Strong and Weak Augmentability in Calculus of Variations

JAVIER F ROSENBLUETH
National Autonomous University of Mexico
Applied Mathematics and Systems Research Institute
Apartado Postal 20-126, Mexico DF 01000
MEXICO
jfrl@unam.mx

Abstract: In this paper we derive sufficient conditions for local optimality, for the Lagrange problem in the calculus of variations involving mixed equality constraints, by means of the notion of augmented Lagrangians. It is well-known that the standard necessary conditions for that problem can be easily obtained under the assumption of augmentability, instead of the usual one of normality. On the other hand, as we show in this paper, the standard sufficient conditions for a strong (weak) minimum imply strong (weak) augmentability. Since this kind of augmentability implies that the extremal under consideration is a local solution, the results given provide an alternative approach to the classical theory of sufficient conditions.

Key–Words: Augmentability, Lagrange problem, calculus of variations, equality constraints, sufficient conditions

1 Introduction

This paper deals with a control problem of Lagrange which corresponds to the classical fixed-endpoint problem in the calculus of variations involving a set of equality constraints.

The problem consists in finding in a class of arcs \( x \) mapping a fixed interval \([t_0, t_1]\) to \( \mathbb{R}^n \) and satisfying a set of differential equations

\[
\varphi_\alpha(t, x(t), \dot{x}(t)) = 0 \quad (\alpha = 1, \ldots, q)
\]

and end conditions \( x(t_0) = \xi_0, \ x(t_1) = \xi_1 \), one which minimizes the integral

\[
I(x) = \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t)) dt.
\]

Necessary and sufficient conditions for this problem are well established in the literature (see [23, 24] for a detailed explanation).

For necessity, a Lagrangian formulation includes the corresponding conditions of Euler, Legendre, Weierstrass and Jacobi, and standard techniques under the assumption of normality yield the required results. Those techniques are not, however, easily applied. As pointed out by Hestenes [12], “the standard necessary conditions for a minimum for the problem of Lagrange are usually derived under normality (controllability) assumptions by means of a very complicated argument.”

A Hamiltonian formulation, on the other hand, expresses the first-order necessary conditions in terms of a minimum or maximum principle and they are equivalent to the Euler and Weierstrass conditions. Again, the assumption of normality plays a fundamental role in this approach.

It is important to mention that, in both cases, first-order necessary conditions for the optimal control problem of which the classical problem of Lagrange in the calculus of variations can be seen as a particular case have been recently derived in the literature under a very general setting which includes, in particular, equality and inequality mixed constraints, nonsmooth data and weaker assumptions than the standard ones (see [3, 5, 6]).

On the other hand, sufficiency for local optimality is based on a slight strengthening of the necessary conditions and it is usually established by invoking general embedding or field theorems of the theory of differential equations. Those theorems are an integral component of the usual proofs (see, for example, [10]).

Other approaches applicable to optimal control problems include optimal fields, the solution to a certain matrix Riccati equation, an appropriate form of local convexity of the Lagrangian, generalized conjugate points, the Hamilton-Jacobi theory, or even indirect sufficiency proofs (see for example [4, 9, 10, 13, 14, 19, 20] and references therein).

An entirely different approach to derive both necessary and sufficient conditions for problems in optimization theory is that of augmentability. In particular, when dealing with constrained minimum problems in the finite dimensional case, there is one type
of augmentability which yields the derivation of the Lagrange multiplier rules in a much simpler way than under the usual assumption of regularity. Also, the standard sufficient conditions for optimality imply that type of augmentability, and it provides a method of multipliers for finding numerical solutions to constrained minimum problems (see [11, 12]).

In order to clearly understand these statements, let us briefly describe the main ideas for constrained minimum problems in the finite dimensional case. For such problems, one is usually interested in proving the linearity of the gradients of the constraints in order to derive the Lagrange multiplier rules as necessary conditions for optimality.

In other words, if the problem is that of minimizing a function \( f: S \rightarrow \mathbb{R} \) on the set

\[
S = \{ x \in \mathbb{R}^n | g_\alpha(x) = 0 (\alpha \in A) \}
\]

with \( A = 1, \ldots, m \), then the Lagrange multiplier rules state that, for some \( \lambda \in \mathbb{R}^m \),

\[
F'(x_0) = 0, \quad F''(x_0; h) \geq 0
\]

for all \( h \in \mathbb{R}^n \) satisfying \( g_\alpha(x_0; h) = 0 (\alpha \in A) \), where

\[
F(x) = f(x) + \langle \lambda, g(x) \rangle
\]

denotes the standard Lagrangian. These conditions become necessary for a solution to the problem if the linear equations

\[
g'_\alpha(x_0; h) = 0 \quad (\alpha \in A)
\]

in \( h \) are linearly independent. By strengthening the inequality in the second order condition to be strict for all \( h \in \mathbb{R}^n, h \neq 0 \), satisfying \( g'_\alpha(x_0; h) = 0 (\alpha \in A) \), one obtains sufficiency for local minima.

For this kind of problems the approach of augmentability deals with an augmented Lagrangian of the type

\[
H(x) = f(x) + \langle \lambda, g(x) \rangle + \sigma G(x)
\]

where

\[
G(x) = \frac{1}{2} \sum_{\alpha=1}^{m} g_\alpha(x)^2.
\]

The problem is called augmentable at a point \( x_0 \) if \( x_0 \) affords an unconstrained minimum to \( H \) and, as one can easily show, it implies the Lagrange multiplier rules at \( x_0 \) together with the fact that the point affords a local minimum to \( f \) on \( S \). Moreover, the standard sufficient conditions imply augmentability.

