Regularity and conjugacy for constrained variational problems

JAVIER F ROSENBLUETH
IIMAS, Universidad Nacional Autónoma de México
Apartado Postal 20-126, CDMX 01000
MEXICO
jfrl@unam.mx

Abstract: For problems in the calculus of variations involving equality and inequality mixed constraints we characterize, in terms of an extended notion of conjugate points, the sign of a quadratic form which corresponds to the second variation of the integral to be minimized. Second order necessary conditions are then derived assuming the well-known constraint qualification of regularity in the sense that, with respect to the set of mixed constraints, both the tangent cone and the set of tangential constraints coincide.

Key–Words: Calculus of variations, regularity, conjugate points, tangent cone, tangential constraints


1 Introduction

The notion of conjugate points for constrained variational problems has played a fundamental role in the derivation of necessary and sufficient optimality conditions. In particular, for the simple fixed-endpoint problem in the calculus of variations, Jacobi’s necessary condition states that there are no conjugate points to an endpoint in the underlying open time interval. The corresponding sufficient condition, a slight strengthening of the previous one, asks for the nonexistence of conjugate points in the open time interval and the other endpoint. In both cases, as it is well-known, a crucial assumption is the nonsingularity or the satisfaction of the strengthened condition of Legendre of the trajectory under consideration. For a full account of this theory we refer the reader to the classical book by Hestenes [9].

Several attempts to extend the definition of conjugacy to problems with singular trajectories have been made. In particular, some of the references found in the literature include [3, 11–14] where the definitions proposed provide different approaches in order to consider such trajectories and still derive necessary and sufficient conditions, as well as to generalize the classical notion to certain classes of optimal control problems.

The approach given in [3] yields a definition of extended conjugate points which characterizes the nonnegativity of the second variation with respect to the integral to be minimized even if the trajectory under consideration does not satisfy the strengthened condition of Legendre. The problem studied in [3] is the simple fixed-endpoint in the calculus of variations and, even for such well-known problem, the results extend the applicability of the classical theory since no nonsingularity assumptions are required. We refer the reader to [5] and references therein where, instead of optimality conditions without classical nonsingularity assumptions, the authors treat existence of solutions to the autonomous Lagrange optimal control problem without classical convexity assumptions.

In this paper we generalize the definition given in [3] to problems in calculus of variations which include equality and inequality mixed constraints (see also [4], where necessary conditions of optimality are derived for state constrained problems). As we shall show, the emptiness of the new set of extended conjugate points, as defined in this paper, is equivalent to the nonnegativity of a quadratic form for all trajectories belonging to a convex cone. On the other hand, second order conditions can be derived in terms of the tangent cone of a subset of the set of mixed constraints with an appropriate norm. This yields, under a regularity assumption (in the sense that the tangent cone and the set of tangential constraints coincide), second order necessary conditions in terms of the new notion of extended conjugate points.

2 The simple fixed-endpoint problem

In order to clearly situate the contribution of this paper and compare our main result with the classical setting, let us give a brief explanation of conjugacy for the simple fixed-endpoint problem in the calculus of variations.

Suppose we are given an interval $T := [t_0, t_1]$ in $\mathbb{R}$, two points $\xi_0, \xi_1$ in $\mathbb{R}^n$, and a function $L$ mapping $T \times \mathbb{R}^n \times \mathbb{R}^n$ to $\mathbb{R}$. Denote by $X$ the space of
piecewise $C^1$ functions mapping $T$ to $\mathbb{R}^n$, set

$$S := \{ x \in X \mid x(t_0) = \xi_0, \ x(t_1) = \xi_1 \},$$

and consider the functional $I: X \to \mathbb{R}$ given by

$$I(x) := \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))dt \quad (x \in X).$$

The (unconstrained) classical fixed-endpoint problem, which we label (P0), is that of minimizing $I$ over $S$.

Elements of $X$ will be called arcs or trajectories, and they are admissible if they belong to $S$. An arc $x$ solves (P0) if it is admissible and $I(x) \leq I(y)$ for all admissible $y$. Given $x \in X$, we shall use the notation $(\tilde{x}(t))$ to represent $(t, x(t), \dot{x}(t))$ and assume that $L$ is $C^2$.

