminus \emptyset(x_0, d)$. (9.2)-(9.4) follows.

Now suppose $Q_g(x_0, d) = \emptyset$. Then the assertion follows from Lemma 3.3 if we put $l_f = 0$ and $l_h = 0$.

qed

The above second order condition contains the famous first order condition by Dubobitskii-Milyutin [10, Theorem 6.1] as a special case. To see this, put $d = 0$ in Theorem 9.1. Obviously, (9.1) is satisfied. The second order sets in (9.2) become (see (2.3), (2.6) and (2.8))

$$\Lambda(Q_f(x_0, 0),) = -D_f^<(x_0)^+, \Lambda(V_h(x_0, 0)) = -T_h(x_0)^+$$

$$\Lambda(Q_{g_i}(x_0, 0),) = -D_{g_i}^<(x_0)^+$$

for $i \in I(x_0)$.

(9.2) reduces to the trivial inequality $0 \leq 0$. We get

**Corollary 9.2** Let $x_0$ be an optimal point for problem (MP) and suppose $D_f^<(x_0)$, $T_h(x_0)$ and $D_{g_i}^<$ for $i \in I(x_0)$ are non-empty and convex. Then there are $l_f \in D_f^<(x_0)^+$, $l_h \in T_h(x_0)^+$ and $l_{g_i} \in D_{g_i}^<$ for $i \in I(x_0)$, not all zero, such that

$$l_f + \sum_{i \in I(x_0)} l_{g_i} = 0.$$
The following differentiable version of Theorem 9.1 is easily proved by specializing Theorem 8.2.

**Theorem 9.3** Let $x_0$ be an optimal point for problem (MP). Let $X$ and $W$ be Banach spaces, suppose $f$, $h$ and $g_i$ for $i=1,2,...,n$ are twice $F$-differentiable and the range of $h'(x_0)$ is closed. Then for every $d$ satisfying

$$f'(x_0)d \leq 0, \quad g_i'(x_0)d \leq 0 \quad \text{for} \quad i \in I(x_0), \quad h'(x_0)d = 0$$

(9.5)

there correspond real numbers $y_0 \geq 0$ and $y_i \geq 0$ for $i \in J(x_0,d)$ and a functional $w^* \in W^*$, not all zero, such that

$$y_0f'(x_0) + \sum_{i \in J(x_0,d)} y_i g_i'(x_0) + w^*h'(x_0) = 0$$

(9.6)

and

$$y_0f''(x_0) + \sum_{i \in J(x_0,d)} y_i g_i''(x_0) + w^*h''(x_0) = 0.$$ 

(9.7)

**Remark 1** For dim $W < \infty$ the range of $h'(x_0)$ is always closed and this assumption can be omitted. Theorem 9.3 reduces to a result by BEN-TAL [5]; see also COX (in HESTENES [13, Chapter 6 Theorem 10.4]) where a similar statement is proved for finite dimensional $X$ and for those $d$'s only satisfying (9.5) with equality.

**Remark 2** Suppose under the assumptions of Theorem 9.3 that (9.5) holds for $d$. Then the corresponding $y_0$ can be chosen as 1 if the following constraint qualification is met (see Corollary 8.3 and Proposition 5.3):

There exists $z \in X$ for which

$$g_i'(x_0)z + g_i''(x_0)(d,d) < 0 \quad \text{for} \quad i \in J(x_0,d)$$

$$h'(x_0)z + h''(x_0)(d,d) = 0.$$
Remark 3: For $d = 0$ and $W = \mathbb{R}^m$, Theorem 9.3 reduces to the FRITZ-JOHN condition. The condition guaranteeing $y_0 = 1$ in (9.6) becomes the MANGASARIAN-FROMOVITZ constraint qualification (see e.g. GAUVIN-TOLLE [9]):

There exists $z \in X$ such that:

$\nabla g^i(x_0)z < 0$ for $i \in I(x_0)$

$\nabla h^j(x_0)z = 0$ for $j = 1, 2, \ldots, m$.

We point out once more that the multipliers in Theorem 9.3 depend on $d$ (compare Remark 2 in Section 4). The question arises under what conditions one can state a necessary condition where the same multipliers can be taken for all $d$.