This concept of augmentability also provides a method of multipliers for finding numerical solutions to constrained minimum problems. A brief explanation can be given as follows. Using the notation

\[
H(x, \lambda, \sigma) = f(x) + \langle \lambda, g(x) \rangle + \frac{\sigma}{2} g(x)^2
\]

select \( \lambda_0 \) and \( \sigma > 0 \), hopefully so that \( H(x, \lambda_0, \sigma) \) is convex in \( x \). Choose \( \xi_0, \xi_1, \ldots \) with \( \xi_k \geq \xi_0 > 0 \) and choose \( x_k, \lambda_k \) successively so that \( x_k \) minimizes \( H(x, \lambda_{k-1}, \sigma + \xi_{k-1}) \). Set

\[
\lambda_k = \lambda_{k-1} + \xi_{k-1} g(x_k).
\]

Then, as explained in Hestenes [10], usually \( \{ x_k \} \) converges to a solution \( x_0 \) to the problem and \( \{ \lambda_k \} \) converges to the Lagrange multiplier associated with \( x_0 \).

The importance of this theory in the finite dimensional case (see [11, 12, 25]) has been recognized particularly in the development of computational procedures (see, for example, [1, 2, 7, 8, 16, 18, 26] and references therein, where a wide range of applications illustrate the use of the theory) but it has received little attention in the development of other areas of optimization. It is important to mention that the method of multipliers for finding numerical solutions has been generalized to certain classes of problems in convex programming [17]. Also, the idea of generalizing this theory to optimal control problems involving mixed constraints has been recently developed in [21–24].

Now, the results given in [12] include a sketch of how some aspects of this theory can be applied to infinite dimensional problems such as the problem of Lagrange mentioned above. The approach is based on earlier results given by Hestenes [9] and McShane [15] and our aim in this paper is to develop that theory by explaining clearly the role played by a generalized Lagrangian on which the notion of augmentability can be based.

In [23] we deal with the same problem and provide some of the basic ideas to define different notions of augmentability according to the local nature of the solution to the problem. Some of the proofs were given in [24] where it is shown, in particular, that the well-known sufficient conditions for a weak local minimum imply weak augmentability.

In this paper we turn to local optimality in terms of strong minima and state the corresponding result assuming the standard sufficient conditions for a strong local minimum. We shall prove that those standard conditions imply strong augmentability which in turn implies that the arc under consideration is a strong local minimum.

Clearly this is a crucial aspect of the theory since it provides an alternative approach not only to the derivation of necessary conditions but also for sufficiency results.
2 Statement of the problem

The problem we shall be dealing with can be stated as follows. Suppose we are given an interval \( T := [0, t_1] \) 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 \( \mathbb{R} \).

Let \( X \) denote the space of all absolutely continuous functions mapping \( T \) to \( \mathbb{R}^n \). For any \( C \) subset of \( T \times \mathbb{R}^n \times \mathbb{R}^n \) let

\[
X(C) := \{ x \in X \mid (t, x(t), \dot{x}(t)) \in C \text{ a.e. in } T \},
\]

\[
X_\epsilon(C) := \{ x \in X(C) \mid x(t_0) = \xi_0, x(t_1) = \xi_1 \},
\]

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

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

Denote by \( P(I, C) \) the problem of minimizing \( I \) over \( X_\epsilon(C) \).

Suppose also that a (relatively) open set \( A \) of \( T \times \mathbb{R}^n \times \mathbb{R}^n \), and a function \( \varphi \) mapping \( T \times \mathbb{R}^n \times \mathbb{R}^n \) to \( \mathbb{R}^q \), are given. Let

\[
B := \{ (t, x, \dot{x}) \in A \mid \varphi(t, x, \dot{x}) = 0 \}.
\]

The problem we shall be concerned with is \( P(I, B) \), that is, the problem of minimizing

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

subject to

\( a. \) \( x : T \to \mathbb{R}^n \) absolutely continuous;

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

\( c. \) \( (t, x(t), \dot{x}(t)) \in \mathcal{A} \) and \( \varphi(t, x(t), \dot{x}(t)) = 0 \)
a.e. in \( T \).

Elements of \( X \) will be called trajectories and, for any \( C \subset T \times \mathbb{R}^n \times \mathbb{R}^n \), a trajectory \( x \) solves \( P(I, C) \) if \( x \in X_\epsilon(C) \) and \( I(x) \leq I(y) \) for all \( y \in X_\epsilon(C) \).

For local minima, a trajectory \( x \) will be called a strong or a weak minimum of \( P(I, C) \) if, for some \( \epsilon > 0 \), \( x \) solves \( P(I, T_0(x; \epsilon) \cap C) \) or \( P(I, T_1(x; \epsilon) \cap C) \) respectively where, for all \( x \in X \) and \( \epsilon > 0 \),

\[
T_0(x; \epsilon) := \{ (t, y, v) \in T \times \mathbb{R}^n \times \mathbb{R}^n : |x(t) - y| < \epsilon \}
\]

(called in some references a “tube” around \( x \)), and

\[
T_1(x; \epsilon) := \{ (t, y, v) \in T_0(x; \epsilon) : |\dot{x}(t) - v| < \epsilon \}
\]

(corresponding to a “restricted tube” around \( x \)).

Let us assume that the functions \( L(t, x, \dot{x}) \) and \( \varphi(t, x, \dot{x}) \) and their derivatives with respect to \( x \) and \( \dot{x} \) are continuous on \( \mathcal{A} \) and the matrix \( A \) has rank \( q \) on \( \mathcal{A} \). For any \( x \in X \) the notation \( (\tilde{x}(t)) \) will be used to represent \( (t, x(t), \dot{x}(t)) \).

3 Sufficient conditions

In the theory to follow it will be convenient to first consider the unconstrained problem \( P(I, A) \) usually referred to as the simple fixed endpoint problem in the calculus of variations. This will be done not only for comparison reasons but also to explain clearly the role played by this result in the theory of augmentability.

For \( x \in X \), define the first variation of \( I \) along \( x \) by

\[
I'(x; y) := \int_{t_0}^{t_1} \{ L_x(\tilde{x}(t))y(t) + L_{\dot{x}}(\tilde{x}(t))\dot{y}(t) \}dt
\]

and the second variation of \( I \) along \( x \) by

\[
I''(x; y) := \int_{t_0}^{t_1} 2\Omega(t, y(t), \dot{y}(t))dt \quad (y \in X)
\]

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

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

Define the set of admissible variations by

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

and consider the following sets:

\[
\mathcal{E} := \{ x \in X \mid I'(x; y) = 0 \text{ for all } y \in Y \},
\]

\[
\mathcal{H}' := \{ x \in X \mid I''(x; y) > 0 \text{ for all } y \in Y, y \neq 0 \},
\]

\[
\mathcal{L} := \{ x \in X \mid L_{xx}(\tilde{x}(t)) > 0 \text{ for all } t \in T \},
\]

\[
\mathcal{W}(A, \epsilon) := \{ x_0 \in X(A) \mid \mathcal{E}_L(t, x, \dot{x}, u) \geq 0 \text{ for all } (t, x, \dot{x}, u) \in T \times \mathbb{R}^n \text{ and } (t, x, \dot{x}, u) \in \mathcal{A} \}
\]

where \( E_L : T \times \mathbb{R}^n \to \mathbb{R} \), the Weierstrass “excess function” with respect to \( L \), is given by

\[
E_L(t, x, \dot{x}, u) := L(t, x, \dot{x}) - L(t, x, \dot{x}) - L_{\dot{x}}(t, x, \dot{x})(u - \dot{x}).
\]

Following the terminology of \([10]\), elements of \( \mathcal{E} \cap \mathcal{C}^1 \) are called extremals, elements of \( \mathcal{L}' \) are said to satisfy the strengthened condition of Legendre, and elements of \( \mathcal{W}(A, \epsilon) \) to satisfy the strengthened condition of Weierstrass. Also, recall that \( x \) belongs to \( \mathcal{E} \) if and only if there exists \( c \in \mathbb{R}^n \) such that

\[
L_{\dot{x}}(\tilde{x}(t)) = \int_{t_0}^{t_1} L_{x}(\tilde{x}(s))ds + c \quad (t \in T).
\]
The following theorem gives sufficient conditions for a local solution to $P(I, A)$. We refer to [10] for a full account of this theory.

**3.1 Theorem:** Let $x \in X_c(A) \cap C^1$. If $x$ belongs to $E \cap H' \cap L'$ then $x$ is a strict weak minimum of $P(I, A)$. If also $x$ belongs to

$$W(A; \epsilon) \quad \text{for some} \ \epsilon > 0$$

then $x$ is a strict strong minimum of $P(I, A)$.

Let us turn now to our original constrained problem $P(I, B)$. Sufficiency for this problem is usually expressed in terms of the following functions and sets.

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

$$F(t, x, \dot{x}, \mu) := L(t, x, \dot{x}) + (\mu, \varphi(t, x, \dot{x}))$$

and consider the second variation of the functional $J(x, \mu) := \int_{t_0}^{t_1} F(t, x(t), \dot{x}(t), \mu(t))dt$ given by

$$J''(x, \mu; y) := \int_{t_0}^{t_1} 2\Omega_x(t, y(t), \dot{y}(t))dt$$

where

$$2\Omega_x(t, y, \dot{y}) := \langle y, F_{xx}(\ddot{x}(t), \mu(t))y \rangle + 2\langle y, F_{x\dot{x}}(\ddot{x}(t), \mu(t))\dot{y} \rangle + \langle \dot{y}, F_{\dot{x}\dot{x}}(\ddot{x}(t), \mu(t))\dot{y} \rangle.$$ 

For all $x \in X$ define the set $Y(B, x)$ of $B$-admissible variations along $x$ as the set of all $y \in X$ satisfying $y(t_0) = y(t_1) = 0$ and

$$\varphi_x(\ddot{x}(t))y(t) + \varphi_x(\dddot{x}(t))\dot{y}(t) = 0 \quad \text{(a.e. in} \ T\text{)}.$$ 

Define now

$$E(\mu) := \{ x \in X \mid \text{there exists} \ c \in \mathbb{R}^n \text{ such that} \}$$

$$F_x(\ddot{x}(t), \mu(t)) = \int_{t_0}^{t_1} F_x(\ddot{x}(s), \mu(s))ds + c \quad \text{(} t \in \ T \text{)\},}$$

$$H'(\mu) := \{ x \in X \mid J''(x, \mu; y) > 0 \}$$

for all $y \in Y(B, x)$, $y \neq 0$, $H'(\mu)$, $\mu(\mu) := \{ x \in X \mid \langle h, F_{z\dot{x}}(\ddot{x}(t), \mu(t))h \rangle > 0 \}

\text{for all} \ h \in \mathbb{R}^n, \ h \neq 0$ such that

$$\varphi_z(\dddot{x}(t))h = 0 \ (t \in \ T).$$

$$W(B, \mu; \epsilon) := \{ x_0 \in X(B) \mid \}$$

$$E_F(t, x, \dot{x}, u, \mu(t)) \geq 0$$

for all $(t, x, \dot{x}, u) \in T \times \mathbb{R}^n$ with $(t, x, \dot{x}) \in T_1(x_0; \epsilon) \cap B$ and $(t, x, u) \in B$ where $E_F$ denotes the Weierstrass excess function with respect to $F$, that is,

$$E_F(t, x, \dot{x}, u, \mu) := F(t, x, \mu) - F(t, x, \dot{x}, \mu)$$

$$- F(t, x, \dot{x}, \mu)(u - \dot{x}).$$

The next result gives sufficient conditions for the constrained Lagrange problem we are dealing with (see [9, 10]).

**3.2 Theorem:** Let $x \in X_c(B) \cap C^1$ and let $\mu$ be an absolutely continuous function mapping $T$ to $\mathbb{R}^q$. If $x$ belongs to

$$E(\mu) \cap H'(\mu) \cap L'(\mu)$$

then $x$ is a strict weak minimum of $P(I, B)$. If also $x$ belongs to

$$W(B, \mu; \epsilon) \quad \text{for some} \ \epsilon > 0$$

then $x$ is a strict strong minimum of $P(I, B)$.

As mentioned in the introduction, the proof of this result is usually established by invoking general embedding or field theorems of the theory of differential equations, and those theorems are an integral component of the usual proofs (see, for example, [10]). In the following section we turn to an entirely different approach.

**4 Augmentability**

For a given function $\sigma$ mapping $A$ to $\mathbb{R}$ and for all $(t, x, \dot{x}, \mu) \in T \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^q$, define

$$\tilde{F}(t, x, \dot{x}, \mu) := L(t, x, \dot{x}) + (\mu, \varphi(t, x, \dot{x}))$$

$$+ \sigma(t, x, \dot{x})G(t, x, \dot{x})$$

where

$$G(t, x, \dot{x}) := \frac{1}{2} \sum_{i=1}^{q} \varphi_{\alpha}(t, x, \dot{x})^2.$$ 

Note that

$$\tilde{F}(t, x, \dot{x}, \mu) = F(t, x, \dot{x}, \mu) + \frac{\sigma(t, x, \dot{x})}{2} |\varphi(t, x, \dot{x})|^2.$$ 

**4.1 Definition:** For any $x_0 \in X_c(B)$ we shall say that $P(I, B)$ is strongly (weakly) augmentable at $x_0$ if
there exist \( \sigma : \mathcal{A} \to \mathbb{R} \) and \( \mu : T \to \mathbb{R}^q \) such that \( x_0 \) is a strong (weak) minimum of the unconstrained problem \( P(J, \mathcal{A}) \), where

\[
\bar{J}(x, \mu) := \int_{t_0}^{t_1} \bar{F}(t, x(t), \dot{x}(t), \mu(t))dt.
\]

Note that, in this event, \( x_0 \) is a strong (weak) minimum of \( P(I, \mathcal{B}) \) since, for any \( x \in \mathcal{X}_e(\mathcal{B}) \), we have \( \tilde{J}(x, \mu) = I(x) \).

The definitions given above of strong and weak augmentability yield one of the main aspects of the theory related to necessity. If the problem is augmentable at an arc, then that arc satisfies the first and second order necessary conditions for the constrained problem posed above and this in turn implies the necessary conditions for the constrained case. This result holds without the usual assumption of normality and is applicable to both strong or weak minima (see [24] for details).

For sufficiency, we have the following crucial result which, in particular, implies that the notion of augmentability can be seen as an alternative approach to establish sufficiency results.

**4.2 Theorem:** Let \( x_0 \in \mathcal{X}_e(\mathcal{B}) \cap C^1 \) and let \( \mu \) be an absolutely continuous function mapping \( T \) to \( \mathbb{R}^q \). If \( x_0 \) belongs to

\[
\mathcal{E}(\mu) \cap \mathcal{H}(\mu) \cap L'(\mu)
\]

then \( P(I, \mathcal{B}) \) is weakly augmentable at \( x_0 \). If also \( x_0 \) belongs to

\[
\mathcal{W}(\mathcal{B}, \mu; \epsilon)
\]

for some \( \epsilon > 0 \) then \( P(I, \mathcal{B}) \) is strongly augmentable at \( x_0 \).

This result was first stated in [23] and a proof of the first part, corresponding to weak augmentability, can be found in [24]. For completeness, and being an integral part of the second statement, we shall find convenient to give a sketch of that proof and then turn entirely to the second part.

Let us denote by \( \mathcal{E}(\mu, \sigma), \mathcal{H}(\mu, \sigma), \bar{L}(\mu, \sigma) \) and \( \mathcal{W}(\mathcal{A}, \mu, \sigma; \epsilon) \) the four sets defined before Theorem 3.1 but now replacing \( L \) with the augmented Lagrangian \( \tilde{F} \),\n
Explicitly,

\[
\mathcal{E}(\mu, \sigma) := \{ x \in X \mid \text{there exists } c \in \mathbb{R}^q \text{ such that } \bar{F}_x(x(t), \mu(t)) = \int_{t_0}^{t_1} \bar{F}_x(s, \mu(s))ds + c \}
\]

\[
\mathcal{H}(\mu, \sigma) := \{ x \in X \mid J''(x, \mu; y) > 0 \text{ for all } y \in Y, y \neq 0 \},
\]

\[
L'(\mu, \sigma) := \{ x \in X \mid \langle h, \bar{F}_x(x(t), \mu(t))h \rangle > 0 \text{ for all } h \in \mathbb{R}^n, h \neq 0, (t \in T) \},
\]

\[
\mathcal{W}(\mathcal{A}, \mu, \sigma; \epsilon) := \{ x_0 \in X(\mathcal{A}) \mid E_{\tilde{F}}(x, t, \dot{x}, u, \mu(t)) \geq 0 \}
\]

These sets are precisely the ones that define the classical sufficient conditions in terms of the function \( F \), that is,\n
\[
E_{\tilde{F}}(t, x, \dot{x}, u, \mu) := \tilde{F}(t, x, u, \mu) - \tilde{F}(t, x, \dot{x}, \mu)
\]

\[
- \tilde{F}_x(t, x, \dot{x}, \mu)(u - \dot{x}).
\]

**Proof:** To prove (a), for all \( t \in T \) and \( h \in \mathbb{R}^n \), define

\[
P(t, h) := \langle h, F_{\tilde{F}_x}(\tilde{x}_0(t), \mu(t), 1)h \rangle,
\]

\[
Q(t, h) := \| \tilde{F}_x(\tilde{x}_0(t))h \|^2.
\]

Since \( x_0 \in \mathcal{L}'(\mu) \), we have

\[
P(t, h) > 0 \text{ for all } t \in T \text{ and } h \neq 0
\]

with \( Q(t, h) = 0 \). We claim that, for some constant \( \theta > 0 \),

\[
P(t, h) + \theta Q(t, h) > 0 \text{ for all } t \in T \text{ and } h \neq 0.
\]

Suppose the contrary. Then, for all \( q \in \mathbb{N} \), there exist \( (t_q, h_q) \in T \times \mathbb{R}^n \) with \( h_q \neq 0 \) such that

\[
P(t_q, h_q) + \theta Q(t_q, h_q) \leq 0.
\]
Let \( k_q := h_q/|h_q| \) so that \( P(t_q, k_q) + qQ(t_q, k_q) \leq 0 \) and \( |k_q| = 1 \). Thus there exist a subsequence (we do not relabel), \( t_0 \in T \) and a unit vector \( k_0 \) such that \( (t_q, k_q) \to (t_0, k_0) \). Therefore \( P(t_0, k_0) \leq 0 \) and \( Q(t_0, k_0) = 0 \), contrary to the assumption \( x_0 \in L'(\mu) \). Now, let \( \sigma(t, x, \dot{x}) \) be such that \( \sigma(\tilde{x}_0(t)) > \theta (t \in T) \). Hence

\[
\langle h, \tilde{F}_{\pm}(\tilde{x}_0(t), \mu(t))h \rangle = P(t, h) + \sigma(\tilde{x}_0(t))Q(t, h) > 0
\]

for all \( h \in \mathbb{R}^n, h \neq 0 \), and \( t \in T \), showing that \( x_0 \in L'(\mu, \sigma) \).

The next auxiliary result shows that the strengthened condition of Legendre together with the positivity of the second variation with respect to \( F \) imply the existence of a function \( \sigma \) for which the second variation with respect to \( \tilde{F} \) is also positive. The main ideas of the proof are based on the theory developed by Hestenes in [9].

4.4 Lemma: If \( x_0 \in L'(\mu) \cap H'(\mu) \), then there exists \( \theta_0 > 0 \) such that, if \( \sigma(t, x, \dot{x}) \geq \theta_0 \), then \( x_0 \in H'(\mu, \sigma) \).

Proof: Define

\[
\Phi(t, y, \dot{y}) := \varphi_x(\tilde{x}_0(t)) + \varphi_x(\tilde{x}_0(t))\dot{y},
\]

\[
P(y) := J''(x_0, \mu; y),
\]

\[
Q(y) := \int_{t_0}^{t_1} |\Phi(t, y(t), \dot{y}(t))|^2 dt.
\]

Since \( x_0 \in H'(\mu) \) we have

\[
P(y) > 0 \quad \text{for all } y \in X'', \ y \neq 0,
\]

satisfying \( \Phi(t, y(t), \dot{y}(t)) = 0 \) a.e. in \( T \) and \( y(t_0) = y(t_1) = 0 \). As one readily verifies,

\[
\tilde{J}''(x_0, \mu; y) = P(y) + \int_{t_0}^{t_1} \sigma(\tilde{x}_0(t))|\Phi(t, y(t), \dot{y}(t))|^2 dt
\]

and so

\[
\tilde{J}''(x_0, \mu; y) \geq P(y) + \theta_0 Q(y) \quad \text{if } \sigma(\tilde{x}_0(t)) \geq \theta_0
\]

the equality holding when \( \sigma(\tilde{x}_0(t)) = \theta_0 \).

Let us suppose the conclusion of the theorem is false. Then, for all \( q \in \mathbb{N} \), if \( \sigma(t, x, \dot{x}) \geq q \) we have \( x_0 \notin H'(\mu, \sigma) \). That is, for all \( q \in \mathbb{N} \) there exists \( y_q \in X'' \) nonnull with \( y_q(t_0) = y_q(t_1) = 0 \) such that

\[
P(y_q) + qQ(y_q) \leq \tilde{J}''(x_0, \mu; y_q) \leq 0 \quad (1)
\]

if \( \sigma(t, x, \dot{x}) \geq q \).

Since the functions at hand are homogeneous in \( y \) we can suppose that \( y_q \) has been chosen so that

\[
\int_{t_0}^{t_1} \{|y_q(t)|^2 + |\dot{y}_q(t)|^2 \} dt = 1. \quad (2)
\]

Therefore we can replace the sequence \( \{y_q\} \) by a subsequence (we do not relabel) which converges to a variation \( y_0 \) in the sense that

\[
\lim_{q \to \infty} y_q(t) = y_0(t) \quad \text{uniformly on } T. \quad (3)
\]

Obviously \( y_0(t_0) = y_0(t_1) = 0 \). By Lemma 4.3 there exists \( \theta > 0 \) such that

\[
\lim \inf \{P(y_q) + \theta Q(y_q)\} \geq P(y_0) + \theta Q(y_0). \quad (4)
\]

This inequality, together with (1) and \( Q(y) \geq 0 \), implies that

\[
\lim \inf Q(y_q) \leq 0.
\]

But since the Legendre condition holds for \( Q(y) \) we have that

\[
\lim \inf Q(y_q) \geq Q(y_0) \geq 0.
\]

Consequently \( Q(y_0) = 0 \). Clearly this can be the case only if \( \Phi(t, y_0(t), \dot{y}_0(t)) = 0 \) a.e. in \( T \). Suppose that \( y_0 \neq 0 \). Then \( P(y_0) > 0 \). However, by (4) with \( Q(y_0) = 0 \) one has, for large values of \( q \) that

\[
P(y_q) + \theta Q(y_q) > 0
\]

Contradicting the inequality in (1). Hence \( y_0 \equiv 0 \).

Let us complete the proof by showing that \( y_0 \) cannot be the null variation. Suppose that this is the case. Take \( \sigma = \theta \) as described in Lemma 4.3. Then, by (4), we have

\[
\lim \inf \tilde{J}''(x_0, \mu; y_q) = \lim \inf \{P(y_q) + \theta Q(y_q)\} \geq 0
\]

since \( P(y_0) = Q(y_0) = 0 \). Using (1), we see that the equality must hold. Consequently, by (3) and the assumption \( y_0 \equiv 0 \), we have

\[
0 = \lim \inf \tilde{J}''(x_0, \mu; y_q) = \lim \inf \int_{t_0}^{t_1} \langle \dot{y}_q(t), \tilde{F}_{\pm}(\tilde{x}_0(t), \mu(t)) \rangle dt. \quad (5)
\]

Since, by Lemma 4.3, the last integrand is a positive definite form, there is a constant \( c > 0 \) such that

\[
\langle h, \tilde{F}_{\pm}(\tilde{x}_0(t), \mu(t))h \rangle \geq \langle h, ch \rangle \geq c|h|^2.
\]
Consequently equation (5) implies that

$$\lim_{q \to \infty} \int_{t_0}^{t_1} |\dot{y}_q(t)|^2 dt = 0.$$  

Using (2) and (3) we see that

$$\lim_{q \to \infty} \int_{t_0}^{t_1} |\dot{y}_q(t)|^2 dt = 1.$$  

This contradiction completes the proof. \[ \square \]

Let us turn now to the second part of Theorem 4.2. We are assuming that $x_0$ belongs to

$$\mathcal{E}(\mu) \cap \mathcal{H}'(\mu) \cap \mathcal{L}'(\mu) \cap \mathcal{W}(B, \mu; \epsilon)$$

and we want to prove that $P(J, B)$ is strongly augmentable at $x_0$. This follows if we show that there exists a function $\sigma$ mapping $A$ to $\mathbb{R}$ such that $x_0$ is a strong minimum of the problem $P(J, A)$, and we have well-known sufficient conditions for such unconstrained problem stated in Theorem 3.1.

In other words, the result will follow if we show that, for some function $\sigma$, $x_0$ satisfies the classical sufficient conditions for a strong minimum in terms of the function $\tilde{F}$.

Thus we want to show that, for some function $\sigma(t, x, \dot{x})$ and $\epsilon > 0$, $x_0$ belongs to

$$\tilde{\mathcal{E}}(\mu, \sigma) \cap \tilde{\mathcal{H}}'(\mu, \sigma) \cap \tilde{\mathcal{L}}'(\mu, \sigma) \cap \tilde{\mathcal{W}}(A, \mu, \sigma; \epsilon).$$

Observe first that, for any $\sigma$, $x_0 \in \tilde{\mathcal{E}}(\mu, \sigma)$. To prove it note that, since $x_0 \in \mathcal{E}(\mu)$, there exists $c \in \mathbb{R}^n$ such that, for all $t \in T$,

$$F_x(\tilde{x}_0(t), \mu(t)) = \int_{t_0}^{t} F_x(\tilde{x}_0(s), \mu(s)) ds + c.$$  

On the other hand, we have

$$F_x(t, x, \dot{x}, \mu) = F_x(t, x, \dot{x}, \mu)$$

$$+ \sigma(t, x, \dot{x}) \sum \varphi_{\alpha}(t, x, \dot{x}) \varphi_{\alpha}(t, x, \dot{x})$$

$$+ \frac{\sigma(z(t, x, \dot{x}))}{2} |\varphi(t, x, \dot{x})|^2$$

and similarly for $\tilde{F}_x$. Thus $\tilde{F}_x = F_x$ and $\tilde{F}_x = F_x$ on $\mathcal{B}$, and therefore

$$\tilde{F}_x(\tilde{x}_0(t), \mu(t)) = \int_{t_0}^{t} \tilde{F}_x(\tilde{x}_0(s), \mu(s)) ds + c$$

showing that $x_0$ belongs to $\tilde{\mathcal{E}}(\mu, \sigma)$.

Now, by Lemma 4.3(a), the assumption $x_0 \in \mathcal{L}'(\mu)$ implies the existence of a positive constant $\theta$ such that, if $\sigma(t, x, \dot{x}) \geq \theta$, then $x_0 \in \tilde{\mathcal{L}}'(\mu, \sigma)$. Similarly, by Lemma 4.4, the assumption $x_0 \in \mathcal{L}'(\mu) \cap \mathcal{H}'(\mu) \cap \mathcal{W}(B, \mu; \epsilon)$ implies the existence of $\theta_0 > 0$ such that, if $\sigma(t, x, \dot{x}) \geq \theta_0$, then $x_0 \in \mathcal{H}'(\mu, \sigma)$.

It remains to prove that the assumption $x_0 \in \mathcal{W}(B, \mu; \epsilon)$ implies the existence of $\sigma(t, x, \dot{x})$ such that $x_0 \in \mathcal{W}(A, \mu, \sigma; \epsilon)$.

This can be done as follows. Let $E_{\psi}(\dot{x}, u)$ be the Weierstrass excess function for

$$\psi(\dot{x}) := (1 + |\dot{x}|^2)^{1/2}.$$  

As mentioned in [9], $x_0$ satisfies the strengthened condition of Weierstrass

$$E_F(t, x, \dot{x}, u, \mu(t)) \geq 0$$

whenever $(t, x, \dot{x}) \in T_1(x_0; \epsilon) \cap \mathcal{B}$, $(t, x, u) \in \mathcal{B}$ (that is, $x_0 \in \mathcal{W}(B, \mu; \epsilon)$), and the arc is nonsingular in the sense that

$$|F_{\dot{x}x} - \varphi_{\dot{x}}| \neq 0$$

along $x_0$.

if and only if there exist a neighborhood $B_0$ of $x_0$ relative to $\mathcal{B}$ and $\tau > 0$ such that

$$E_F(t, x, \dot{x}, u, \mu(t)) \geq \tau E_{\psi}(\dot{x}, u)$$

whenever $(t, x, \dot{x}) \in B_0$, $(t, x, u) \in \mathcal{B}$.

Note that, since $\varphi_{\dot{x}}(\tilde{x}_0(t))$ has rank $q$ and

$$\langle h, F_{\dot{x}x}(\tilde{x}_0(t)) h \rangle \geq 0$$

for all $h \in S$ where

$$S = \{ h \in \mathbb{R}^n \mid \varphi_{\dot{x}}(\tilde{x}_0(t)) h = 0 \},$$

the condition of nonsingularity is equivalent to the relation

$$\langle h, F_{\dot{x}x}(\tilde{x}_0(t)) h \rangle > 0$$

for all $h \neq 0$ in $S$ that is, $x_0 \in \mathcal{L}'(\mu)$. Also, as one readily verifies,

$$E_{E_F}(t, x, \dot{x}, u, \mu(t)) \geq \tau E_{\psi}(\dot{x}, u)$$

whenever $(t, x, \dot{x}) \in B_0$, $(t, x, u) \in \mathcal{B}$.

Now, the main idea of the proof consists in showing that there exist a function $\sigma(t, x, \dot{x}) \geq \theta_0$, a positive number $\tau$ and $A_0$ neighborhood of $x_0$ relative to $A$ such that

$$E_{E_F}(t, x, \dot{x}, u, \mu(t)) \geq \tau E_{\psi}(\dot{x}, u)$$

whenever $(t, x, \dot{x}) \in A_0$, $(t, x, u) \in A$.

In view of the above characterization, this will imply that $x_0$ belongs to $\mathcal{W}(A, \mu, \sigma; \epsilon)$ and an application of Theorem 3.1 yields the required result.
This will be proved in several steps. Observe first that, since
\[ F(t, x, \dot{x}, \mu) = F(t, x, \dot{x}, \mu) + \frac{\sigma(t, x, \dot{x})}{2} |\varphi(t, x, \dot{x})|^2, \]
we can assert that, for some \( \tau_1 > 0 \) and \( A_1 \) neighborhood of \( x_0 \),
\[ E_F(t, x, \dot{x}, u) \geq \tau_1 E_{\dot{x}}(\dot{x}, u) \]  
for all \((t, x, \dot{x}) \in B \cap A_1, (t, x, u) \in B\) (for simplicity of notation we have deleted the dependence of the excess function with respect to \( \mu(t) \)).

Since \( x_0 \in \tilde{L}^{2}(\mu, \sigma) \), by Taylor’s theorem and continuity we can diminish \( \tau_1 \) and \( A_1 \) so that (7) holds for all \((t, x, \dot{x}) \) and \((t, x, u) \) in \( A_1 \) (not necessarily in \( B \)).

Now, we claim that there exist \( \tau > 0 \) and \( A_0 \) neighborhood of \( x_0 \) with \( \text{cl} A_0 \subset A_1 \) and let \( 0 < \epsilon < 1 \) be such that, if \((t, x, \dot{x}) \) is in \( A_0 \) and \((t, x, u) \) is in \( A_1 \) and \((t, x, u) \) is exterior to \( A_1 \), then
\[ 3\epsilon |\varphi(u)| \leq E_{\dot{x}}(\dot{x}, u) \leq 2 |\varphi(u)|. \]

Next select \( A_0 \) so small that there exists \( r(t, x, \dot{x}) \) defined on \( A_0 \) such that
\[(t, x, r(t, x, \dot{x})) \in A^* \cap B \quad \text{for all} \quad (t, x, \dot{x}) \in A_0.\]

We have
\[ E_F(t, x, \dot{x}, u) = E_F(t, x, r(t, x, \dot{x}), u) \]
where
\[ K(t, x, \dot{x}, u) = \tilde{F}(t, x, r(t, x, \dot{x})) - \tilde{F}(t, x, \dot{x}) \]
\[ = \tilde{F}(t, x, r(t, x, \dot{x}))u - \tilde{F}(t, x, \dot{x})u \]
\[ = \tilde{F}(t, x, r(t, x, \dot{x}))(\dot{x} - \tilde{F}(t, x, r(t, x, \dot{x}))r(t, x, \dot{x}). \]

If \( A_0 \) is taken sufficiently small we have
\[ |K(t, x, \dot{x}, u)| < \epsilon \tau_1 |\varphi(u)| \quad \text{for all} \quad (t, x, \dot{x}) \in A_0. \]

By (7) and (8), if \((t, x, \dot{x}) \in A_0 \) and \((t, x, u) \) in \( B \) but not in \( A_1 \) then
\[ E_F(t, x, \dot{x}, u) \geq \tau_1 E_{\dot{x}}(r(t, x, \dot{x}), u) - \epsilon \tau_1 |\varphi(u)| \]
\[ \geq 2 \epsilon \tau_1 |\varphi(u)| \geq \epsilon \tau_1 E_{\dot{x}}(\dot{x}, u). \]

Setting \( \tau = \epsilon \tau_1 \) in (6), the claim follows.

Select now open sets \( A_2, A_3, \ldots \) whose union is \( A \) and such that \( \text{cl} A_j \subset A_{j+1} \) \((j = 1, 2, \ldots)\) and let \( \theta_j(t, x, \dot{x}) \) be functions of class \( C^2 \) such that
\[ \theta_j = 0 \text{ on } A_{j-1}, \]
\[ \theta_j \geq 0 \text{ on } A_j, \]
\[ \theta_j = 1 \text{ on } A \sim A_j. \]  

Let \((t, x, u) \in A_0 \). If \((t, x, u) \in A_{j+1} \) but \((t, x, u) \not\in A_j \) \((j \geq 1)\) then (6) holds if \((t, x, u) \in B \) and so, by continuity, if
\[ \varphi_\alpha(t, x, u)^2 < \epsilon_j E_{\dot{x}}(\dot{x}, u) \]  
where \( \epsilon_j \) is a small positive constant. Select \( \delta_j > 0 \) such that
\[ E_F(t, x, \dot{x}, u) \geq \delta_j \quad \text{if} \quad (t, x, u) \in A_{j+1} \sim A_j. \]

Let \( d_j > 0 \) be such that, on this set,
\[ d_j \epsilon_j E_{\dot{x}}(\dot{x}, u) + \delta_j > \tau E_{\dot{x}}(\dot{x}, u). \]

Set \( \sigma(t, x, \dot{x}) := \theta_0 + \sum d_j \theta_j(t, x, \dot{x}). \)

Now, observe that, by (9), we have
\[ \sigma - \theta_0 = 0 \text{ on } A_0, \]
\[ \sigma - \theta_0 \geq 0 \text{ on } A, \]
\[ \sigma - \theta_0 \geq d_j \text{ on } A \sim A_j. \]

Set
\[ h(t, x, \dot{x}) := (\sigma(t, x, \dot{x}) - \theta_0)|\varphi(t, x, \dot{x})|^2. \]

Note that \( h \equiv 0 \) on \( A_0 \) and
\[ E_h(t, x, \dot{x}, u) = h(t, x, u) \geq 0 \]
(provided \((t, x, \dot{x}) \) in \( A_0 \) as we have supposed).

Let \( F^* = \tilde{F} + h \). We have, whenever \((t, x, \dot{x}) \in A_0, \)
\[ E_{F^*}(t, x, \dot{x}, u) = E_F(t, x, \dot{x}, u) + h(t, x, u). \]

If \((t, x, u) \in A_1 \) then (6) holds so that
\[ E_{F^*}(t, x, \dot{x}, u) \geq \tau E_{\dot{x}}(\dot{x}, u). \]

The same is true if \((t, x, \dot{x}) \in A_{j+1} \sim A_j \) \((j \geq 1)\) and (10) holds. If (10) fails to hold then, by (12),
\[ h(t, x, u) \geq (\sigma(t, x, \dot{x}) - \theta_0) \epsilon_j E_{\dot{x}}(\dot{x}, u) \geq d_j \epsilon_j E_{\dot{x}}(\dot{x}, u). \]

It follows from (11) and (13) that (14) holds in this case also. \( \blacksquare \)
5 Conclusion

In this paper we deal with a Lagrange problem in the calculus of variations involving mixed equality constraints. For such a problem, necessary and sufficient conditions are well established in the literature.

However, the classical techniques used both for necessity (under the assumption of normality) and sufficiency (invoking general embedding of field theorems of the theory of differential equations, or matrix Riccati equations, or Jacobi’s theory on conjugate points, or Hamilton-Jacobi theory) can be substantially simplified by using a notion of a certain type of augmentability.

We introduce the notions of weak and strong augmentability and study their implications in sufficiency theory. This aspect of the theory of augmentability is crucial since, being an alternative approach to that of normality for necessary conditions, one is interested in finding conditions that imply that the problem is augmentable at some point. This paper shows that precisely the classical sufficient conditions both for weak and strong minima imply the corresponding property of augmentability. In other words, sufficiency for a weak or strong local minimum can be established directly from the definition of augmented Lagrangians proposed in this paper.

It is of interest to see if the proof of this sufficiency result can be generalized to optimal control problems. Moreover, as in the case of finite dimensional problems, one can expect that this approach may yield a method of multipliers for finding numerical solutions of such problems.

References:


