Mixed constraints in optimal control: an implicit function theorem approach

MARIA DO ROSÁRIO DE PINHO†
Universidade do Porto, Rua Dr Roberto Frias, 4200-465 Porto, Portugal

AND

JAVIER F. ROSENBLUETH‡
IIMAS–UNAM, Apartado Postal 20-726, México DF 01000, México

[Received on 27 May 2005; accepted on 15 December 2005]

This paper concerns a derivation of second-order necessary conditions for a fixed-endpoint control problem of Lagrange involving mixed equality and/or inequality constraints, posed over piecewise continuous controls. These conditions are obtained in a clear and transparent way by reducing the original problem, through an implicit function theorem approach, to an unconstrained control problem.

Keywords: optimal control; mixed constraints; second-order necessary conditions; normality.

1. Introduction

According to a widely quoted paper by Gilbert and Bernstein (1983), ‘mathematical rigorous treatments of second order necessary conditions for problems in optimal control seem to be limited’. In that paper, two different sets of second-order conditions are obtained by applying necessary conditions for an abstract optimization problem developed by the second author (Bernstein, 1984). These conditions are compared with, and shown to be a generalization of, those previously derived by Hestenes (1966) and Warga (1978).

In a brief summary of the problems considered in those two references, Gilbert and Bernstein write: ‘Hestenes, whose work is the earliest, considered a fairly general optimal control problem but made the standard assumption that the control set is open. His main result states that the second variation of a suitably defined function is nonnegative on a set of admissible variations related to the first order necessary conditions. More recently, Warga obtained a similar result, stated in a somewhat different way, for problems where the controls are restricted to a convex, not necessarily open, constrained set.’

Since then, this question has been studied by a large number of authors. In the literature, one can now find second-order conditions for a wide variety of different specific optimal control problems according to the constraints, the spaces of admissible processes, the assumptions on the functions delimiting the problem, and so on (see e.g. Arutyunov & Vereshchagina, 2002; Loewen & Zheng, 1994; Milyutin & Osmolovskii, 1998; Osmolovskii, 1975; Stefani & Zezza, 1996; Zeidan, 1994, 1996; Zeidan & Zezza, 1988, and references therein). Not all coincide, and it may be extremely cumbersome to compare between different problems and the conditions obtained, but this is not the issue of this paper.

†Email: mrpinho@fe.up.pt
‡Corresponding author. Email: jfrl@servidor.unam.mx

© The author 2006. Published by Oxford University Press on behalf of the Institute of Mathematics and its Applications. All rights reserved.
Our aim is to give a clear explanation of how the technique used by Hestenes (1966) to obtain second-order necessary conditions for unconstrained problems can be used to derive such conditions, for problems involving mixed state-control equality and/or inequality constraints, once the original problem is reduced to a problem without constraints. This reduction is based on a uniform implicit function theorem established precisely in Hestenes (1966), where first-order conditions for constrained problems are derived in that way.

For simplicity of exposition and to keep notational complexity to a minimum, we shall deal with the fixed-endpoint control problem of Lagrange (without isoperimetric constraints) as presented in Hestenes (1966, p. 251). A common and concise way of formulating this problem is as follows:

Minimize $I(x, u) = \int_{t_0}^{t_1} L(t, x(t), u(t)) \, dt$ subject to

a. $x: T \rightarrow \mathbb{R}^n$ piecewise $C^1$; $u: T \rightarrow \mathbb{R}^m$ piecewise continuous;

b. $\dot{x}(t) = f(t, x(t), u(t))$ ($t \in T$);

c. $x(t_0) = \xi_0, x(t_1) = \xi_1$;

d. $(t, x(t), u(t)) \in A$ ($t \in T$),

where $T = [t_0, t_1]$ is a fixed compact interval in $\mathbb{R}$. No difficulties arise in the theory to follow if one considers Bolza problems with only one fixed-endpoint as in Loewen & Zheng (1994) and Zeidan & Zezza (1988), or both endpoints varying as in Zeidan (1996), since the convex cones defining the sets of admissible variations depend only on the set of tangential constraints which characterize the set $A$.

We shall consider three cases, namely, when $A$ is a (relatively) open set in $T \times \mathbb{R}^n \times \mathbb{R}^m$, when $A$ is defined by equality constraints and when $A$ is defined by equality–inequality constraints. The first case is taken up from Hestenes (1966) where second-order conditions are obtained assuming normality of the processes under consideration. The notion of normality is introduced so as to assure uniqueness of the Lagrange multipliers, in the first-order necessary conditions, when the cost multiplier is equal to 1. This differs from other authors (see e.g. Loewen & Zheng, 1994; Zeidan, 1996) where normality implies a nonzero cost multiplier. For the second case it is shown that a solution to the problem solves also an unconstrained problem for which the previous first- and second-order conditions can be applied, assuming again a notion of normality which implies the corresponding notion for the unconstrained case.

Finally, we apply a simple technique used by Valentine (1937) to reduce the third case to the second one and, once more, define normality in the event of inequality constraints so as to obtain normality for the reduced, equivalent, equality constrained problem.

2. An unconstrained control problem of Lagrange

Let us point out that, in the three references just mentioned, the equality–inequality constraints are not mixed since they are imposed only on the control functions.

Let us give a clear explanation of how the technique used by Hestenes (1966) to obtain second-order necessary conditions for unconstrained problems can be used to derive such conditions, for problems involving mixed state-control equality and/or inequality constraints, once the original problem is reduced to a problem without constraints. This reduction is based on a uniform implicit function theorem established precisely in Hestenes (1966), where first-order conditions for constrained problems are derived in that way.
and consider the functional $I: Z \to \mathbb{R}$ given by $I(x, u) := \int_0^1 L(t, x(t), u(t))dt \ (x, u) \in Z$. The problem we shall deal with, which we will label $P(A, f, I)$, is that of minimizing $I$ over $Z_e(A, f)$.

Elements of $Z$ will be called \textit{processes}, and a process $(x, u)$ solves $P(A, f, I)$ if $(x, u) \in Z_e(A, f)$ and $I(x, u) \leq I(y, v)$ for all $(y, v) \in Z_e(A, f)$. Above and throughout the paper, ‘$(t \in T)$’ means ‘for all $t$ in $T$’. Also, for any $(x, u) \in Z$, we shall use the notation $(\tilde{x}(t))$ to represent $(t, x(t), u(t))$, and $^T$ will denote transpose. For all $(t, x, u, p, \lambda)$ in $T \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}$ let

$$H(t, x, u, p, \lambda) := \langle p, f(t, x, u) \rangle - \lambda L(t, x, u).$$

The first result we state (maximum principle) is proved in Hestenes (1966, p. 254) under the following assumptions on the data of the problem.

A. The functions $L$, $f$ and their partial derivatives with respect to $x$ are continuous on $\mathcal{A}$, and the set $\mathcal{A}$ is \textit{admissible} in the sense that, for each $(s, v) \in \mathcal{A}$, there exist $\delta > 0$ and $u: [s - \delta, s + \delta] \to \mathbb{R}^m$ continuous such that $u(s) = v$ and $(t, x, (u(t))) \in \mathcal{A}$ for all $(t, x) \in T \times \mathbb{R}^n$ with $|(t, x) - (s, y)| < \delta$.

\textbf{THEOREM 2.1} Assume (A) holds and $(x_0, u_0)$ solves $P(A, f, I)$. Then there exist $\lambda_0 \geq 0$ and $p \in X$, not both zero, such that

\begin{itemize}
  \item[a.] $\dot{p}(t) = -H_x^*(\tilde{x}(t), p(t), \lambda_0)$ on every interval of continuity of $u_0$.
  \item[b.] $H(t, x_0(t), u, p(t), \lambda_0) \leq H(\tilde{x}_0(t), p(t), \lambda_0)$ for all $(t, u) \in T \times \mathbb{R}^m$ with $(t, x_0(t), u) \in \mathcal{A}$.
\end{itemize}

\textbf{NOTE 2.2} In Theorem 2.1, the function $t \mapsto H(\tilde{x}_0(t), p(t), \lambda_0)$ is continuous on $T$. Moreover, if $\mathcal{A}$ is (relatively) open and $L$ and $f$ are continuous on $\mathcal{A}$ and $C^1$ on $\mathcal{A}$ with respect to $u$ then $H_u(\tilde{x}_0(t), p(t), \lambda_0) = 0$. If also $L$ and $f$ are $C^2$ on $\mathcal{A}$ with respect to $u$ then $H_{uu}(\tilde{x}_0(t), p(t), \lambda_0) \leq 0$ ($t \in T$).

In the remaining of this section we assume that the following holds.

B. The functions $L$, $f$ are $C^2$ on $\mathcal{A}$ and $\mathcal{A}$ is (relatively) open.

\textbf{DEFINITION 2.3}

- \textit{Admissible variations}
  Given $(x, u) \in Z$ let $A(t) := f_x(\tilde{x}(t))$, $B(t) := f_u(\tilde{x}(t))$ ($t \in T$), and define the set $Y_0(x, u)$ of $\mathcal{A}$-admissible variations along $(x, u)$ as the set of all $(y, v) \in Z$ satisfying
    \begin{itemize}
      \item[i.] $y(t_0) = y(t_1) = 0$.
      \item[ii.] $\dot{y}(t) = A(t)y(t) + B(t)v(t)$ ($t \in T$).
    \end{itemize}

- The sets $M_0$, $\mathcal{E}_0$ and $\mathcal{H}_0$
  For all $(x, u) \in Z$ let
    \begin{align*}
      M_0(x, u) := \{(\lambda_0, p) \in \mathbb{R} \times X \mid \lambda_0 \geq 0, \lambda_0 + |p| \neq 0, \dot{p}(t) = -H_x^*(t, \lambda_0), H_u(t, \lambda_0) = 0 (t \in T)\},
    \end{align*}
  where $H(t, \lambda_0)$ denotes $H(\tilde{x}(t), p(t), \lambda_0)$, and consider the following sets:
    \begin{align*}
      \mathcal{E}_0 &:= \{(x, u, p) \in Z \times X \mid (x, u) \in D(f) \text{ and } (1, p) \in M_0(x, u)\}, \\
      \mathcal{H}_0 &:= \{(x, u, p) \in Z \times X \mid J_0((x, u); (y, v)) \geq 0 \text{ for all } (y, v) \in Y_0(x, u)\},
    \end{align*}

    where
    \begin{align*}
      J_0((x, u); (y, v)) = \int_0^{t_1} 2\Omega_0(t, y(t), v(t))dt \quad ((y, v) \in Z)
    \end{align*}
and, for all \((t, y, v) \in T \times \mathbb{R}^n \times \mathbb{R}^m\),
\[
2\Omega_0(t, y, v) := -[\langle y, H_{xx}(t, 1)y \rangle + 2\langle y, H_{xu}(t, 1)v \rangle + \langle v, H_{uu}(t, 1)v \rangle],
\]
where \(H(t, 1)\) denotes \(H(\tilde{x}(t), p(t), 1)\).

- **Normality**
  A process \((x, u)\) will be said to be normal to \(P(A, f, I)\) if the equations
  \[
  \dot{p}(t) = -A^*(t)p(t) \quad [= -H^*_x(\tilde{x}(t), p(t), 0)]
  \]
  \[
  0 = B^*(t)p(t) \quad [= H^*_u(\tilde{x}(t), p(t), 0)]
  \]
  have no non-null solution \(p\) on \(T\).

The following theorem provides first- and second-order conditions for problem \(P(A, f, I)\). These conditions can be written in a succinct way in terms of \(E_0\) and \(H_0\).

**Theorem 2.4** Suppose \((x, u)\) solves \(P(A, f, I)\). Then \(M_0(x, u) \neq \emptyset\). If \((x, u)\) is normal to \(P(A, f, I)\), then there exists a unique \(p \in X\) such that \((x, u, p) \in E_0\). Moreover, \((x, u, p) \in H_0\).

It is a common practice to state results of this nature by giving certain names to the assumptions, conditions and functions involved. In contrast, we have chosen this notation not only because it allows us to express the results in a concise way but, mainly, to avoid misleading interpretations which may easily occur when the same words or phrases are used with different meanings.

Now, we are interested in showing how the second-order condition \(\{x, u, p\} \in H_0\) can be easily established. To do so, let us state a crucial auxiliary result from Hestenes (1966, pp. 274, 280).

**Lemma 2.5** Suppose \((x_0, u_0) \in Z_e(A, f)\) is normal to \(P(A, f, I)\) and \((y, v) \in Y_0(x_0, u_0)\). Then there exist \(\delta > 0\) and a one-parameter family \((x(\cdot, \epsilon), u(\cdot, \epsilon)) \in Z_e(A, f) (|\epsilon| < \delta)\) such that

i. \(x(t, 0) = x_0(t), u(t, 0) = u_0(t) \ (t \in T)\).

ii. \(x_\epsilon(t, 0) = y(t), u_\epsilon(t, 0) = v(t) \ (t \in T)\).

**Proof of Theorem 2.4.** Suppose \((x_0, u_0)\) solves \(P(A, f, I)\). By Theorem 2.1 and Note 2.2, \(M_0(x_0, u_0) \neq \emptyset\). Let \((\lambda_0, p) \in M_0(x_0, u_0)\) and suppose \((x_0, u_0)\) is normal to \(P(A, f, I)\). This implies that \(\lambda_0 \neq 0\) and, if \((\lambda_0, q) \in M_0(x_0, u_0)\), then \(r := p - q\) satisfies \(\dot{r}(t) = -A^*(t)r(t), 0 = B^*(t)r(t) \ (t \in T)\), implying that \(p \equiv q\). Clearly, we can choose \(\lambda_0 = 1\) since \((1, p/\lambda_0) \in M_0(x_0, u_0)\). Suppose therefore that \((x_0, u_0, p) \in E_0\). To show that \((x_0, u_0, p) \in H_0\), define

\[
K(x, u) := \langle p(t_1), \tilde{z}_1 \rangle - \langle p(t_0), \tilde{z}_0 \rangle + \int_{t_0}^{t_1} F(t, x(t), u(t))dt \quad ((x, u) \in Z),
\]

where, for all \((t, x, u) \in T \times \mathbb{R}^n \times \mathbb{R}^m\),
\[
F(t, x, u) := L(t, x, u) - \langle p(t), f(t, x, u) \rangle - \langle \dot{p}(t), x \rangle.
\]

Observe that
\[
F(t, x, u) = -H(t, x, u, p(t), 1) - \langle \dot{p}(t), x \rangle
\]
and, if \((x, u) \in Z_e(A, f)\), then \(K(x, u) = I(x, u)\). Let \((y, v) \in Y_0(x_0, u_0)\) and let \((x(\cdot, \epsilon), u(\cdot, \epsilon)) \in Z_e(A, f) (|\epsilon| < \delta)\) be a one-parameter family satisfying Lemma 2.5. Hence,
\[
g(\epsilon) := K(x(\cdot, \epsilon), u(\cdot, \epsilon)) = I(x(\cdot, \epsilon), u(\cdot, \epsilon)) \quad (|\epsilon| < \delta)
satisfies \( g(\epsilon) \geq g(0) = K(x_0, u_0) = I(x_0, u_0) \). Note that
\[
F_x(\tilde{x}_0(t)) = L_x(\tilde{x}_0(t)) - p^*(t) f_x(\tilde{x}_0(t)) - \dot{p}^*(t) = 0, \quad F_u(\tilde{x}_0(t)) = -H_u(\tilde{x}_0(t), p(t), 1) = 0
\]
and therefore
\[
0 \leq g''(0) = \int_{t_0}^{t_1} \left\{ \langle F_x(\tilde{x}_0(t)), x_{ee}(t, 0) \rangle + \langle F_u(\tilde{x}_0(t)), u_{ee}(t, 0) \rangle \right\} dt + K''((x_0, u_0); (y, v)) = K''((x_0, u_0); (y, v)) = J_0((x_0, u_0); (y, v)).
\]

3. Auxiliary results

In this section we state several auxiliary results, taken up from Hestenes (1966), which play a fundamental role in the theory to follow. We begin with the implicit function theorem as stated in Hestenes (1966, p. 23).

**THEOREM 3.1** Let \( f: A \to \mathbb{R}^n \), where \( A \subset \mathbb{R}^m \times \mathbb{R}^n \) is open. Suppose \( f \) and \( f_x(t, x) \) are continuous on \( A \) and there exist \( T_0 \subset \mathbb{R}^m \) compact and \( x_0: T_0 \to \mathbb{R}^n \) continuous such that, for all \( t \in T_0 \),

i. \( (t, x_0(t)) \in A \).
ii. \( f(t, x_0(t)) = 0 \) and \( |f_x(t, x_0(t))| \neq 0 \).

Then there exist \( T \) neighbourhood of \( T_0, \epsilon > 0 \) and \( x: T \to \mathbb{R}^n \) continuous, such that

a. \( x(t) = x_0(t) \) for all \( t \in T_0 \).
b. \( f(t, x(t)) = 0 \) for all \( t \in T \).
c. \( t \in T, f(t, x) = 0 \) and \( |x - x(t)| < \epsilon \) \( \Rightarrow x = x(t) \).
d. \( f \in C^m(A) \Rightarrow x \in C^m(T) \).

The next result is based on the implicit function theorem and, as we shall see in the following section, it allows to transform an optimization problem with equality constraints into an unconstrained problem. We give a partial proof (see Hestenes, 1966, p. 206, for the remaining claims) since some of the functions appearing in it will be used explicitly in the next two sections.

**LEMMA 3.2** Let \( \phi \) be a \( C^1 \) function mapping \( \mathbb{R}^n \times \mathbb{R}^m \) to \( \mathbb{R}^q \). Let
\[
S := \{(x, u) \in \mathbb{R}^n \times \mathbb{R}^m \mid \phi(x, u) = 0\}
\]
and suppose the matrix \( \phi_u(x, u) \) has rank \( q \) on \( S \). Then there exists a neighbourhood \( \tilde{S} \) of \( S \) and a continuous function \( U: \tilde{S} \to \mathbb{R}^m \) such that

a. \( (x, U(x, u)) \in S \) for all \( (x, u) \in \tilde{S} \).
b. \( U(x, u) = u \) for all \( (x, u) \in S \).

If \( \phi \) is of class \( C^r \) then \( U \) can be chosen to be of class \( C^r \) on \( \tilde{S} \).

**Proof.** Let \( h: \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^q \to \mathbb{R}^q \) be given by \( h(x, u, b) := \phi(x, u + \phi_u^*(x, u)b) \). We have \( h \) and \( h_b \) continuous on \( \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^q \) and, for all \( (x, u) \in S \),
\[
h(x, u, 0) = 0 \quad \text{and} \quad |h_b(x, u, 0)| = |\phi_u(x, u)\phi_u^*(x, u)| \neq 0.
\]
Suppose first that $S \subset \mathbb{R}^n \times \mathbb{R}^m$ is compact. By Theorem 3.1 there exists $\tilde{S}$ neighbourhood of $S$, $\epsilon > 0$ and $B: \tilde{S} \to \mathbb{R}^q$ continuous, such that

a. $B(x, u) = 0$ for all $(x, u) \in S$.
b. $h(x, u, B(x, u)) = 0$ for all $(x, u) \in \tilde{S}$.
c. $((x, u) \in \tilde{S}, h(x, u, b) = 0$ and $|b - B(x, u)| < \epsilon) \Rightarrow b = B(x, u)$.
d. $h \in C^r \Rightarrow B \in C^r(\tilde{S})$.

Let $U(x, u) = u + \varphi_u^*(x, u)B(x, u)$ for all $(x, u) \in \tilde{S}$. Then $U$ satisfies the required properties. For the case $S$ not compact and the fact that $U$ can be chosen to be of class $C^r$ on $\tilde{S}$, we refer to Hestenes’ book.

\textit{NOTE 3.3} Observe that the function $B: \tilde{S} \to \mathbb{R}^q$ satisfies

$$B'(x, u) = -(h_b(x, u, 0))^{-1}h_x,u(x, u, 0) \quad \text{for all } (x, u) \in S$$

and so, if $D(x, u) := [\varphi_u(x, u)\varphi_u^*(x, u)]^{-1}$, then

$$B_x(x, u) = -D(x, u)\varphi_x(x, u), \quad B_u(x, u) = -D(x, u)\varphi_u(x, u) \quad \text{for all } (x, u) \in S.$$ 

Therefore $U: \tilde{S} \to \mathbb{R}^m$ defined by $U(x, u) = u + \varphi_u^*(x, u)B(x, u)$ satisfies, for any $(x, u) \in S$,

$$U_x(x, u) = -\varphi_u^*(x, u)D(x, u)\varphi_x(x, u), \quad U_u(x, u) = I - \varphi_u^*(x, u)D(x, u)\varphi_u(x, u).$$

We end with necessary conditions for certain optimization problems, involving equality and/or inequality constraints, which are based on the corresponding results for the finite-dimensional case (see Hestenes, 1966, p. 209).

Suppose we are given an interval $T = [a, b]$ in $\mathbb{R}$ and functions $f$ and $g$ mapping $T \times \mathbb{R}^m$ to $\mathbb{R}$ and $\mathbb{R}^q$, respectively, with $f$, $g$ continuous and having continuous partial derivatives with respect to $u$. Let

$$C := \{(t, u) \in T \times \mathbb{R}^m \mid g(t, u) = 0\}$$

and suppose the matrix $g_u(t, u)$ has rank $q$ on $C$. Denote by $\mathcal{U}$ the space of piecewise continuous functions mapping $T$ to $\mathbb{R}^m$ and let

$$\mathcal{U}(C) := \{u \in \mathcal{U} \mid (t, u(t)) \in C \ (t \in T)\}.$$ 

\textit{LEMMA 3.4} Suppose $u_0 \in \mathcal{U}(C)$ and $f(t, u_0(t)) \leq f(t, u)$ for all $t \in T$ such that $(t, u) \in C$. Then there exists a unique $\mu: T \to \mathbb{R}^q$ such that, if

$$F(t, u) := f(t, u) + \langle \mu(t), g(t, u) \rangle \quad ((t, u) \in T \times \mathbb{R}^m),$$

then $F_u(t, u_0(t)) = 0$ $(t \in T)$. Moreover, $(h, F_{uu}(t, u_0(t))) \geq 0$ for all $h \in \mathbb{R}^m$ such that $g_u(t, u_0(t))h = 0$. The function $\mu$ is piecewise continuous on $T$ and continuous at each point of continuity of $u_0$. 

Suppose now that

$$C = \{(t, u) \in T \times \mathbb{R}^m \mid g_\alpha(t, u) \leq 0 \ (\alpha \in A), \ g_\beta(t, u) = 0 \ (\beta \in B)\},$$

where $A$ and $B$ are finite index sets.
where \( A = \{1, \ldots, r\}, B = \{r + 1, \ldots, q\}, \) and assume that the \( q \times (m + r) \)-dimensional matrix

\[
\left( \frac{\partial g_i}{\partial u^k} \delta_{i\alpha} g_{\alpha} \right) \quad (i = 1, \ldots, q; \; \alpha = 1, \ldots, r; \; k = 1, \ldots, m)
\]

has rank \( q \) on \( C \). Here \( \delta_{\alpha\alpha} = 1, \delta_{\alpha\beta} = 0 \) \((\alpha \neq \beta)\). This condition is equivalent (see Section 5 for details) to the condition that at each point \((t, u) \in C\), the matrix

\[
\left( \frac{\partial g_i}{\partial u^k} \right) \quad (i = i_1, \ldots, i_p; \; k = 1, \ldots, m)
\]

has rank \( p \), where \( I(t, u) = \{\alpha \in A \mid g_{\alpha}(t, u) = 0\} \) and \( I(t, u) \cup B = \{i_1, \ldots, i_p\} \).

**Lemma 3.5** Suppose \( u_0 \in \mathcal{U}(C) \) and \( f(t, u_0(t)) \leq f(t, u) \) for all \( t \in T \) such that \((t, u) \in C\). Then there exists a unique \( \mu: T \rightarrow \mathbb{R}^q \) such that, if

\[
F(t, u) := f(t, u) + \langle \mu(t), g(t, u) \rangle \quad ((t, u) \in T \times \mathbb{R}^m),
\]

then \( F_u(t, u_0(t)) = 0 \) \((t \in T)\). The function \( \mu = (\mu_1, \ldots, \mu_q) \) is piecewise continuous on \( T \) and continuous at each point of continuity of \( u_0 \). Moreover, \( \mu_{\alpha}(t) \geq 0 \) \((\alpha \in A, \; t \in T)\) and \( \mu_{\alpha}(t) = 0 \) whenever \( g_{\alpha}(t, u_0(t)) < 0 \).

**4. Equality constraints**

Consider problem \( P(A, f, I) \) of Section 2. Assume \( L, f \) are \( C^2 \) and

\[
A = \{t, x, u \in T \times \mathbb{R}^n \times \mathbb{R}^m \mid \phi(t, x, u) = 0\},
\]

where \( \phi: T \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^q \) is of class \( C^2 \) and the matrix \( \phi_u(t, x, u) \) has rank \( q \) on \( A \). Denote by \( \mathcal{U}_q \) the space of piecewise continuous functions mapping \( T \) to \( \mathbb{R}^q \).

Let us show how first-order conditions for this problem, involving mixed equality constraints, can be easily obtained from those established for the unconstrained problem of Section 2. We simply invoke Lemma 3.2 and apply Theorem 2.1 to the new problem. We provide a detailed proof for completeness.

For all \((t, x, u, p, \mu, \lambda)\) in \( T \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \) let

\[
H(t, x, u, p, \mu, \lambda) := \langle p, f(t, x, u) \rangle - \lambda L(t, x, u) - \langle \mu, \phi(t, x, u) \rangle.
\]

**Theorem 4.1** Suppose \((x_0, u_0)\) solves \( P(A, f, I) \). Then there exist \( \lambda_0 \geq 0, \; p \in X, \) and \( \mu \in \mathcal{U}_q \) continuous on each interval of continuity of \( u_0 \), not vanishing simultaneously on \( T \), such that

a. \( p(t) = -H_x^*(\tilde{x}_0(t), p(t), \mu(t), \lambda_0) \) and \( H_u(\tilde{x}_0(t), p(t), \mu(t), \lambda_0) = 0 \) on every interval of continuity of \( u_0 \).

b. \( H(t, x_0(t), u, p(t), \mu(t), \lambda_0) \leq H(\tilde{x}_0(t), p(t), \mu(t), \lambda_0) \) for all \((t, u) \in T \times \mathbb{R}^m \) with \((t, x_0(t), u) \in A\).

**Proof.** By Lemma 3.2 there exists a neighbourhood \( \mathcal{B} \) of \( A \) and a \( C^2 \) function \( U: \mathcal{B} \rightarrow \mathbb{R}^m \) such that \((t, x, U(t, x, u)) \in A \) for all \((t, x, u) \in \mathcal{B} \), and \( U(t, x, u) = u \) for all \((t, x) \in A\). Set

\[
\hat{f}(t, x, u) := f(t, x, U(t, x, u)), \quad \hat{L}(t, x, u) := L(t, x, U(t, x, u)) \quad ((t, x, u) \in \mathcal{B}),
\]

\[
\hat{I}(x, u) := \int_{t_0}^{t_1} \hat{L}(t, x(t), u(t)) \, dt \quad ((x, u) \in Z(\mathcal{B})).
\]
Let us prove that \((x_0, u_0)\) solves \(P(B, \hat{f}, \hat{I})\), i.e.

i. \((x_0, u_0) \in Z_e(B, \hat{f})\).

ii. \(\hat{I}(x_0, u_0) \leq \hat{I}(\hat{x}, \hat{u})\) for all \((\hat{x}, \hat{u}) \in Z_e(B, \hat{f})\).

Since \((x_0, u_0)\) solves \(P(A, f, I)\) we have \((x_0, u_0) \in Z_e(A, f)\) and \(I(x_0, u_0) \leq I(x, u)\) for all \((x, u) \in Z_e(A, f)\). Hence, \((t, x_0(t), u_0(t)) \in A \subset B\) \((t \in T)\), \(\dot{x}_0(t) = \hat{f}(t, x_0(t), u_0(t))\), and so (i) holds. To prove (ii) let \((\hat{x}, \hat{u}) \in Z_e(B, \hat{f})\) and define

\[
x(t) := \hat{x}(t), \quad u(t) := U(t, \hat{x}(t), \hat{u}(t)) \quad (t \in T).
\]

Then \((t, x(t), u(t)) \in A, \dot{x}(t) = f(t, x(t), u(t))\) \((t \in T)\), and so \((x, u) \in Z_e(A, f)\). Also \(I(x, u) = \hat{I}(\hat{x}, \hat{u})\) and \(\hat{I}(x_0, u_0) = I(x_0, u_0)\) and so (ii) holds.

Now, by Theorem 2.1 applied to \((x_0, u_0)\) with respect to \(P(B, \hat{f}, \hat{I})\), there exist \(\lambda_0 \geq 0\) and \(p \in X\), not both zero, such that if

\[
\hat{H}(t, x, u) := \langle p(t), \hat{f}(t, x, u) \rangle - \lambda_0 \hat{L}(t, x, u) \quad ((t, x, u) \in B),
\]

then

a. \(\hat{p}(t) = -\hat{H}_x^*(\bar{x}_0(t))\) on every interval of continuity of \(u_0\).

b. \(\hat{H}(t, x_0(t), u) \leq \hat{H}(\bar{x}_0(t))\) for all \((t, u) \in T \times \mathbb{R}^m\) with \((t, x_0(t), u) \in B\).

Define

\[
G(t, x, u) := \langle p(t), f(t, x, u) \rangle - \lambda_0 L(t, x, u)
\]

so that \(\hat{H}(t, x, u) = G(t, x, U(t, x, u))\) for all \((t, x, u) \in B\). Since \(U(t, x, u) = u\) for all \((t, x, u) \in A\), it follows from (b) that

\[
G(t, x_0(t), u) \leq G(\bar{x}_0(t)) \quad \text{for all} \quad (t, u) \in T \times \mathbb{R}^m \quad \text{with} \quad (t, x_0(t), u) \in A.
\]

Let \(g(t, u) := \varphi(t, x_0(t), u)\) and \(h(t, u) := -G(t, x_0(t), u)\) for all \((t, u) \in T \times \mathbb{R}^m\) and set

\[
S := \{(t, u) \in T \times \mathbb{R}^m \mid g(t, u) = 0\}.
\]

Since \(h(t, u_0(t)) \leq h(t, u)\) for all \((t, u) \in S\), by Lemma 3.4 there exists a unique \(\mu \in U_q\) such that, if we set

\[
F(t, u) := h(t, u) + \langle \mu(t), g(t, u) \rangle,
\]

then \(F_u(t, u_0(t)) = 0\). In other words, if

\[
\hat{G}(t, x, u) := G(t, x, u) - \langle \mu(t), \varphi(t, x, u) \rangle,
\]

so that \(\hat{G}(t, x_0(t), u) = -F(t, u)\), then

\[
0 = \hat{G}_u(\bar{x}_0(t)) = G_u(\bar{x}_0(t)) - \mu^*(t)\varphi_u(\bar{x}_0(t)) \quad (t \in T).
\]

The function \(\mu\) is continuous on each interval of continuity of \(u_0\). Since, for all \((t, x, u) \in B\), we have \((t, x, U(t, x, u)) \in A\) and \(\hat{H}(t, x, u) = G(t, x, U(t, x, u))\), it follows that

\[
\hat{H}(t, x, u) = \hat{G}(t, x, U(t, x, u)) \quad ((t, x, u) \in B).
\]
Consequently,

\[ \hat{H}_x(\tilde{x}_0(t)) = \hat{G}_x(\tilde{x}_0(t)) + \hat{G}_u(\tilde{x}_0(t))U_x(\tilde{x}_0(t)) = \hat{G}_x(\tilde{x}_0(t)). \]

Since \( \hat{G}(t, x, u) = H(t, x, u, p(t), \mu(t), \lambda_0) \), the conditions of the theorem hold.

**Corollary 4.2** Suppose \((x_0, u_0)\) solves \(P(A, f, I)\). Let \((p, \mu, \lambda_0)\) be as in Theorem 4.1. Then

\[ \langle h, H_{uu}(\tilde{x}_0(t), p(t), \mu(t), \lambda_0)h \rangle \leq 0 \]

for all \( h \in \mathbb{R}^m \) such that \( \varphi_u(\tilde{x}_0(t))h = 0 \).

**Note 4.3** Suppose that, for some \((x, u) \in Z, p \in X, \lambda \in \mathbb{R} \) and a function \( \mu : T \to \mathbb{R}^q \), we have

\[ 0 = H_u(\tilde{x}(t), p(t), \mu(t), \lambda) = p^*(t)f_u(\tilde{x}(t)) - \lambda L_u(\tilde{x}(t)) - \mu^*(t)\varphi_u(\tilde{x}(t)) \quad (t \in T). \]

Then \( \mu \in U_q \) and

\[ \mu^*(t) = \left[ p^*(t)f_u(\tilde{x}(t)) - \lambda L_u(\tilde{x}(t))\right]\varphi_u^*(\tilde{x}(t))\left[\varphi_u(\tilde{x}(t))\varphi_u^*(\tilde{x}(t))\right]^{-1} \quad (t \in T). \]

**Definition 4.4**

- **Admissible variations**
  Given \((x, u) \in Z\) let \( A(t) := f_x(\tilde{x}(t)), B(t) := f_u(\tilde{x}(t)) \) \( (t \in T) \), and define the set \( Y_1(x, u) \) of \( A \)-admissible variations along \((x, u)\) as the set of all \((y, v) \in Z\) satisfying
  
  i. \( y(t_0) = y(t_1) = 0 \).
  
  ii. \( \dot{y}(t) = A(t)y(t) + B(t)v(t) \) \( (t \in T) \).
  
  iii. \( \varphi_x(\tilde{x}(t))y(t) + \varphi_u(\tilde{x}(t))v(t) = 0 \) \( (t \in T) \).

- **The sets \( M_1, E_1 \) and \( H_1 \)**
  For all \((x, u) \in Z\) let

  \[ M_1(x, u) := \{ (\lambda_0, p, \mu) \in \mathbb{R} \times X \times U_q \mid \lambda_0 \geq 0, \lambda_0 + |p| \neq 0, \}
  \[ \dot{p}(t) = -H^*_x(t, \lambda_0), \ H_u(t, \lambda_0) = 0 \ (t \in T) \}, \]

  where \( H(t, \lambda_0) \) denotes \( H(\tilde{x}(t), p(t), \mu(t), \lambda_0) \), and consider the following sets:

  \[ E_1 := \{ (x, u, p, \mu) \in Z \times X \times U_q \mid (x, u) \in D(f) \) and \(1, p, \mu) \in M_1(x, u)\}, \]

  \[ H_1 := \{ (x, u, p, \mu) \in Z \times X \times U_q \mid J_1((x, u); (y, v)) \geq 0 \ for \ all \ (y, v) \in Y_1(x, u)\}, \]

  where

  \[ J_1((x, u); (y, v)) = \int_{t_0}^{t_1} 2\Omega_1(t, y(t), v(t))dt \quad ((y, v) \in Z) \]

  and, for all \((t, y, v) \in T \times \mathbb{R}^n \times \mathbb{R}^m \),

  \[ 2\Omega_1(t, y, v) := -\langle (y, H_{xx}(t, 1)y) + 2\langle y, H_{sx}(t, 1)v \rangle + \langle v, H_{uu}(t, 1)v \rangle, \]

  where \( H(t, 1) \) denotes \( H(\tilde{x}(t), p(t), \mu(t), 1) \).
Normality
A process \((x, u)\) will be said to be normal to \(P(A, f, I)\) if, given \((p, \mu)\) \(\in X \times \mathcal{U}_q\) such that, for all \(t \in T\),

\[
\dot{p}(t) = -A^*(t)p(t) + f^*(\tilde{x}(t))\mu(t) \quad [=-H^*_A(\tilde{x}(t), p(t), \mu(t), 0)],
\]

\[
0 = B^*(t)p(t) - \varphi^*_u(\tilde{x}(t))\mu(t) \quad [=H^*_u(\tilde{x}(t), p(t), \mu(t), 0)],
\]

then \(p \equiv 0\). In this event, by Note 4.3, also \(\mu \equiv 0\). Note that, if \((x, u)\) is normal to \(P(A, f, I)\) and \((\lambda_0, p, \mu)\) \(\in M_1(x, u)\), then \(\lambda_0 \neq 0\).

The notion of normality for equality constraints, as defined above, is implied by a different condition stated in the following remark which generalizes the way normality is defined in other references (see e.g. Loewen & Zheng, 1994; Zeidan, 1996; Zeidan & Zezza, 1988).

**Remark 4.5** Let \((x, u)\) \(\in Z\) and suppose that, given \((p, \mu)\) \(\in X \times \mathcal{U}_q\) such that, for all \(t \in T\),

\[
\dot{p}(t) = -A^*(t)p(t) + f^*(\tilde{x}(t))\mu(t) \quad p^*(t)B(t)h = 0 \quad \text{for all } h \in \mathbb{R}^m \text{ satisfying } \varphi_u(\tilde{x}(t))h = 0
\]
necessarily \(p \equiv 0\). Then \((x, u)\) is normal to \(P(A, f, I)\).

Let us now show that normality for the problem with equality constraints, as defined in Definition 4.4, implies normality for the unconstrained problem as defined in Definition 2.3.

**Lemma 4.6** Consider problem \(P(B, \hat{f}, \hat{I})\) given in the proof of Theorem 4.1 and suppose \((x_0, u_0)\) is normal to \(P(A, f, I)\). Then \((x_0, u_0)\) is normal to \(P(B, \hat{f}, \hat{I})\).

**Proof.** We want to prove that \(p \equiv 0\) is the only solution of

\[
\dot{p}(t) = -\hat{A}^*(t)p(t), \quad 0 = \hat{B}^*(t)p(t) \quad (t \in T),
\]

where

\[
\hat{A}(t) = \hat{f}^*(\tilde{x}_0(t)) = A(t) + B(t)U_x(\tilde{x}_0(t)) \quad \text{and} \quad \hat{B}(t) = \hat{f}^*_u(\tilde{x}_0(t)) = B(t)U_u(\tilde{x}_0(t)).
\]

This follows since, by Note 3.3, the above relations correspond to

\[
\dot{p}(t) = -A^*(t)p(t) - U^*_x(\tilde{x}_0(t))B^*(t)p(t) = -A^*(t)p(t) + \varphi^*_u(\tilde{x}_0(t))\mu(t),
\]

\[
0 = U^*_u(\tilde{x}_0(t))B^*(t)p(t) = B^*(t)p(t) - \varphi^*_u(\tilde{x}_0(t))\mu(t),
\]

where

\[
\mu(t) = [\varphi_u(\tilde{x}_0(t))\varphi^*_u(\tilde{x}_0(t))]^{-1}\varphi_u(\tilde{x}_0(t))B^*(t)p(t).
\]

Thus, the normality of \((x_0, u_0)\) to \(P(A, f, I)\) implies that \(p \equiv 0\). \(\square\)

We are now in a position to establish first- and second-order conditions for problem \(P(A, f, I)\) where \(A\) is defined by mixed equality constraints.

**Theorem 4.7** Suppose \((x, u)\) solves \(P(A, f, I)\). Then \(M_1(x, u) \neq \emptyset\). If \((x, u)\) is normal to \(P(A, f, I)\) then there exists a unique \((p, \mu)\) \(\in X \times \mathcal{U}_q\) such that \((x, u, p, \mu)\) \(\in E_1\). Moreover, \((x, u, p, \mu)\) \(\in H_1\).
Proof. Suppose \((x_0, u_0)\) solves \(P(A, f, I)\). By Theorem 4.1, \(M_1(x_0, u_0) \neq \emptyset\). Let \((\lambda_0, p, \mu) \in M_1(x_0, u_0)\) and suppose \((x_0, u_0)\) is normal to \(P(A, f, I)\). This implies that \(\lambda_0 \neq 0\) and, if \((\lambda_0, q, v) \in M_1(x_0, u_0)\) then \(r := p - q\) satisfies

\[
\dot{r}(t) = -A^*(t)r(t) + \varphi_x^*(\tilde{x}_0(t))[\mu(t) - v(t)],
\]

\[
0 = B^*(t)r(t) - \varphi_u^*(\tilde{x}_0(t))[\mu(t) - v(t)],
\]

implying that \(p \equiv q\) and \(\mu \equiv v\).

Let \((p, \mu) \in X \times U_q\) be the unique pair such that \((x_0, u_0, p, \mu) \in E_1\) and define

\[
K(x, u) := \langle p(t_1), \xi_1 \rangle - \langle p(t_0), \xi_0 \rangle + \int_{t_0}^{t_1} F(t, x(t), u(t))dt \quad ((x, u) \in \mathbb{Z}),
\]

where, for all \((t, x, u) \in T \times \mathbb{R}^n \times \mathbb{R}^m\),

\[
F(t, x, u) := L(t, x, u) - \langle p(t), f(t, x, u) \rangle + \langle \mu(t), \varphi(t, x, u) \rangle - \langle \hat{\mu}(t), x \rangle.
\]

Observe that

\[
F(t, x, u) = -H(t, x, u, p(t), \mu(t), 1) - \langle \hat{\mu}(t), x \rangle
\]

and, if \((x, u) \in Z_e(A, f)\), then \(K(x, u) = I(x, u)\). Let \((y, v) \in Y_1(x_0, u_0)\). By Note 3.3 we have

\[
U_x(\tilde{x}_0(t))y(t) + U_u(\tilde{x}_0(t))v(t) = -\varphi_u^*(\tilde{x}_0(t))D(t)[\varphi_x(\tilde{x}_0(t))y(t) + \varphi_u(\tilde{x}_0(t))v(t)] + v(t),
\]

where \(D(t) = [\varphi_u(\tilde{x}_0(t))\varphi_u^*(\tilde{x}_0(t))]^{-1}\). But the right-hand expression equals \(v(t)\) and, therefore,

\[
\dot{y}(t) = A(t)y(t) + B(t)v(t)
\]

\[
= A(t)y(t) + B(t)[U_x(\tilde{x}_0(t))y(t) + U_u(\tilde{x}_0(t))v(t)]
\]

\[
= \hat{A}(t)y(t) + \hat{B}(t)v(t).
\]

Thus, \((y, v) \in Y_0(x_0, u_0)\) with respect to \(P(B, \hat{f}, \hat{I})\). Also, by Lemma 4.6, \((x_0, u_0)\) is normal to \(P(B, \hat{f}, \hat{I})\). Therefore, by Lemma 2.5, there exist \(\delta > 0\) and a one-parameter family \((\hat{x}(\cdot, \epsilon), \hat{u}(\cdot, \epsilon)) \in Z_e(B, \hat{f})\) \((|\epsilon| < \delta)\) such that

i. \(\hat{x}(t, 0) = x_0(t), \hat{u}(t, 0) = u_0(t) \quad (t \in T)\).

ii. \(\hat{x}_\epsilon(t, 0), \hat{u}_\epsilon(t, 0) = v(t) \quad (t \in T)\).

Define \(x(t, \epsilon) := \hat{x}(t, \epsilon)\) and \(u(t, \epsilon) := U(t, \hat{x}(t, \epsilon), \hat{u}(t, \epsilon))\). Clearly, \((x(\cdot, \epsilon), u(\cdot, \epsilon)) \in Z_e(A, f)\).

Let

\[
g(\epsilon) := K(x(\cdot, \epsilon), u(\cdot, \epsilon)) = I(x(\cdot, \epsilon), u(\cdot, \epsilon)) \quad (|\epsilon| < \delta).
\]

Thus, \(g(\epsilon) \geq g(0) \quad (|\epsilon| < \delta)\). Since

\[
u_t(\epsilon, t) = U_x(\tilde{x}_0(t))y(t) + U_u(\tilde{x}_0(t))v(t) = v(t) \quad (|\epsilon| < \delta),
\]

we have \(u_\epsilon(t, 0) = U_x(\tilde{x}_0(t))y(t) + U_u(\tilde{x}_0(t))v(t) = v(t)\) and, therefore,

\[
0 \geq g''(0) = K''((x_0, u_0); (y, v)) = J_1((x_0, u_0); (y, v)).
\]

\qed
5. Full-rank matrices

In the following section we shall study problem \( P(A, f, I) \) in the event that \( A \) is defined by mixed inequality and equality constraints, i.e.

\[
A = \{(t, x, u) \in T \times \mathbb{R}^n \times \mathbb{R}^m \mid \varphi_\alpha(t, x, u) \leq 0 (\alpha \in R), \ \varphi_\beta(t, x, u) = 0 (\beta \in Q)\},
\]

where \( R = \{1, \ldots, r\}, \ Q = \{r + 1, \ldots, q\} \) and \( \varphi: T \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^q \) is given by \( \varphi = (\varphi_1, \ldots, \varphi_q) \).

We shall impose a rank condition on \( \varphi \) which differs from others that can be found in the literature, and we devote this section to a brief discussion on the subject.

Recall first that, if \( X, Y \) are finite-dimensional vector spaces and \( L: X \to Y \) a linear transformation, the rank of \( L \) is defined as the dimension of the subspace \( \text{Im}(L) \). The rank of a matrix \( A \), which we shall denote by \( \rho(A) \), is the rank of the associated linear transformation, and it coincides with the maximum number of linearly independent rows or columns of \( A \). In particular, if \( a_i = (a^1_i, \ldots, a^n_i) \in \mathbb{R}^n \ (i = 1, \ldots, m) \), consider the \( m \times n \) matrix

\[
A = \begin{pmatrix}
a^1_1 & \cdots & a^1_n \\
\vdots & \ddots & \vdots \\
a^m_1 & \cdots & a^m_n
\end{pmatrix}.
\]

Then \( \rho(A) = m \Leftrightarrow \{a_1, \ldots, a_m\} \) is linearly independent. In terms of linear functionals, if \( L_i: \mathbb{R}^n \to \mathbb{R} \ (i = 1, \ldots, m) \) is given, for all \( x = (x^1, \ldots, x^n) \in \mathbb{R}^n \), by

\[
L_i(x) = \sum_{j=1}^n a^j_i x^j \quad (i = 1, \ldots, m),
\]

then \( \rho(A) = m \Leftrightarrow \{L_1, \ldots, L_m\} \) is linearly independent, i.e.

\[
\sum_{i=1}^m \lambda_i L_i(x) = 0 \quad \text{for all } x \in \mathbb{R}^n \implies \lambda_i = 0 \ (i = 1, \ldots, m).
\]

A matrix \( A \in \mathbb{R}^{m \times n} \) (the set of all \( m \times n \) matrices with real elements) is said to be of full rank if \( \rho(A) = \min\{m, n\} \). Clearly, \( A \) is of full rank if and only if its column vectors are linearly independent in \( \mathbb{R}^m \) if \( m \geq n \) while, for \( m \leq n \), its row vectors must be linearly independent in \( \mathbb{R}^n \). Finally, if \( A \in \mathbb{R}^{m \times n} \) and \( m \leq n \), then \( A \) has full rank \( \Leftrightarrow \det AA^* > 0 \) (see e.g. Lancaster & Tismenetsky, 1985).

Now, let us define the matrices we shall be dealing with. Suppose we are given an interval \( T := [t_0, t_1] \) and \( C^1 \) functions \( \varphi_i: T \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R} \ (i = 1, \ldots, q) \) with \( q \leq m \).

Let

\[
A := \{(t, x, u) \in T \times \mathbb{R}^n \times \mathbb{R}^m \mid \varphi_\alpha(t, x, u) \leq 0 (\alpha \in R), \ \varphi_\beta(t, x, u) = 0 (\beta \in Q)\},
\]

where \( R = \{1, \ldots, r\}, \ Q = \{r + 1, \ldots, q\} \). Fix \( (t, x, u) \in A \). In what follows, \( \varphi_i \) and its partial derivatives are all evaluated at \( (t, x, u) \). Let us consider the \( q \times (m + r) \) matrix

\[
A := \left( \begin{array}{c}
\frac{\partial \varphi_i}{\partial t} \\
\delta_{\alpha \beta} \frac{\partial \varphi_i}{\partial u^k} \\
\end{array} \right) \quad (i = 1, \ldots, q; \ \alpha = 1, \ldots, r; \ k = 1, \ldots, m)
\]

(here \( \delta_{\alpha \alpha} = 1, \delta_{\alpha \beta} = 0 (\alpha \neq \beta) \)) and the \( p \times m \) matrix

\[
B := \left( \begin{array}{c}
\frac{\partial \varphi_i}{\partial \tilde{u}^k} \\
\end{array} \right) \quad (i = i_1, \ldots, i_p; \ k = 1, \ldots, m),
\]
where $i_1, \ldots, i_p$ are the indices $i \in R \cup Q$ such that $\varphi_i(t, x, u) = 0$. Explicitly,

$$
A = \begin{pmatrix}
\frac{\partial \varphi_1}{\partial u^1} & \cdots & \frac{\partial \varphi_1}{\partial u^m} & \varphi_1 & \cdots & 0 \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
\frac{\partial \varphi_r}{\partial u^1} & \cdots & \frac{\partial \varphi_r}{\partial u^m} & 0 & \cdots & \varphi_r \\
\frac{\partial \varphi_{r+1}}{\partial u^1} & \cdots & \frac{\partial \varphi_{r+1}}{\partial u^m} & 0 & \cdots & 0 \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
\frac{\partial \varphi_q}{\partial u^1} & \cdots & \frac{\partial \varphi_q}{\partial u^m} & 0 & \cdots & 0
\end{pmatrix}, \quad B = \begin{pmatrix}
\frac{\partial \varphi_{i_1}}{\partial u^1} & \cdots & \frac{\partial \varphi_{i_1}}{\partial u^m} \\
\vdots & \ddots & \vdots \\
\frac{\partial \varphi_{i_p}}{\partial u^1} & \cdots & \frac{\partial \varphi_{i_p}}{\partial u^m}
\end{pmatrix}.
$$

The first question posed is whether the statements ‘$\rho(A) = q$’ and ‘$\rho(B) = p$’ are equivalent or not. Let us prove that the equivalence does certainly hold. Note that this means that $A$ is of full rank if and only if $B$ is of full rank.

**Proposition 5.1** $\rho(A) = q \iff \rho(B) = p$.

**Proof.** Let us begin by defining $L_i: \mathbb{R}^{m+r} \rightarrow \mathbb{R}$ ($i = 1, \ldots, q$) and $M_j: \mathbb{R}^m \rightarrow \mathbb{R}$ ($j = i_1, \ldots, i_p$) as follows. For all $x = (x^1, \ldots, x^{m+r}) \in \mathbb{R}^{m+r}$ and $y = (y^1, \ldots, y^m) \in \mathbb{R}^m$ let

$$
L_i(x) = \sum_{k=1}^m \frac{\partial \varphi_i}{\partial u^k} x^k + \varphi_i x^{m+i} \quad (i = 1, \ldots, r), \quad L_i(x) = \sum_{k=1}^m \frac{\partial \varphi_i}{\partial u^k} x^k \quad (i = r + 1, \ldots, q),
$$

$$
M_j(y) := \sum_{k=1}^m \frac{\partial \varphi_j}{\partial u^k} y^k \quad (j = i_1, \ldots, i_p).
$$

Let $I := \{i_1, \ldots, i_p\}$. We then have

$$
\rho(A) = q \iff \left[ \sum_{i=1}^q \lambda_i L_i(x) = 0 \quad \text{for all} \quad x \in \mathbb{R}^{m+r} \Rightarrow \lambda_i = 0 \quad (i = 1, \ldots, q) \right], \quad (1)
$$

$$
\rho(B) = p \iff \left[ \sum_{i \in I} \mu_i M_i(y) = 0 \quad \text{for all} \quad y \in \mathbb{R}^m \Rightarrow \mu_i = 0 \quad (i \in I) \right]. \quad (2)
$$

‘$\Rightarrow$’ Let $\mu_i \in \mathbb{R}$ ($i \in I$) and suppose

$$
\sum_{i \in I} \mu_i M_i(y) = 0 \quad \text{for all} \quad y \in \mathbb{R}^m.
$$

For all $i = 1, \ldots, q$ let $\lambda_i := \mu_i$ if $i \in I$ and $\lambda_i := 0$ otherwise. Let $x \in \mathbb{R}^{m+r}$. Then

$$
\sum_{i=1}^q \lambda_i L_i(x) = \sum_{i \in I} \mu_i \left( \sum_{k=1}^m \frac{\partial \varphi_i}{\partial u^k} x^k \right) = \sum_{i \in I} \mu_i M_i(x^1, \ldots, x^m) = 0.
$$
By (1), \( \lambda_i = 0 \) \((i = 1, \ldots, q)\). Therefore, \( \mu_i = 0 \) \((i \in I)\) and, by (2), \( \rho(B) = p \).

\( \Leftarrow \): Let \( \lambda_i \in \mathbb{R} \) \((i = 1, \ldots, q)\) and suppose
\[ \sum_{i=1}^{q} \lambda_i L_i(x) = 0 \quad \text{for all } x \in \mathbb{R}^{m+r}. \]

Note that
\[ \sum_{i=1}^{q} \lambda_i L_i(x) = \sum_{i=1}^{m} \left( \sum_{i=1}^{q} \lambda_i \frac{\partial \phi_i}{\partial u^k} \right) x^k + \sum_{i=1}^{r} \lambda_i \phi_i x^{m+i} = 0. \]

Since this holds for all \( x \in \mathbb{R}^{m+r} \) we have \( \lambda_i \phi_i = 0 \) \((i = 1, \ldots, r)\). This implies that \( \lambda_i = 0 \) if \( i \notin I \).

Now, let \( y \in \mathbb{R}^m \) and set \( x := (y^1, \ldots, y^m, 0, \ldots, 0) \) \((\text{the last } r \text{ terms are in fact arbitrary})\). Then
\[ \sum_{i \in I} \lambda_i M_i(y) = \sum_{i \in I} \lambda_i \left( \sum_{k=1}^{m} \frac{\partial \phi_i}{\partial u^k} x^k \right) = \sum_{i=1}^{q} \lambda_i L_i(x) = 0 \]
and so, by (2), also \( \lambda_i = 0 \) if \( i \in I \). By (1), \( \rho(A) = q \).

In what follows we shall write explicitly the dependence on \( t \in T \) of the two matrices \( A \) and \( B \).

Suppose that \( (x, u): T \to \mathbb{R}^n \times \mathbb{R}^m \) is such that, for all \( t \in T \), \((t, x(t), u(t)) \in \mathcal{A} \) and \( \phi_i \) and its partial derivatives are all evaluated at \((t, x(t), u(t))\). For all \( t \in T \), consider the \( q \times (m + r) \) matrix
\[ A(t) := \left( \frac{\partial \phi_i}{\partial u^k} \cdot \delta_{i\alpha} \phi_{\alpha} \right) \quad (i = 1, \ldots, q; \ \alpha = 1, \ldots, r; \ k = 1, \ldots, m) \]
and the \( p \times m \) matrix
\[ B(t) := \left( \frac{\partial \phi_i}{\partial u^k} \right) \quad (i = i_1, \ldots, i_p; \ k = 1, \ldots, m) \]
where \( i_1, \ldots, i_p \) are the indices \( i \in R \cup Q \) such that \( \phi_i(t, x(t), u(t)) = 0 \).

In de Pinho & Ilchmann (2002), a weak maximum principle is derived for problems with possible non-smooth data, involving mixed constraints. One finds there a discussion on different rank assumptions on matrices depending on \( \phi \) expressed in terms of a notion of full rankness which does not coincide with the one given above. According to de Pinho & Ilchmann (2002), a matrix \( F: T \to \mathbb{R}^{m \times n} \) \((m \leq n)\) is of full rank if there exists \( K > 0 \) such that \( \det F(t) F^*(t) \geq K \) a.e. in \( T \). Let us call this property \( \mathcal{R}\text{-full rankness} \). The first-order necessary conditions obtained in de Pinho & Ilchmann (2002) are then derived assuming \( \mathcal{R}\text{-full rankness of the matrix } A \), and it is shown through an example that the statement

\( \text{‘} A \text{ is of } \mathcal{R}\text{-full rank if and only if } B \text{ is of } \mathcal{R}\text{-full rank’} \)

is incorrect.

Let us consider the following statements:

a. \( F(t) \) is of full rank for a.a. \( t \in T \).

b. \( \det F(t) F^*(t) > 0 \) for a.a. \( t \in T \), i.e. for a.a. \( t \in T \) there exists \( K > 0 \) \((\text{depending on } t)\) such that \( \det F(t) F^*(t) \geq K \).

c. \( F \) is of \( \mathcal{R}\text{-full rank} \), i.e. there exists \( K > 0 \) such that \( \det F(t) F^*(t) \geq K \) for a.a. \( t \in T \).
Clearly, (c) ⇒ (b) and we have (a) ⇔ (b), but one can easily show that (b) does not necessarily imply (c). For example, if \( F(t) = t \) for \( t \in [0, 1] \), then (b) holds but not (c). On the other hand, by Proposition 5.1, we have

\[ B \text{ is of } R\text{-full rank } \Rightarrow B(t) \text{ is of full rank for a.a. } t \in T \Leftrightarrow A(t) \text{ is of full rank for a.a. } t \in T \]

but, as the above example shows, this last statement does not necessarily imply that \( A \) is of \( R\)-full rank.

To clarify these remarks, let us take up the example provided in de Pinho & Ilchmann (2002).

**EXAMPLE 5.2** For all \((t, x, u) \in [0, 1] \times \mathbb{R} \times \mathbb{R}^3\), let

\[
\varphi_1(t, x, u) = u_1 + u_2u_3, \quad \varphi_2(t, x, u) = tu_1 + u_2u_3 - t,
\]

so that, for \( t \in [0, 1] \) and \((x, u) \equiv (0, 0)\),

\[
A(t) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -t \end{pmatrix}, \quad B(0) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad B(t) = (1, 0, 0) \quad (t \in (0, 1]).
\]

Therefore,

\[
\det A(t)A^*(t) = t^2 \quad (t \in [0, 1]), \quad \det B(0)B^*(0) = 0, \quad \det B(t)B^*(t) = 1 \quad (t \in (0, 1]).
\]

Clearly, \( B \) is of \( R\)-full rank (implying that \( A(t) \) is of full rank a.e. in \([0, 1]) \) but \( A \) is not of \( R\)-full rank.

Note also that this example does not contradict Proposition 5.1. Indeed, for \( t \neq 0 \), \( A(t) \) and \( B(t) \) are both of full rank and, for \( t = 0 \), neither of them is so.

**6. Inequality constraints**

Consider problem \( P(A, f, I) \) of Section 2. Assume \( L, f \) are \( C^2 \) and

\[
A = \{(t, x, u) \in T \times \mathbb{R}^n \times \mathbb{R}^m \mid \varphi_{\alpha}(t, x, u) \leq 0 \ (\alpha \in R), \ \varphi_{\beta}(t, x, u) = 0 \ (\beta \in Q)\},
\]

where \( R = \{1, \ldots, r\} \), \( Q = \{r + 1, \ldots, q\} \). Assume \( \varphi: T \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^q \) given by \( \varphi = (\varphi_1, \ldots, \varphi_q) \) is of class \( C^2 \) and the \( q \times (m + r) \)-dimensional matrix

\[
\left( \frac{\partial \varphi_i}{\partial u^k} \right)_{i = 1, \ldots, q; \ \alpha = 1, \ldots, r; \ k = 1, \ldots, m} \]

has rank \( q \) on \( A \). As shown in the previous section, this condition is equivalent to the condition that at each point \((t, x, u) \) in \( A \), the matrix

\[
\left( \frac{\partial \varphi_i}{\partial u^k} \right)_{i = i_1, \ldots, i_p; \ k = 1, \ldots, m}
\]

has rank \( p \), where \( i_1, \ldots, i_p \) are the indices \( i \in \{1, \ldots, q\} \) such that \( \varphi_i(t, x, u) = 0 \).

Let us now transform problem \( P(A, f, I) \) into a problem involving only equality constraints. This can be achieved by a simple device used by Valentine (1937).

**Note 6.1** Let

\[
\hat{A} := \{(t, x, u, w) \in T \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^r \mid \psi(t, x, u, w) = 0\},
\]
where \( \psi = (\psi_1, \ldots, \psi_q) \) and
\[
\psi_\alpha(t, x, u, w) = \varphi_\alpha(t, x, u) + (w^\alpha)^2 \quad (\alpha \in R), \quad \psi_\beta(t, x, u, w) = \varphi_\beta(t, x, u) \quad (\beta \in Q).
\]
Define \( \hat{f}(t, x, u, w) := f(t, x, u) \), \( \hat{L}(t, x, u, w) := L(t, x, u) \), and
\[
\hat{I}(x, u, w) := \int_0^t \hat{L}(t, x(t), u(t), w(t))dt.
\]

Then, as one readily verifies, the following are equivalent:

a. \((x_0, u_0)\) solves \( P(\mathcal{A}, f, I) \) and \( w_0^\alpha(t) = [-\varphi_\alpha(\tilde{x}_0(t))]^{1/2} \quad (\alpha \in R, \ t \in T) \).

b. \((x_0, u_0, w_0)\) solves \( P(\hat{\mathcal{A}}, \hat{f}, \hat{I}) \).

For all \((t, x, u, p, \mu, \lambda)\) in \( T \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^q \times \mathbb{R} \), let
\[
H(t, x, u, p, \mu, \lambda) := \langle p, f(t, x, u) \rangle - \lambda L(t, x, u) - \langle \mu, \varphi(t, x, u) \rangle.
\]

Note that this function coincides with the function \( H \) defined in Section 4.

**Theorem 6.2** Suppose \((x_0, u_0)\) solves \( P(\mathcal{A}, f, I) \). Then there exist \( \lambda_0 \geq 0, \ p \in X \) and \( \mu \in U_q \) continuous on each interval of continuity of \( u_0 \), not vanishing simultaneously on \( T \), such that

a. \( \dot{p}(t) = -H_x^*(\tilde{x}_0(t), p(t), \mu(t), \lambda_0) \) and \( H_{\mu}(\tilde{x}_0(t), p(t), \mu(t), \lambda_0) = 0 \) on every interval of continuity of \( u_0 \).

b. \( H(t, x_0(t), u, p(t), 0, \lambda_0) \leq H(\tilde{x}_0(t), p(t), 0, \lambda_0) \) for all \((t, u, p)\) in \( T \times \mathbb{R}^m \) with \((t, x_0(t), u) \in \mathcal{A} \).

In addition, \( \mu_\alpha(t) > 0 \quad (\alpha \in R) \) with \( \mu_\alpha(t) = 0 \) whenever \( \varphi_\alpha(\tilde{x}_0(t)) < 0 \).

**Proof.** Observe first that our assumptions on \( \varphi \) imply that the \( q \times (m + r) \)-dimensional matrix
\[
\psi_{\alpha\varphi}(w) = \left( \frac{\partial \varphi_i}{\partial u^\varphi} 2\delta_i\mu(w^\alpha) \right) \quad (i = 1, \ldots, q; \ \alpha = 1, \ldots, r; \ k = 1, \ldots, m)
\]
has rank \( q \) on \( \hat{\mathcal{A}} \). By Note 6.1 and Theorem 4.1 there exist \( \lambda_0 \geq 0, \ p \in X \) and \( \mu \in U_q \) continuous on each interval of continuity of \( u_0 \), not vanishing simultaneously on \( T \), such that if
\[
\hat{H}(t, x, u, w) := \langle p(t), f(t, x, u, w) \rangle - \lambda_0 \hat{L}(t, x, u, w) - \langle \mu(t), \psi(t, x, u, w) \rangle,
\]
then
\[
a'. \quad \dot{p}(t) = -\hat{H}_x^*(\tilde{x}_0(t), w_0(t)) \quad \text{and} \quad \hat{H}_{\alpha\varphi}(\tilde{x}_0(t), w_0(t)) = 0 \quad \text{on each interval of continuity of } (u_0, w_0).
\]
\[
b'. \quad \hat{H}(t, x_0(t), u, w) \leq \hat{H}(\tilde{x}_0(t), w_0(t)) \quad \text{for all } (t, u, w) \in T \times \mathbb{R}^m \times \mathbb{R}^r \text{ such that } (t, x_0(t), u, w) \in \hat{\mathcal{A}}.
\]

Observe now that, if
\[
G(t, x, u) := H(t, x, u, p(t), \mu(t), \lambda_0) - \langle p(t), f(t, x, u) \rangle - \lambda_0 L(t, x, u) - \langle \mu(t), \varphi(t, x, u) \rangle,
\]
then \( \hat{H}(t, x, u, w) = G(t, x, u) - \sum_1^r \mu_\alpha(t)(w^\alpha)^2 \). Therefore (a) holds. Also
\[
\hat{H}_{\alpha\varphi}(\tilde{x}_0(t), w_0(t)) = -2\mu_\alpha(t)w_0^\alpha(t) = 0.
\]
Hence, \( \mu_a(t) = 0 \) whenever \( w_0^a(t) > 0 \), and hence, whenever \( \varphi_a(\tilde{x}_0(t)) < 0 \). Now, since \( \psi(t, x_0(t), u, w) = 0 \) for any \((t, x_0(t), u, w) \in \hat{A}\),

\[
H(t, x_0(t), u, p(t), 0, \lambda_0) = \dot{H}(t, x_0(t), u, w) \leq \dot{H}(\tilde{x}_0(t), w_0(t)) = H(\tilde{x}_0(t), p(t), 0, \lambda_0)
\]

for all \((t, u) \in T \times \mathbb{R}^m\) with \((t, x_0(t), u) \in A\), and condition (b) holds.

To prove that \( \mu_a(t) \geq 0 \) \((\alpha \in R)\), let \( g(t, u) := \varphi(t, x_0(t), u) \) and \( h(t, u) := -H(t, x_0(t), u, p(t), 0, \lambda_0) \) for all \((t, u) \in T \times \mathbb{R}^m\) and set

\[
C := \{(t, u) \in T \times \mathbb{R}^m \mid g_a(t, u) \leq 0 \ (\alpha \in R), \ g_\beta(t, u) = 0 \ (\beta \in Q)\}.
\]

Since \( h(t, u_0(t)) \leq h(t, u) \) for all \((t, u) \in C\), by Lemma 3.5 there exists a unique \( v \in U_g\) such that, if we set

\[
F(t, u) := h(t, u) + \langle v(t), g(t, u) \rangle,
\]

then \( F_a(t, u_0(t)) = 0 \). Moreover, \( v_a(t) \geq 0 \) \((\alpha \in R\), \( t \in T\)) and \( v_a(t) = 0 \) whenever \( g_a(t, u_0(t)) < 0 \).

Since, on the other hand,

\[
0 = -H_a(\tilde{x}_0(t), p(t), \mu(t), \lambda_0) = h_a(t, u_0(t)) + \mu^*(t)g_a(t, u_0(t)),
\]

it follows by uniqueness that \( \mu = v \).

An alternate proof can be made as follows. By Corollary 4.2 we have

\[
\langle h, G_{uu}(\tilde{x}_0(t))h \rangle - 2 \sum_{\alpha=1}^{r} \mu_a(t)(\eta^\alpha)^2 \leq 0
\]

for all \((h, \eta) \in \mathbb{R}^m \times \mathbb{R}^r\) satisfying

\[
\sum_{k=1}^{m} \frac{\partial \varphi_a(\tilde{x}_0(t))}{\partial u_k} h_k + 2w^a_0(t)\eta^a = 0 \ (\alpha \in R), \quad \sum_{k=1}^{m} \frac{\partial \varphi_\beta(\tilde{x}_0(t))}{\partial u_k} h_k = 0 \ (\beta \in Q).
\]

At a point \( t \) at which \( \varphi_a(\tilde{x}_0(t)) = -(w^a_0(t))^2 = 0 \), the variable \( \eta^a \) is arbitrary. The above inequality therefore can hold at such a point if and only if \( \mu_a(t) \geq 0 \).

\( \Box \)

**Note 6.3** Observe that (b) can be put in the alternate form

\[
H(t, x_0(t), u, p(t), \mu(t), \lambda_0) + \langle \mu(t), \varphi(t, x_0(t), u) \rangle \leq H(\tilde{x}_0(t), p(t), \mu(t), \lambda_0)
\]

for all \((t, u) \in T \times \mathbb{R}^m\) with \((t, x_0(t), u) \in A\). Also, the statement that \( \mu_a(t) = 0 \) whenever \( \varphi_a(\tilde{x}_0(t)) < 0 \) is equivalent to the relation \( \mu_a(t)\varphi_a(\tilde{x}_0(t)) = 0 \) \((\alpha \in R)\).

Let us now show that, as in Note 4.3, the function \( \mu \) appearing in Theorem 6.2 can be expressed in terms of the other functions involved.

**Note 6.4** Suppose that, for some \((x, u) \in Z(A), \ p \in X, \ \lambda \in \mathbb{R}\) and a function \( \mu : T \rightarrow \mathbb{R}^q\) satisfying \( \mu_a(t) = 0 \) whenever \( \varphi_a(\tilde{x}(t)) < 0 \), we have

\[
0 = H_u(\tilde{x}(t), p(t), \mu(t), \lambda) \quad [= p^*(t)f_u(\tilde{x}(t)) - \lambda L_u(\tilde{x}(t)) - \mu^*(t)\varphi_u(\tilde{x}(t))] \quad (t \in T).
\]

Let \( \hat{\phi} = (\varphi_{i_1}, \ldots, \varphi_{i_p}) \) with \( i_1, \ldots, i_p \) the indices \( i \in R \cup Q \) such that \( \varphi_i(\tilde{x}(t)) = 0 \), and let \( \hat{\mu} = (\mu_{i_1}, \ldots, \mu_{i_p}) \). Then \( \mu \in U_q \) and

\[
\hat{\mu}^*(t) = \left[ p^*(t)f_u(\tilde{x}(t)) - \lambda L_u(\tilde{x}(t)) \right] \hat{\phi}_u(\tilde{x}(t))^{-1} \quad (t \in T).
\]
DEFINITION 6.5

• Admissible multipliers
  For all \((x, u) \in Z(A)\) let \(\mathcal{V}(x, u) := \{\mu \in \mathcal{U}_q \mid \mu_a(t) = 0\ \text{whenever}\ \varphi_a(\tilde{x}(t)) < 0\}\).

• Admissible variations
  Given \((x, u) \in Z(A)\) let \(A(t) := f_x(\tilde{x}(t)), B(t) := f_u(\tilde{x}(t))\) \((t \in T)\), let \(\hat{\phi} = (\varphi_{i_1}, \ldots, \varphi_{i_p})\) with \(i_1, \ldots, i_p\) the indices \(i \in R \cup Q\) such that \(\varphi_i(\tilde{x}(t)) = 0\), and define the set \(Y_2(x, u)\) of \(A\)-admissible variations along \((x, u)\) as the set of all \((y, v) \in Z\) satisfying
    i. \(y(t_0) = y(t_1) = 0\).
    ii. \(\dot{y}(t) = A(t)y(t) + B(t)v(t)\) \((t \in T)\).
    iii. \(\hat{\phi}_x(\tilde{x}(t))y(t) + \hat{\phi}_u(\tilde{x}(t))v(t) = 0\) \((t \in T)\).

• The sets \(M_2, E_2\) and \(H_2\)
  For all \((x, u) \in Z(A)\) let
  \[
  M_2(x, u) := \{(\lambda_0, p, \mu) \in \mathbb{R} \times X \times \mathcal{V}(x, u) \mid \lambda_0 \geq 0, \mu_a(t) \geq 0, \lambda_0 + |p| \neq 0, \\
  \dot{p}(t) = -H_x^*(t, \lambda_0), H_u(t, \lambda_0) = 0 \quad (t \in T)\},
  \]
  where \(H(t, \lambda_0)\) denotes \(H(\tilde{x}(t), p(t), \mu(t), \lambda_0)\), and consider the following sets:
  \[
  E_2 := \{(x, u, p, \mu) \in Z \times X \times \mathcal{U}_q \mid (x, u) \in D(f) \text{ and } (1, p, \mu) \in M_2(x, u)\},
  \]
  \[
  H_2 := \{(x, u, p, \mu) \in Z \times X \times \mathcal{U}_q \mid J_2((x, u); (y, v)) \geq 0 \text{ for all } (y, v) \in Y_2(x, u)\},
  \]
  where
  \[
  J_2((x, u); (y, v)) = \int_{t_0}^{t_1} 2\Omega_2(t, y(t), v(t)) dt \quad ((y, v) \in Z)
  \]
  and, for all \((t, y, v) \in T \times \mathbb{R}^n \times \mathbb{R}^m\),
  \[
  2\Omega_2(t, y, v) := -[\langle y, H_{xx}(t, 1)y \rangle + 2\langle y, H_{xu}(t, 1)v \rangle + \langle v, H_{uu}(t, 1)v \rangle],
  \]
  where \(H(t, 1)\) denotes \(H(\tilde{x}(t), p(t), \mu(t), 1)\).

• Normality
  A process \((x, u) \in Z(A)\) will be said to be normal to \(P(A, f, I)\) if, given \((p, \mu) \in X \times \mathcal{V}(x, u)\) such that, for all \(t \in T\),
  \[
  \dot{p}(t) = -A^*(t)p(t) + \varphi_u^*(\tilde{x}(t))\mu(t) \quad [= -H_u^*(\tilde{x}(t), p(t), \mu(t), 0)],
  \]
  \[
  0 = B^*(t)p(t) - \varphi_u^*(\tilde{x}(t))\mu(t) \quad [= H_u^*(\tilde{x}(t), p(t), \mu(t), 0)],
  \]
  then \(p \equiv 0\). In this event, by Note 6.4, also \(\mu \equiv 0\), and clearly \((\lambda_0, p, \mu) \in M_2(x, u) \Rightarrow \lambda_0 \neq 0\).

REMARK 6.6 Let \((x, u) \in Z(A)\) and suppose that, given \((p, \mu) \in X \times \mathcal{V}(x, u)\) such that, for all \(t \in T\),
  \[
  \dot{p}(t) = -A^*(t)p(t) + \varphi_u^*(\tilde{x}(t))\mu(t)
  \]
  \[
  p^*(t)B(t)h = 0, \quad \text{for all } h \in \mathbb{R}^m \text{ satisfying } \varphi_u(\tilde{x}(t))h = 0,
  \]
  necessarily \(p \equiv 0\). Then \((x, u)\) is normal to \(P(A, f, I)\).
Note that, with the notation of Note 6.4, if \( \mu \in \mathcal{V}(x, u) \), then

\[
\varphi^*_a(\tilde{x}(t))\mu(t) = \tilde{\varphi}^*_a(\tilde{x}(t))\hat{\mu}(t) \quad \text{and} \quad \varphi^*_p(\tilde{x}(t))\mu(t) = \tilde{\varphi}^*_p(\tilde{x}(t))\hat{\mu}(t).
\]

Thus, in the definition of normality and in Remark 6.6 we can replace \( \varphi \) and \( \mu \) with \( \tilde{\varphi} \) and \( \hat{\mu} \), respectively.

**Theorem 6.7** Suppose \((x, u)\) solves \(P(\mathcal{A}, f, I)\). Then \(M_2(x, u) \neq \emptyset\). If \((x, u)\) is normal to \(P(\mathcal{A}, f, I)\), then there exists a unique \((p, \mu) \in X \times \mathcal{U}_q\) such that \((x, u, p, \mu) \in \mathcal{E}_2\). Moreover, \((x, u, p, \mu) \in \mathcal{H}_2\).

**Proof.** Suppose \((x_0, u_0)\) solves \(P(\mathcal{A}, f, I)\). By Theorem 6.2, \(M_2(x_0, u_0) \neq \emptyset\). Let \((\lambda_0, p, \mu) \in M_2(x_0, u_0)\) and suppose \((x_0, u_0)\) is normal to \(P(\mathcal{A}, f, I)\). This implies that \(\lambda_0 \neq 0\) and, if \((\lambda_0, q, v) \in M_2(x_0, u_0)\), then \(r := p - q\) satisfies

\[
\dot{r}(t) = -A^*r(t) + \varphi^*_a(\tilde{x}_0(t))[\mu(t) - v(t)],
\]

\[
0 = B^*r(t) - \varphi^*_p(\tilde{x}_0(t))[\mu(t) - v(t)],
\]

implying that \(p \equiv q\) and \(\mu \equiv v\).

Let \((p, \mu) \in X \times \mathcal{U}_q\) be the unique pair such that \((x_0, u_0, p, \mu) \in \mathcal{E}_2\). Let \(w_0^\alpha(t) := [-\varphi_a(\tilde{x}_0(t))]^{1/2}\) \((\alpha \in \mathbb{R}, t \in T)\), so that \((x_0, u_0, w_0)\) solves \(P(\hat{\mathcal{A}}, \hat{f}, \hat{I})\). Observe that the assumption that \((x_0, u_0)\) is normal to \(P(\mathcal{A}, f, I)\) implies that \((x_0, u_0, w_0)\) is also normal to \(P(\hat{\mathcal{A}}, \hat{f}, \hat{I})\), i.e. if \((q, v) \in X \times \mathcal{U}_q\) is such that

\[
\dot{q}(t) = -\hat{H}_q^*(t, 0) = -H_q^*(\tilde{x}_0(t), q(t), v(t), 0)),
\]

\[
0 = \hat{H}_{(uw)}(t, 0) = (H_u(\tilde{x}_0(t), q(t), v(t), 0), -2w_0^1(t)v_1(t), \ldots, -2w_0^r(t)v_r(t)),
\]

then \((q, v) \equiv (0, 0)\). Here \(\hat{H}(t, 0)\) denotes \(\hat{H}(\tilde{x}_0(t), w_0(t), q(t), v(t), 0)\) and

\[
\hat{H}(t, x, u, w, q, v, \lambda) = \langle q, \hat{f}(t, x, u, w) \rangle - \lambda \hat{L}(t, x, u, w) - \langle v, \psi(t, x, u, w) \rangle = H(t, x, u, q, v, \lambda) - \sum_{\alpha=1}^{r} v_\alpha(w_\alpha^\alpha)^2.
\]

By Theorem 4.7, there exists a unique \((q, v) \in X \times \mathcal{U}_q\) such that, with respect to \(P(\hat{\mathcal{A}}, \hat{f}, \hat{I})\), we have \((x_0, u_0, w_0, q, v) \in \mathcal{E}_1\). Moreover, \((x_0, u_0, w_0, q, v) \in \mathcal{H}_1\).

The first contention means that \((x_0, u_0, w_0) \in D(\hat{f})\) and

\[
\dot{q}(t) = -\hat{H}_q^*(t, 1) = -H_q^*(t, 1),
\]

\[
0 = \hat{H}_{(uw)}(t, 1) = (H_u(1, -2w_0^1(t)v_1(t), \ldots, -2w_0^r(t)v_r(t)),
\]

where \(\hat{H}(t, 1)\) denotes \(\hat{H}(\tilde{x}_0(t), w_0(t), q(t), v(t), 1)\), and \(H(t, 1)\) denotes \(H(\tilde{x}_0(t), q(t), v(t), 1)\). Since also \((x_0, u_0, w_0, p, \mu) \in \mathcal{E}_1\), it follows that \((p, \mu) \equiv (q, v)\).

Now, the fact that \((x_0, u_0, w_0, p, \mu) \in \mathcal{H}_1\) with respect to \(P(\hat{\mathcal{A}}, \hat{f}, \hat{I})\) means that, for any \((y, v, z) \in Z \times \mathcal{U}_z\) satisfying

i. \(y(t_0) = y(t_1) = 0\),

ii. \(y(t) = \hat{f}(\tilde{x}_0(t), w_0(t))y(t) + \hat{f}_{(uw)}(\tilde{x}_0(t), w_0(t))(v(t), z(t))^*(t \in T)\),

iii. \(\psi(\tilde{x}_0(t), w_0(t))y(t) + \psi_{(uw)}(\tilde{x}_0(t), w_0(t))(v(t), z(t))^* = 0\ (t \in T)\),
where we have
\[
\hat{J}((x_0, u_0, w_0); (y, v, z)) := \int_{t_0}^{t_1} 2\tilde{\Omega}(t, y(t), v(t), z(t))dt \geq 0,
\]
where
\[
2\tilde{\Omega}(t, y, v, z) = -[\langle y, \hat{H}_{xx}(t, 1)y \rangle + 2\langle y, \hat{H}_{xw}(t, 1)v(z) \rangle^* + \langle v, z \rangle^*, \hat{H}_{uw}(t, 1)v(z) \rangle^*].
\]
Note that condition (ii) corresponds to \(\dot{y}(t) = A(t)y(t) + B(t)v(t) (t \in T)\), and condition (iii) to
\[
\varphi_x(\tilde{x}_0(t))y(t) + \varphi_u(\tilde{x}_0(t))v(t) + 2(w_0^1(t)z_1(t), \ldots, w_0^r(t)z_r(t), 0, \ldots, 0)^* = 0.
\]
Also, as one readily verifies,
\[
2\tilde{\Omega}(t, y, v, z) = 2\Omega_2(t, y, v) + 2\sum_{i=1}^r \mu_{\alpha}(t)z_{\alpha}^2(t).
\]
Now, let \((y, v) \in Y_2(x_0, u_0)\) so that \(y(t_0) = y(t_1) = 0, \dot{y}(t) = A(t)y(t) + B(t)v(t) (t \in T)\), and
\[
\hat{\varphi}_x(\tilde{x}_0(t))y(t) + \hat{\varphi}_u(\tilde{x}_0(t))v(t) = 0 \quad (t \in T),
\]
where \(\hat{\varphi} = (\varphi_{i_1}, \ldots, \varphi_{i_p})\) with \(i_1, \ldots, i_p\) the indices \(i \in R \cup Q\) such that \(\varphi_i(\tilde{x}_0(t)) = 0\). For all \(\alpha \in R\) define \(z_{\alpha}(t) = 0\) if \(\varphi_{\alpha}(\tilde{x}_0(t)) = 0\) and
\[
z_{\alpha}(t) = -[2\omega_0^\alpha(t)]^{-\frac{1}{2}} \left[ \frac{\partial \varphi_{\alpha}}{\partial x}(\tilde{x}_0(t))y(t) + \frac{\partial \varphi_{\alpha}}{\partial u}(\tilde{x}_0(t))v(t) \right] \quad \text{(if } \varphi_{\alpha}(\tilde{x}_0(t)) < 0).\]
In this event we have, for all \(\alpha \in R\),
\[
\frac{\partial \varphi_{\alpha}}{\partial x}(\tilde{x}_0(t))y(t) + \frac{\partial \varphi_{\alpha}}{\partial u}(\tilde{x}_0(t))v(t) + 2\omega_0^\alpha(t)z_{\alpha}(t) = 0
\]
and, for all \(\beta \in Q\),
\[
\frac{\partial \varphi_{\beta}}{\partial x}(\tilde{x}_0(t))y(t) + \frac{\partial \varphi_{\beta}}{\partial u}(\tilde{x}_0(t))v(t) = 0.
\]
Therefore, \((y, v, z)\) satisfies conditions (i)–(iii) and we conclude that
\[
J_2((x_0, u_0); (y, v)) = \hat{J}((x_0, u_0, w_0); (y, v, z)) \geq 0,
\]
showing that \((x_0, u_0, p, \mu) \in H_2\).

Let us end with a simple example which illustrates the main result of the paper.

**Example 6.8** Consider the problem of minimizing
\[
I(x, u) = \frac{1}{2} \int_0^8 [2[x_1(t) + t(u_1(t) - u_2(t))] + t[u_2^2(t) - u_1^2(t)]]dt,
\]
where \(x = (x_1, x_2)^*\) and \(u = (u_1, u_2)^*\), subject to \(\dot{x}(t) = u(t), x(0) = x(8) = 0, -x_2(t) - u_1(t) \leq 0\).
In this event we have \(T = [0, 8]\),
\[
L(t, x, u) = x_1 + t(u_1 - u_2) + t(u_2^2 - u_1^2)/2, \quad f(t, x, u) = u, \quad \varphi(t, x, u) = -x_2 - u_1.
\]
Observe first that, in the sense of Definition 6.5, any \((x, u) \in Z\) is normal to the problem. Indeed, since 
\(f_x(\tilde{x}(t)) = 0, f_u(\tilde{x}(t)) = I, \varphi_x(\tilde{x}(t)) = (0, -1)\) and 
\(\varphi_u(\tilde{x}(t)) = (-1, 0)\), if \((p, \mu) \in X \times \mathcal{V}(x, u)\) is such that
\[
\begin{pmatrix}
\dot{p}_1(t) \\
\dot{p}_2(t)
\end{pmatrix} =
\begin{pmatrix}
0 \\
-\mu(t)
\end{pmatrix}, \quad \begin{pmatrix}
0 \\
p_1(t) + \mu(t)
\end{pmatrix} = \begin{pmatrix}
0 \\
p_2(t)
\end{pmatrix}
\]
then necessarily \(p \equiv 0\). On the other hand, the process \((x_0, u_0) \equiv (0, 0)\) is such that, for a (unique) \((p, \mu) \in X \times \mathcal{U}_1\), \((x_0, u_0, p, \mu)\) belongs to \(E_2\). To prove it, note first that \(H(t, x, u, p, \mu, 1)\) is given by
\[
p_1u_1 + p_2u_2 - x_1 - t(u_1 - u_2) - t(u_2^2 - u_1^2)/2 + \mu(x_2 + u_1)
\]
so that
\[
H_x(\tilde{x}_0(t), p(t), \mu(t), 1) = (-1, \mu(t)),
\]
\[
H_u(\tilde{x}_0(t), p(t), \mu(t), 1) = (p_1(t) - t + \mu(t), p_2(t) + t).
\]
Thus, if \(H(t, 1)\) denotes \(H(\tilde{x}_0(t), p(t), \mu(t), 1)\), the relations
\[
\dot{p}(t) = \begin{pmatrix}
\dot{p}_1(t) \\
\dot{p}_2(t)
\end{pmatrix} = \begin{pmatrix}
1 \\
\mu(t)
\end{pmatrix} = -H_x(\tilde{x}_0(t), 1), \quad \begin{pmatrix}
0 \\
0
\end{pmatrix} = \begin{pmatrix}
p_1(t) - t + \mu(t) \\
p_2(t) + t
\end{pmatrix} = H_u^*(t, 1)
\]
are satisfied with \((p_1(t), p_2(t)) = (t - 1, -t)\) and \(\mu \equiv 1\).

Therefore, \((x_0, u_0)\) belongs to \(Z_c(A, f)\), it is normal (see Definition 6.5) and, for some \((p, \mu) \in X \times \mathcal{U}_1\), \((x_0, u_0, p, \mu)\) belongs to \(E_2\). However, by an application of Theorem 6.7, it can be shown that \((x_0, u_0)\) is not a solution to the problem.

To prove this claim observe that the set \(Y_2(x_0, u_0)\) of \(A\)-admissible variations along \((x_0, u_0)\) is given by those \((y, v) \in Z\) satisfying
\begin{enumerate}
  \item \(y(0) = y(8) = 0\);
  \item \(\dot{y}(t) = v(t) (t \in T)\);
  \item \(-y_2(t) - v_1(t) = 0 (t \in T)\).
\end{enumerate}
Define \((y, v) \in Z\) with \(y = (y_1, y_2)^*\) and \(v = (v_1, v_2)^*\) by
\[
y_1(t) := \begin{cases}
-t^2/2, & \text{if } t \in [0, 2], \\
4 - 4t + t^2/2, & \text{if } t \in [2, 6], \\
-32 + 8t - t^2/2, & \text{if } t \in [6, 8],
\end{cases}
\]
\[
y_2(t) := \begin{cases}
t, & \text{if } t \in [0, 2], \\
4 - t, & \text{if } t \in [2, 6], \\
t - 8, & \text{if } t \in [6, 8],
\end{cases}
\]
\[
v_1(t) := -y_2(t) \quad \text{and} \quad v_2(t) := y_2(t) (t \in [0, 8]).
\]
As one readily verifies, the conditions defining membership of \(Y_2(x_0, u_0)\) hold for this particular process \((y, v)\). Moreover,
\[
2\Omega_2(t, y(t), v(t)) = \langle v(t), L_{uu}(\tilde{x}(t))v(t) \rangle = t\{v_2^2(t) - v_1^2(t)\}
\]
and, therefore,
\[
J_2((x_0, u_0); (y, v)) = \int_0^8 t\{v_2^2(t) - v_1^2(t)\}dt
\]
\[
= \int_0^2 t(1 - t^2)dt + \int_2^6 t(1 - (4 - t)^2)dt + \int_6^8 t(1 - (t - 8)^2)dt = -32/3 < 0.
\]
We conclude that \((x_0, u_0, p, \mu)\) does not belong to \(\mathcal{H}_2\) and therefore, by Theorem 6.7, \((x_0, u_0)\) is not a solution to the problem.

**References**