One such case will be when the multipliers in the EULER-LAGRANGE equation (9.6) are determined uniquely by (9.6). To state such a condition denote by $g^i_1(x_0)(x_0)$, the linear map of $X$ into $\mathbb{R}^{\text{card } I(x_0)}$

$g^i_1(x_0)(x_0)x = (g^i_1(x_0)x)^T i \in I(x_0) \quad \text{for } x \in X.$

We say that the constraint qualification $\text{CQ1}$ holds if

$\text{CQ1: } \{x = (g^i_1(x_0)(x_0)x, h^j(x_0)x) \text{ maps } X \text{ onto } \mathbb{R}^{\text{card } I(x_0)} \times W.$

For $\dim X < \infty$ and $h = (h_1, \ldots, h_m)$ (i.e., $W = \mathbb{R}^m$), $\text{CQ1}$ is equivalent to the well-known condition (see e.g. FIACCO-MCCORMICK [8] or LUENBERGER [19])

The gradients $\nabla g^i(x_0)$ and $\nabla h^j(x_0)$, $i \in I(x_0)$ and $j = 1, \ldots, m$, are linearly independent.

Another type of regularity assumption is the following
For every $d$ satisfying (9.5) there is $z \in X$ such that:

\[ g_i'(x_0)z + g_i''((x_0)(d,d) = 0 \quad \text{for } i \in I(x_0) \]
\[ h'(x_0)z + h''((x_0)(d,d) = 0. \]

For $\dim X < \infty$ and $\dim W < \infty$, CQ2 is implied by the so-called second order constraint qualification given by FIACCO-McCOR - MICK [8, p. 25].

We get the following result which for $\dim X < \infty$ and $\dim W < \infty$ can be found e.g. in FIACCO-McCORMICK [8 p. 25], LUENBERGER [19] or HESTENES [13 p. 27].

**Corollary 9.4:** Consider problem (MP) with twice differentiable functions $f$, $g_i$ and $h$. Let $X$ and $W$ be Banach spaces and the range of $h'(x_o)$ be closed. Let $x_0$ be optimal and suppose that either of the following conditions holds

(i) CQ1,

(ii) CQ2 together with CQ(d) for all $d$ satisfying (9.5).

Then there are multipliers $y_i$ for $i \in I(x_0)$ and $w^* \in W^*$ such that

\[ f'(x_0) + \sum_{i \in I(x_0)} y_i g_i'(x_0) + w^* h^r(x_0) = 0. \quad (9.8) \]

and

\[ (f''(x_0) + \sum_{i \in I(x_0)} y_i g_i''(x_0) + w^* h''((x_0))(d,d) \geq 0. \quad (9.9) \]

for all $d$ satisfying

\[ g_i'(x_0)d = 0 \quad \text{for } i \in I(x_0) \text{ and } y_i > 0 \]
\[ g_i'(x_0)d \leq 0 \quad \text{for } i \in I(x_0) \text{ and } y_i = 0 \quad (9.10) \]
\[ h'(x_0)d = 0. \]

**Proof:** (i) If CQ1 holds, then $y_0$ in (9.6) can be norma-
lized to 1 for all $d$ satisfying (9.5) ($y_0 = 0$ leads to a contradiction to the onto condition and the fact that not all multipliers are zero in 9.3). Once $y_0$ is put equal to 1, CQ1 also implies that the $y_i$ and $w^*$ in (9.6) are uniquely determined. Hence (9.6) and (9.7) or (9.8) and (9.9), respectively, hold with fixed multipliers for all $d$ satisfying (9.5). Using (9.8) it is easily seen that the $d$'s for which (9.10) holds do also satisfy (9.5).

(ii) We know from Remark 2 above, that we can choose $y_0 = 1$ in (9.6), (9.7) for all $d$ satisfying (9.5). Now suppose (9.6), (9.7) hold for a fixed $d$ with $y_i$, $w^*$ and for some other $\bar{d}$ with $\bar{y}_i$, $\bar{w}^*$; further let $\bar{z}$ be the element which corresponds to $\bar{d}$ via CQ2. All we have to show is that (9.7) also holds for $\bar{d}$ with $\bar{y}_i$ and $\bar{w}^*$. But this follows by some easy calculation:

\[
\begin{align*}
(f''(x_0) + \sum_{i \in J(x_0,d)} y_i g''_i(x_0) + w^* h''(x_0))(\bar{d}, \bar{d}) \\
= f''(x_0)(\bar{d}, \bar{d}) - \sum_{i \in J(x_0,d)} y_i g''_i(x_0)\bar{z} - w^* h''(x_0)\bar{z} \\
= f''(x_0)(\bar{d}, \bar{d}) - \sum_{i \in J(x_0,d)} \bar{y}_i g''_i(x_0)\bar{z} - \bar{w}^* h''(x_0)\bar{z} \\
= (f''(x_0) + \sum_{i \in J(x_0,d)} \bar{y}_i g''_i(x_0) + \bar{w}^* h''(x_0))(\bar{d}, \bar{d})
\end{align*}
\]

$= 0$.

qed
10. An example

We study an example for Theorem 2.1 and 9.1. Consider the non-differentiable problem (compare FIACCO-McCORMICK [8, p.24])

\[ \text{minimize } f(x) = (x_1 - 1)^2 + x_2^2 \]
\[ \text{subject to } g(x) = |x_1| - \frac{1}{k} x_2^2 \leq 0 \]

where \( k \) is a positive parameter. For what values of \( k \), \( x_0 := (0,0)^T \) can be optimal?

We calculate (see Proposition 6.1 (i), 6.2 (i) and e.g. (6.3))

\[ \nabla f(x_0) = \begin{pmatrix} -2 \\ 0 \end{pmatrix}, \quad \nabla^2 f(x_0) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \]
\[ D_f(x_0) = \{ d : d^T \nabla f(x_0) \leq 0 \} = \{ (0,d_2) : d_2 \in \mathbb{R} \} \]
\[ Q_f(x_0,d) = \{ z : -2z_1 + d_2^2 < 0 \} \text{ for } d \in D_f(x_0). \]

The last relation shows that \( f \) is regular at \( x_0 \). By Proposition 8.1 (i)

\[ \delta_f(y|Q_f(x_0,d)) = \begin{cases} \frac{1}{2} \lambda_0 d^T \nabla^2 f(x_0) d & \text{if } y = \lambda_0 \nabla f(x_0) \text{ for some } \lambda_0 > 0 \\ \infty & \text{otherwise} \end{cases} \]

Further, by Proposition 6.1 (ii)

\[ D_g(x_0) = \{ d : g'(x_0; d) \leq 0 \} \]
\[ = \{ d : \lim_{t \to 0^+} t^{-1} [t|d_1| - \frac{1}{k} t^2 d_2^2] \leq 0 \} \]
\[ = \{ (0,d_2) : d_2 \in \mathbb{R} \} \]

and, by Corollary 6.3 (ii),

\[ D_g(x_0) = \{ d : g'(x_0; d) < 0 \} = \emptyset \]
so $D_g(x_0) = D_{g_0}(x_0)$.

Now for $d \in D_g(x_0)$, it follows from Proposition 6.2(ii) that $Q_g(x_0,d) = \{z : g''(x_0,d;z) < 0\}$ which (after some arithmetic) becomes

$Q_g(x_0,d) = \{(z_1^2) : |z_1| < \frac{1}{k} d_2^2, z_2 \in \mathbb{R}\}$.

Hence $g$ is $d$-regular at $x_0$ if $d_2 \neq 0$. Finally, by a direct computation,

$$
\delta^*(y|Q_g(x_0,d)) = \begin{cases} 
\frac{1}{k} |y_1| d_2^2 & \text{if } y = (\frac{y_1}{0})^T, y_1 \in \mathbb{R} \\
\infty & \text{otherwise}
\end{cases}
$$

The EULER-LAGRANGE equation (2.11) is then

$$
\lambda_0(0) + (\frac{y_1}{0}) = (0), \quad \lambda_0 > 0, y_1 \in \mathbb{R} \quad (10.1)
$$

and the LEGENDRE inequality (2.12) is

$$
-\lambda_0 d_2^2 + \frac{1}{k} |y_1| d_2^2 \leq 0, \quad d_2 \in \mathbb{R} \quad (10.2)
$$

From (10.1) it follows that $|y_1| = 2\lambda_0$, which, when substituted in (10.2), yields

$$
\lambda_0 d_2^2 (1 - \frac{2}{k}) \geq 0.
$$

Hence $k \geq 2$, which answers the question.

In fact, it can be shown that $k > 2$ is also sufficient for the optimality of $x_0$. Note that the first order condition Corollary 4.1, does not provide any information since $D_{g_0}(x_0) = Q_g(x_0,0) = \emptyset$ (for any $k$) and therefore the Euler-LAGRANGE equation 4.3 holds trivially.
11. **Sufficient conditions**

We start with the differentiable case. Let us call a point $x_0$ of the feasible set $F$ **strict** optimal for $(P)$ if there is a neighborhood $N$ of $x_0$ such that for $x \in N \cap F$ 

$$f(x) \leq f(x_0) \implies x = x_0.$$ 

For finite-dimensional $X$ we can prove a counterpart of Theorem 8.2. For $U = \mathbb{R}^n$, $V = \mathbb{R}^m$ and $W = \mathbb{R}^m$ this reduces to a result by BEN-TAL [2]; see also HESTENES [13, chapter 4, Theorem 7.4].

**Theorem 11.1:** Consider problem $(P)$ with finite-dimensional $X$, closed convex cone $C$ and twice $F$-differentiable functions $f$, $g$ and $h$. Let $x_0$ be feasible for $(P)$. Then $x_0$ is strictly optimal if either of the following two conditions holds:

(i) There is no solution $d$ to the system 

$$f'(x_0)d \leq 0, \quad g'(x_0)d \in -K, \quad h'(x_0)d = 0, \quad d \neq 0.$$ 

(ii) For every solution $d$ of $(11.1)$ there correspond multipliers $u^* \in C^+$, $v^* \in K^+$ and $w^* \in W^+$ such that

$$u^*f'(x_0)d = 0, \quad v^*g'(x_0)d = 0, \quad v^*g(x_0) = 0 \quad (11.2)$$

$$u^*f'(x_0) + v^*g'(x_0) + w^*h'(x_0) = 0 \quad (11.3)$$

and

$$u^*f''(x_0)(d,d) + v^*g''(x_0)(d,d) + w^*h''(x_0)(d,d) > 0. \quad (11.4)$$

**Proof:** The proof is indirect. Suppose $x_0$ is not strict optimal. Then there is a sequence $\{x_n\}_{n=1}^\infty$, $x_n \neq x_0$ and $x_n \to x_0$, such that for all $n$
and \( g(x_n) \in K \), \( h(x_n) = 0 \). From the two last relations we get
\[
\begin{align*}
g(x_n) - g(x_0) & \in -Kg(x_0) \\
h(x_n) - h(x_0) & = 0.
\end{align*}
\] (11.6)

Put \( z_n := x_n - x_0 \) and \( t_n := \|z_n\| \). Since the boundary of the unit ball in the finite-dimensional space is compact the following limit exists (at least for a suitable subsequence)
\[
d_o := \lim_{n \to \infty} t_n^{-1}z_n.
\]

(11.5), (11.6) imply (note that \( C \) and \( Kg(x_0) \) are closed)
\[
\begin{align*}
f'(x_0)d_o & = \lim_{n \to \infty} t_n^{-1}(f(x_n) - f(x_0)) \in -C \\
g'(x_0)d_o & = \lim_{n \to \infty} t_n^{-1}(g(x_n) - g(x_0)) \in -Kg(x_0) \\
h'(x_0)d_o & = \lim_{n \to \infty} t_n^{-1}(h(x_n) - h(x_0)) = 0.
\end{align*}
\]

Hence \( d_o \) satisfies (11.1). Further \( d_o \not\in 0 \). If (i) holds then we reach a contradiction. Hence suppose (ii) holds, that is, (11.2) - (11.4) hold with suitable \( u^*_o, v^*_o \) and \( w^*_o \).

Now by definition of the derivative
\[
f(x_n) - f(x_0) = f'(x_0)z_n + \frac{1}{2}f''(x_0)(z_n, z_n) + o(t_n^2)
\]
and, because of (11.5),
\[
f'(x_0)z_n + \frac{1}{2}f''(x_0)(z_n, z_n) + o(t_n^2) \in -C.
\]

Similarly,
\[
\begin{align*}
g'(x_0)z_n + \frac{1}{2}g''(x_0)(z_n, z_n) + o(t_n^2) & \in -Kg(x_0) \\
h'(x_0)z_n + \frac{1}{2}h''(x_0)(z_n, z_n) + o(t_n^2) & = 0.
\end{align*}
\]

Multiplying these equations with \( u^*_o, v^*_o, w^*_o \) and summing up while using (11.3) we get (note that \( v^*_o \in Kg(x_0) + \) since
\[ v_o^* \in K^+ \text{ and } v_o^* g(x_o) = 0. \]
\[ (u_o^* f''(x_o) + v_o^* g''(x_o) + w_o^* h''(x_o))(z_n, z_n) + o(t_n^2) \leq 0. \]

Dividing by \( t_n^2 \) and letting \( n \to \infty \) we get
\[ (u_o^* f''(x_o) + v_o^* g''(x_o) + w_o^* h''(x_o)(d_0, d_0) \leq 0 \]
which contradicts (11.4).

\[ \text{qed} \]

A simple modification of the above proof gives us a similar condition which for \( U = R \) and finite-dimensional \( V \) and \( W \) becomes Theorem 7.3 in chapter 4 of HESTENES [3].

The assumption \( \dim X < \infty \) is essential for the above statement; see the following example.

**Example 1**: Let \( X \) be the real sequence space \( l_2 \) and \( U = R \).
Choose some \( s \in X \) with components \( s_n > 0 \) for all \( n \) and consider the free problem

\[
\text{minimize } f(x) = \sum_{i=1}^{\infty} s_i x_i^2 - \sum_{i=1}^{\infty} x_i^4.
\]

Let \( x_o \) be the origin in \( l_2 \). Then \( f'(x_o) \) is the zero-functional on \( l_2 \) and
\[ f''(x_o)(d,d) = 2 \sum s_i d_i^2 \text{ for all } d \in l_2. \]
Hence (11.2) \( - \) (11.4) holds with \( u^* = 1 \) for all \( d \in l_2, d \neq 0 \). Nevertheless, \( x_o \) is not optimal. Take the sequence
\[ x^n := 2 \sqrt{s_n} \cdot 1^n \text{ for } n=1,2,\ldots \text{ where } 1^n \text{ is the vector with all components 0 except the } n^\text{th} \text{ component which is 1. Then } x^n \to x_o \text{ but } \]
\[ f(x^n) = 4s_n^2 - 16s_n^2 < 0 \text{ for all } n. \]
For infinite-dimensional $X$, the following sufficient second order condition can be proved; see MAURER-ZOWE [21].

**Theorem 11.2** Consider problem $(P)$ with $J = \mathbb{R}$, with real Banach spaces $X$, $V$, $W$ and twice $F$-differentiable functions. Let $x_0$ be feasible for $(P)$ and suppose there is a map $x \to d(x)$ of the feasible set $F$ into the linear approximation $\{d \in X : g'(x_0)d \in -Kg(x_0), h'(x_0)d = 0\}$ of $F$ at $x_0$ such that

$$\|d(x) - (x-x_0)\| = o(\|x-x_0\|) \text{ for all feasible } x. \quad (11.7)$$

Suppose there are multipliers $w^* \in W^*$ and $v^* \in K^+$ with $v^*g(x_0) = 0$ and, further, positive reals $\beta$, $\delta$ such that

$$f'(x_0) + v^*g'(x_0) + w^*h'(x_0) = 0$$

and

$$(f''(x_0) + v^*g''(x_0) + w^*h''(x_0))(d,d) \geq \delta \|d\|^2$$

for all $d \in \{y \in X : g'(x_0)y \in -Kg(x_0), h'(x_0)y = 0, v^*g'(x_0)y \leq \beta \|y\|\}$. Then $x_0$ is strictly optimal.

For conditions guaranteeing (11.7) see MAURER-ZOWE [21].
12. Summary of necessary and sufficient conditions

Necessary conditions for the general problem (P)

$x_0$ is an optimal solution for (P)

<table>
<thead>
<tr>
<th>Statement</th>
<th>Source</th>
<th>Remark</th>
</tr>
</thead>
<tbody>
<tr>
<td>For every $d \in D_f(x_0) \cap D_g(x_0) \cap T_h(x_0)$ for which $f$, $g$ and $h$ are $d$-regular there are continuous linear functionals $1_f \in \mathcal{A}(Q_f(x_0,d))$, $1_g \in \mathcal{A}(Q_g(x_0,d))$, $1_h \in \mathcal{A}(V_h(x_0,d))$, not all zero, such that $1_f + 1_g + 1_h = 0$ $\delta^<em>(1_f(Q_f(x_0,d)) + \delta^</em>(1_g(Q_g(x_0,d)) + \delta^*(1_h(V_h(x_0,d))) \leq 0$</td>
<td>Theorem 2.1</td>
<td>If $CQ(d)$ holds then $1_f$ is not the zero-functional (Remark 2 in Section 4)</td>
</tr>
<tr>
<td>For every $d$ satisfying $f'(x_0)d \leq 0$, $g'(x_0)d \in -\text{cone}(K+g(x_0))$, $h'(x_0)d = 0$, there are continuous linear functionals $u^* \in K^<em>$, $v^</em> \in W^*$, not all zero, such that $u^*f'(x_0)d = 0$, $v^*g(x_0) = 0$, $w^*h'(x_0) = 0$ $u^*f'(x_0) + v^*g(x_0) + w^*h'(x_0) = 0$ $u^*f''(x_0)d + v^*g''(x_0) + w^*h''(x_0)(d,d) \geq 0$</td>
<td>Theorem 8.2</td>
<td>$X$, $W$ Banach spaces, $h'(x_0)X$ closed. If $CQ(d)$ holds then $h^*$ is not the zero functional (Corollary 8.3)</td>
</tr>
</tbody>
</table>

cone($K+g(x_0)$) = \{ $k(k+g(x_0)) \mid k \in K$, $k \geq 0$ \}

$K_g(x_0) = \text{cl cone}(K+g(x_0))$

$K_g(x_0), g'(x_0)d = \text{cl cone}(K_g(x_0) + g'(x_0)d)$

$CQ(d): Q_g(x_0,d) \cap V_h(x_0,d) \neq \emptyset$

There exists $z \in X$ such that

$CQ(d): \begin{cases} g'(x_0)z + g''(x_0)(d,d) \notin \text{int } K_g(x_0), g'(x_0)d \\ h'(x_0)z + h''(x_0)(d,d) = 0 \end{cases}$ (differentiable version)
### Necessary conditions for the infinite-dimensional mathematical programming problem (MP)

$x_0$ is an optimal solution for (MP)

<table>
<thead>
<tr>
<th>Statement</th>
<th>Source</th>
<th>Remark</th>
</tr>
</thead>
<tbody>
<tr>
<td>For every $d \in D_2(x_0) \cap \bigcap_{i \in I(x_0)} D_{g_i}(x_0) \cap T_h(x_0)$, for which $f$, $h$ and $g_i$ $i \in I(x_0)$, are $d$-regular, there are continuous linear functionals $l_f \in \mathcal{A}(Q_f(x_0,d))$, $l_h \in \mathcal{A}(v_h(x_0,d))$, $l_{g_i} \in \mathcal{A}(g_{g_i}(x_0,d))$ for $i \in J(x_0,d)$, not all zero, such that $l_f + \sum_{i \in J(x_0,d)} l_{g_i} + l_h = 0$ and $\delta^*(l_f</td>
<td>Q_f(x_0,d)) + \sum_{i \in J(x_0,d)} \delta^*(l_{g_i}</td>
<td>g_{g_i}(x_0,d)) + \delta^*(l_h</td>
</tr>
</tbody>
</table>

### Non-differentiable case

For every $d$ satisfying $f'(x_0)d \leq 0$, $g_i'(x_0)d \leq 0$ for $i \in I(x_0)$, $h'(x_0)d = 0$ (*), there are real numbers $y_0 \geq 0$, $y_1 \geq 0$ for $i \in J(x_0,d)$, and a functional $w* \in W^*$, not all zero, such that

$$y_0 f''(x_0) + \sum_{i \in I(x_0,d)} y_i g_i''(x_0) + w* h''(x_0) = 0$$

and

$$(f''(x_0) + \sum_{i \in I(x_0,d)} g_i''(x_0) + w* h''(x_0))(d,d) > 0$$

for all $d$ satisfying $g_i'(x_0)d = 0$ for $i \in I(x_0)$ and $y_1 > 0$ $g_i'(x_0)d \leq 0$ for $i \in I(x_0)$ and $y_1 = 0$ $h'(x_0)d = 0$.

### Corollary 9.4

If either $\text{CQ1}$ or $\text{CQ2} + \text{CQ}(d)$ for all $d$ satisfying (*), then there are multipliers $y_i$ for $i \in I(x_0)$ and $w* \in W^*$ such that

$$f''(x_0) + \sum_{i \in I(x_0,d)} y_i g_i''(x_0) + w* h''(x_0) = 0$$

and

$$(f''(x_0) + \sum_{i \in I(x_0,d)} g_i''(x_0) + w* h''(x_0))(d,d) > 0$$

for all $d$ satisfying $g_i'(x_0)d = 0$ for $i \in I(x_0)$ and $y_1 > 0$ $g_i'(x_0)d \leq 0$ for $i \in I(x_0)$ and $y_1 = 0$ $h'(x_0)d = 0$.

### CQ1

$x = (g_i'(x_0)x_0 x_0 h'(x_0)x_0)$ maps $X$ onto $\mathbb{R}^{\text{card}I(x_0)} \times W^*$. (dim $X < \infty$, dim $W < \infty$; $\text{CQ}(d)$ for $i \in I(x_0)$ and $v_i(x_0)$ for all $j$ are lin. independent)

### CQ2

For every $d$ satisfying (*), there is $z \in X$ such that

$$h_i'(x_0)z + g_i''(x_0)(d,d) = 0$$

for $i \in I(x_0)$, $h'(x_0)z + h''(x_0)(d,d) = 0$.

(dim $X < \infty$; dim $W < \infty$; $\text{CQ2}$ becomes the second order constraint qualification (for a definition, see Chapter 7).)
Sufficient conditions for problem (P) under differentiability assumptions

$x_0$ is a feasible point for (P).

<table>
<thead>
<tr>
<th>Finite-dimensional $X$</th>
<th>Statement</th>
<th>Source</th>
<th>Remark</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_0$ is strictly optimal if either of the two conditions holds:</td>
<td>$f'(x_0)d \leq 0$, $g'(x_0)d \in -K_g(x_0)$, $h'(x_0)d = 0$, $d \neq 0$ (*)</td>
<td>Theorem 11.1</td>
<td>Théorem fails for infinite-dimensional $X$ (see Example 1 in Section 11)</td>
</tr>
<tr>
<td>(ii) For every solution $d$ of (<em>) there are multipliers $u^</em> \in C^r$, $v^* \in K^r$, $w^* \in W^r$ such that</td>
<td>$u^*f'(x_0)d = 0$, $v^*g'(x_0)d = 0$, $v^*g(x_0) = 0$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$u^*f'(x_0) + v^*g'(x_0) + w^*h'(x_0) = 0$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$(u^*f''(x_0) + v^*g''(x_0) + w^*h''(x_0))(d, d) &gt; 0$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Infinite-dimensional $X$</th>
<th>Statement</th>
<th>Source</th>
<th>Remark</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_0$ is strictly optimal if the approximation property holds (see below) and if there are multipliers $w^* \in W^r$ and $v^* \in K^r$ with $v^*g(x_0) = 0$ and further real $\beta &gt; 0$, $\epsilon &gt; 0$ such that</td>
<td>$f'(x_0) + v^*g'(x_0) + w^*h'(x_0) = 0$</td>
<td>Theorem 11.2</td>
<td>U = $R$</td>
</tr>
<tr>
<td>$f''(x_0) + v^*g''(x_0) + w^*h''(x_0)(d, d) \geq \delta |d|^2$</td>
<td></td>
<td>$X_tV,W$ Banach spaces.</td>
<td></td>
</tr>
<tr>
<td>for all $d \in {y \in X : g'(x_0)y \in -K_g(x_0), h'(x_0)y = 0, v^*g'(x_0)y \leq \beta |y|}$</td>
<td></td>
<td>Constraint qualification (see Approximation property below)</td>
<td></td>
</tr>
</tbody>
</table>

$K_g(x_0) = \text{cl} \text{cone}(K+g(x_0))$

**Approximation property** (holds for finite-dimensional $X$): there is a map $x \rightarrow d(x)$ of the feasible set into the linear approximation $\{d' \in X : g'(x_0)d' \in -K_g(x_0), h'(x_0)d' = 0\}$ of the feasible set at $x_0$, such that $\|d(x) - (x-x_0)\| = o(\|x-x_0\|)$ for all feasible $x$. 

Technion - Computer Science Department - Technical Report CS0156 - 1979
13. References