For all $x \in X$ consider the second variation of $I$ along $x$ given by

$$J(x; y) := \int_{t_0}^{t_1} 2\Omega(t, y(t), \dot{y}(t))dt \quad (y \in X)$$

where, for all $(t, y, \dot{y}) \in T \times \mathbb{R}^n \times \mathbb{R}^n$,

$$2\Omega(t, y, \dot{y}) := \langle y, L_{xx}(\tilde{x}(t))y \rangle + 2\langle y, L_{x\dot{x}}(\tilde{x}(t))\dot{y} \rangle + \langle \dot{y}, L_{\dot{x}\dot{x}}(\tilde{x}(t))\dot{y} \rangle.$$ 

Denote by $\mathcal{H}$ the set of arcs $x$ for which $J(x; y) \geq 0$ for all $y \in X$, where

$$Y = \{ y \in X \mid y(t_0) = y(t_1) = 0 \}$$

is the set of admissible variations. It is well-known (see, for example, [9]) that, if $x$ solves (P0) then $x$ belongs to $\mathcal{H}$.

The set $\mathcal{H}$ therefore plays a fundamental role in the set of necessary conditions, but it might be difficult to check if a certain trajectory belongs to it. Jacobi’s theory helps to solve this issue.

To explain it, let us introduce the following notation. For all $s \in [t_0, t_1]$ set $T_s := \{ \tilde{t}, \tilde{s} \}$, let $X_s$ be the space of piecewise $C^1$ functions mapping $T_s$ to $\mathbb{R}^n$ and denote by $Y_s$ the set of trajectories $y \in X_s$ for which $y(t_0) = y(s) = 0$.

Whenever we are given $x \in X$ and $y \in Y_s$, we shall consider the functions $v, w: T_s \to \mathbb{R}^n$ (depending on both $x$ and $y$) defined by

$$v(t) := \Omega_2(\dot{y}(t))$$
$$= L_{xx}(\tilde{x}(t))y(t) + L_{x\dot{x}}(\tilde{x}(t))\dot{y}(t)$$
$$w(t) := \Omega_2(\dot{y}(t))$$
$$= L_{xx}(\tilde{x}(t))y(t) + L_{x\dot{x}}(\tilde{x}(t))\dot{y}(t).$$

**Definition 1** For $x \in X$ denote by $C(x)$ the set of points $s \in (t_0, t_1)$ for which there exists $y \in Y_s$ with $y \neq 0$ such that $\dot{v}(t) = w(t)$ ($t \in T_s$).

Elements of $C(x)$ are called points conjugate to $t_0$ on $x$ and Jacobi’s necessary condition, relating this set with the nonnegativity of $J(x; y)$ for all $y \in Y$, can be stated as follows.

**Theorem 2** Let $x$ be an admissible $C^1$ trajectory satisfying Legendre’s strengthened condition, that is, $L_{x\dot{x}}(\tilde{x}(t)) > 0$ ($t \in T$). If $x \in \mathcal{H}$, then $C(x) \cap (t_0, t_1) = \emptyset$.

Let us point out that, as it is also well-known, the assumption of nonsingularity of $x$ in the sense that $|L_{x\dot{x}}(\tilde{x}(t))| \neq 0$ for all $t \in T$ is essential in the theorem.

In the next section we shall consider the same fixed-endpoint problem except for a new element which, for second order conditions, makes the problem much more difficult to analyse. Our problem will involve equality and inequality mixed constraints.

### 3 Statement of the problem

Let us state the problem we shall be concerned with, a fixed-endpoint problem in the calculus of variations posed over piecewise $C^1$ functions involving equality and inequality mixed constraints.

It is important to mention that, in contrast with constrained optimization problems in the finite dimensional case, the type of constraints we shall deal with, as explained in [6, p 335], “make the problem much more complex than the mathematical programming problem, or even the isoperimetric problem. In part, this is because we now have infinitely many constraints, one for each $t$.” For the finite dimensional case, we refer the reader to [10] and, for the isoperimetric problem in the calculus of variations, we refer to [1, 2].

Suppose the data are as before, but we are also given a function $\varphi$ mapping $T \times \mathbb{R}^n \times \mathbb{R}^n$ to $\mathbb{R}^d$. Set now

$$D := \{ x \in X \mid x(t_0) = \xi_0, \ x(t_1) = \xi_1 \},$$

$$S := \{ x \in D \mid (t, x(t), \dot{x}(t)) \in U \ (t \in T) \}$$

(not to be confused with the previous set $S$), where the set $U$ of constraints, involving equalities and inequalities, is given by

$$U := \{ (t, x, \dot{x}) \in T \times \mathbb{R}^n \times \mathbb{R}^n \mid \varphi_\alpha(t, x, \dot{x}) \leq 0, \ \varphi_\beta(t, x, \dot{x}) = 0 \ (\alpha \in R, \ \beta \in Q) \},$$

$$R = \{ 1, \ldots, r \} \text{ and } Q = \{ r + 1, \ldots, q \}.$$ As before, consider the functional $I: X \to \mathbb{R}$ given by

$$I(x) := \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))dt \quad (x \in X).$$
The problem we shall deal with, which we label (P1), is that of minimizing $I$ over $S$. A common and concise way of formulating this problem is as follows:

Minimize

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

subject to $x \in X$ and

$$x(t_0) = \xi_0, \quad x(t_1) = \xi_1;$$
$$\varphi_\alpha(t, x(t), x(t)) \leq 0, \quad \varphi_\beta(t, x(t), x(t)) = 0 \quad \left(\alpha \in R, \quad \beta \in Q, \quad t \in T\right).$$

We assume that $L$ and $\varphi$ are $C^2$ and the $q \times (n + r)$-dimensional matrix

$$(\frac{\partial \varphi_i}{\partial x^k}, \delta_{\alpha i} \varphi_\alpha)$$

$(i = 1, \ldots, q; \alpha = 1, \ldots, r; k = 1, \ldots, n)$

has rank $q$ on $U$. This is equivalent to the condition that, at each point $(t, x, \dot{x})$ in $U$, the matrix

$$(\frac{\partial \varphi_i}{\partial x^k})$$

$(i = 1, \ldots, q; k = 1, \ldots, n)$

has rank $p$, where $i_1, \ldots, i_p$ are the indices $i$ in $\{1, \ldots, q\}$ such that $\varphi_i(t, x, \dot{x}) = 0$ (see [8]).

4 Necessary conditions

First order conditions for this problem are well-known, and a Hamiltonian formulation (see, for example, [9], p 254) yields the following result.

For all $(t, x, \dot{x}, p, \mu, \lambda) \in T \times R^n \times R^n \times R^n \times R^q \times R$

let

$$H(t, x, \dot{x}, p, \mu, \lambda) := \dot{p} - \lambda L(t, x, \dot{x}) - (\sigma, \mu(t, \dot{x}))$$

and denote by $U_q$ the space of piecewise continuous functions mapping $T$ to $R^q$.

**Theorem 3** If $x_0$ solves (P1), \( \exists \lambda_0 \geq 0, p \in X, \) and $\mu \in U_q$ continuous such that $x_0$ not vanishing simultaneously on $T$, such that

a. $\mu_\alpha(t) \geq 0$ and $\mu_\alpha(t) \varphi_\alpha(\dot{x}_0(t)) = 0$ ($\alpha \in R, \ t \in T$);

b. $\dot{p}(t) = -H^*_x(\dot{x}_0(t), p(t), \mu(t), \lambda_0)$ on every interval of continuity of $\dot{x}_0$;

c. $H^*_x(\dot{x}_0(t), p(t), \mu(t), \lambda_0) = 0$.

Let us denote by $E$ the set of trajectories and their corresponding multipliers for which the conditions of Theorem 3 hold, and having a cost multiplier $\lambda_0$ equal to 1.

**Definition 4** Denote by $E$ the set of all $(x, p, \mu) \in S \times X \times U_q$ such that

i. $\mu_\alpha(t) \geq 0$ and $\mu_\alpha(t) \varphi_\alpha(\dot{x}(t)) = 0$ ($\alpha \in R, \ t \in T$);

ii. $\dot{p}(t) = L^*_x(\dot{x}(t)) + \varphi^*_x(\dot{x}(t))\mu(t)$

$$\left.\begin{array}{l}
= -H^*_x(\dot{x}(t), p(t), \mu(t), 1) \quad (t \in T);
\end{array}\right\}$$

iii. $0 = p(t) - L^*_x(\dot{x}(t)) - \varphi^*_x(\dot{x}(t))\mu(t)$

$$\left.\begin{array}{l}
= H^*_x(\dot{x}(t), p(t), \mu(t), 1) \quad (t \in T).
\end{array}\right\}$$

For second order conditions define, for any $(x, p, \mu) \in X \times X \times U_q$,

$$J(x, p, \mu; y) := \int_{t_0}^{t_1} 2\omega(t, y(t), y(t)) dt \quad (y \in X)$$

where, for all $(t, y, \dot{y}) \in T \times R^n \times R^n$,

$$2\omega(t, \dot{y}) := -\langle [y, H_{xx}(t)y] + 2[y, H_{xy}(t)\dot{y}] + \langle \dot{y}, H_{xx}(t)\dot{y} \rangle \rangle$$

and $H(t)$ denotes $H(\dot{x}(t), p(t), \mu(t), 1)$.

Second order conditions can be expressed in terms of tangent cones and regularity. To introduce these concepts, let us endow the space $X$ with the norm

$$\|x\| := \sup_{t \in T} \{|x(t)|^2 + |\dot{x}(t)|^2\}^{1/2} \quad (x \in X).$$

**Definition 5** A sequence $\{x_q\} \subset X$ converges to $x_0$ in the direction $y$ if $y$ is a unit arc (that is, $\|y\| = 1$), $x_q \neq x_0$, and

$$\lim_{q \to \infty} \|x_q - x_0\| = 0, \quad \lim_{q \to \infty} \|x_q - x_0\| = y.$$ 

Given $x_0 \in S \subset X$, the tangent cone of $S$ at $x_0$, denoted by $\mathcal{T}_S(x_0)$, is the (closed) cone determined by the unit vectors $y$ for which there exists $\{x_q\} \subset S$ converging to $x_0$ in the direction $y$.

Equivalently, $\mathcal{T}_S(x_0)$ is the set of all $y \in X$ for which $\exists \{x_q\} \subset S$ and $\{\epsilon_q > 0\}$ such that

$$\lim_{q \to \infty} \epsilon_q = 0, \quad \lim_{q \to \infty} \frac{x_q - x_0}{\epsilon_q} = y.$$ 

**Definition 6** A trajectory $x \in S$ is said to be regular relative to $S$ if $\mathcal{T}_S(x)$ coincides with the set $\mathcal{R}_S(x)$ of trajectories satisfying the tangential constraints at $x$ with respect to $S$, given by all $y \in Y$ satisfying

$$\varphi^{\alpha}(\dot{x}(t))y(t) + \varphi^{\alpha}(\dot{x}(t))\dot{y}(t) \leq 0$$

$$(\alpha \in I_0(\dot{x}(t)), \ t \in T),$$
\[ \varphi_{\beta x}(\tilde{x}(t))y(t) + \varphi_{\beta x}(	ilde{x}(t))\dot{y}(t) = 0 \]

(\beta \in Q, t \in T)

where

\[ Y = \{ y \in X \mid y(t_0) = y(t_1) = 0 \} \]

is as in Section 2, and

\[ I_a(t, x, \dot{x}) := \{ \alpha \in R \mid \varphi_{\alpha}(t, x, \dot{x}) = 0 \} \]

denotes the set of active indices at \((t, x, \dot{x}) \in U\).

In a recent paper (see [7]) second order conditions were obtained in terms of tangent cones and regularity with respect to the set \( S_1 \) defined below, which coincides with

\[ S_1 = \{ x \in D \mid \varphi_{\alpha}(\tilde{x}(t)) \leq 0 \} \]

\( (\alpha \in R, \mu_{\alpha}(t) = 0, t \in T), \)

\[ \varphi_{\beta x}(\tilde{x}(t)) = 0 \]

(\beta \in R with \( \mu_{\beta}(t) > 0 \), or \( \beta \in Q, t \in T \}).

Note that \( R_{\alpha}(x_0) \) is given by the set of all \( y \in Y \) satisfying

\[ \varphi_{\alpha x}(\tilde{x}_0(t))y(t) + \varphi_{\alpha x}(	ilde{x}_0(t))\dot{y}(t) \leq 0 \]

\( (\alpha \in I_a(\tilde{x}_0(t)), \mu_{\alpha}(t) = 0, t \in T), \)

\[ \varphi_{\beta x}(	ilde{x}_0(t))y(t) + \varphi_{\beta x}(	ilde{x}_0(t))\dot{y}(t) = 0 \]

(\beta \in R with \( \mu_{\beta}(t) > 0 \), or \( \beta \in Q, t \in T \}).

**Theorem 7** Let \( x_0 \in S \) and suppose \( \exists (p, \mu) \) such that \((x_0, p, \mu) \in E\). If \( x_0 \) solves \((P)\) and

\[ S_1 := \{ x \in S \mid \varphi_{\alpha x}(\tilde{x}(t)) = 0 \} \]

\( (\alpha \in R, \mu_{\alpha}(t) > 0, t \in T), \)

then \( J(x_0, p, \mu; y) \geq 0 \) for all \( y \in R_{S_1}(x_0) \). In particular, if \( x_0 \) is regular relative to \( S_1 \), then \( J(x_0, p, \mu; y) \geq 0 \) for all \( y \in R_{S_1}(x_0) \).

5 **Extended conjugate points**

As in the simple fixed-endpoint problem posed in Section 2, we are interested in finding conditions for which \( J(x, p, \mu; y) \) is nonnegative.

Let \( T_s, X_s \) and \( Y_s \) have the same meaning as before and, for any \( B \subset X \), denote by \( Y_s(B) \) the set of all \( y \in Y_s \) such that \( \zeta \), where

\[ \zeta(t) = \begin{cases} y(t) & \text{if } t \in [t_0, s] \\ 0 & \text{if } t \in [s, t_1] \end{cases} \]

For convenience we shall denote \( Y_{t_1}(B) \) as \( Y(B) \) which, clearly, coincides with \( Y \cap B \).

Consider the set

\[ \mathcal{H}(B) := \{ (x, p, \mu) \in X \times X \times U_q \mid J(x, p, \mu; y) \geq 0 \}

for all \( y \in Y(B) \}) \]

where the second variation is nonnegative for arcs belonging to \( Y \cap B \).

- Whenever we are given \((x, p, \mu) \in X \times X \times U_q \) and \( y \in Y_s \), we shall consider the functions \( \sigma, \rho \) mapping \( T_s \) to \( R^n \) defined by

\[ \sigma(t) := -H_{xx}(t)y(t) - H_{xx}(t)\dot{y}(t), \]

\[ \rho(t) := -H_{xx}(t)y(t) - H_{xx}(t)\dot{y}(t) \]

where \( H(t) \) denotes \( H(\tilde{x}(t), p(t), \mu(t), 1) \).

- Given \( s \in [t_0, t_1] \) and \((x, p, \mu) \in X \times X \times U_q \), define the bilinear form \( F_s : X_s \times X_s \to R \) by

\[ F_s(z, y) := \int_{t_0}^{s} \langle \sigma(t), z(t) \rangle + \dot{z}(t), \rho(t) \rangle dt. \]

Let us now define the notion of extended conjugate points which will allow us to characterize the set \( \mathcal{H}(B) \) where the quadratic form \( J \) is nonnegative.

**Definition 8** Let \( B \subset X \). For any \((x, p, \mu) \in X \times X \times U_q \) denote by \( C(B; x, p, \mu) \) the set of points \( s \in (t_0, t_1] \) for which there exists \( y \in Y_s(B) \) such that

i. \( F_s(y, y) \leq 0 \).

ii. There exists \( z \in Y(B) \) such that \( F_s(z, y) < 0 \).

The following result provides a characterization of \( \mathcal{H}(B) \) for any convex cone \( B \) in \( X \).

**Theorem 9** Suppose \( B \subset X \) is a convex cone and \((x, p, \mu) \in X \times X \times U_q \). Then \((x, p, \mu) \in \mathcal{H}(B) \iff C(B; x, p, \mu) = \emptyset \).

**Proof:**

\( \Rightarrow \): Suppose there exists \( s \in C(B; x, p, \mu) \). Let \( y \) and \( z \) be as in Definition 8, and set

\[ \zeta(t) := \begin{cases} y(t) & \text{if } t \in [t_0, s] \\ 0 & \text{if } t \in [s, t_1] \end{cases} \]

Then \( \zeta \) belongs to \( Y(B) \) and, by Definition 8(i),

\[ J(x, p, \mu; z) = \int_{t_0}^{s} 2\omega(t, \zeta(t), \dot{z}(t))dt \]

\[ = \int_{t_0}^{s} 2\omega(t, \zeta(t), \dot{z}(t))dt \]

\[ \leq 0. \]
Set
\[ k := J(x, p, \mu; z), \quad \beta := \mathcal{F}_z(z, y), \]
\[ \alpha := -(\beta + k/2\beta). \]
Note that \( \alpha > 0 \) since \( k \geq 0 \) and \( \beta < 0 \). Therefore
\[ y_\alpha := z + \alpha \zeta \]
belongs to \( Y(B) \) and
\[ J(x, p, \mu; y_\alpha) = \int_{t_0}^{t_1} 2\omega(t, y_\alpha(t), \dot{y}_\alpha(t))dt \]
\[ = k + \alpha^2 J(x, p, \mu; \zeta) + 2\alpha \mathcal{F}_z(z, y) \]
\[ \leq k + 2\alpha \beta \]
\[ = -2\beta^2 < 0. \]

\("\Leftarrow\) Suppose \( (x, p, \mu) \notin \mathcal{H}(B) \). Let \( y \in Y(B) \) be such that \( J(x, p, \mu; y) \leq 0 \) and let \( z \equiv y \). Then \( t_1 \in \mathcal{C}(B; x, p, \mu) \). \[ \Box \]

A combination of this result and Theorem 7 yields the second order necessary conditions for optimality in terms of the new notion of extended conjugate points.

**Theorem 10** Let \( x_0 \in S \) and suppose \( \exists (p, \mu) \) such that \( (x_0, p, \mu) \in \mathcal{E} \). If \( x_0 \) solves \( (P) \) and
\[ S_1 := \{ x \in S \mid \varphi_\alpha(x(t)) = 0 \} \]
\[ (\alpha \in R, \mu_\alpha(t) > 0, t \in T) \],
then \( \mathcal{C}(T_{S_1}(x_0); x_0, p, \mu) = \emptyset \). In particular, if \( x_0 \) is regular relative to \( S_1 \), then \( \mathcal{C}(R_{S_1}(x_0); x_0, p, \mu) = \emptyset \).

6 Conclusion

For a fixed-endpoint problem in the calculus of variations involving equality and inequality mixed constraints, a characterization of the nonnegativity of the second variation along convex cones is derived. It is expressed in terms of an extended notion of conjugate points which does not require the standard assumption of nonsingularity of the extremal under consideration, thus extending significantly the applicability of the classical theory, and it allows to obtain second order necessary conditions under regularity assumptions.

It is of interest to generalize these results under normality assumptions and for more general problems in optimal control.

Acknowledgement. The author is grateful to Dirección General de Asuntos del Personal Académico from Universidad Nacional Autónoma de México for the support given as part of the PASPA program during a sabbatical stay at the Department of Mathematical Sciences, University of Bath, United Kingdom.

References:


