Generalized Lagrangians for the Lagrange problem with equality constraints

Javier F. Rosenblueth*

Abstract—In this paper we deal with necessary and sufficient conditions for the Lagrange problem in the calculus of variations involving equality constraints. The approach we follow is based on adding a penalty term to the standard Lagrangian. This type of augmentability has been successfully applied to constrained minimum problems in finite dimensional spaces, particularly in the development of computational procedures, and the main purpose of this paper is to generalize this approach by introducing the notions of weak and strong augmentability which yield first and second order necessary conditions for local minima in the calculus of variations. Moreover, we provide also a simple proof to show that the standard sufficient conditions for a weak local minimum imply weak augmentability.

Keywords—Augmentability, normality, Lagrange problem, calculus of variations, equality constraints

I. Introduction

Dealing with constrained minimum problems in finite dimensional spaces, one is usually interested in proving the linearly independence of the gradients of the constraints in order to derive the Lagrange multiplier rules as necessary conditions for optimality. To be precise, if the problem at hand consists in minimizing a given function $f\colon S \to \mathbf{R}$ on the set

$$S = \{x \in \mathbf{R}^n \mid g_\alpha(x) = 0 \ (\alpha \in A)\}$$

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

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

for all $h \in \mathbf{R}^n$ satisfying $g'_\alpha(x_0; h) = 0$ $(\alpha \in A)$, where

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

denotes the standard Lagrangian, and they 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, one obtains sufficiency for local minima.

An entirely different approach, much simpler to apply, is that of augmentability, where one deals with an augmented Lagrangian of the type

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

where

$$G(x) = \frac{1}{2} \sum_1^m g_\alpha(x)^2.$$

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.

In 1980, Hestenes devoted one of his last papers [10] to call attention to the role of augmentability in optimization theory. As mentioned above, one type of augmentability applied to constrained minimum problems in the finite dimensional case yields the derivation of the Lagrange multiplier rules in a simple way, much simpler than under the assumption of regularity usually used.

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

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

*IIMAS–UNAM, Universidad Nacional Autónoma de México, Apartado Postal 20-726, México DF 01000, México (e-mail: jfrl@unam.mx).

select λ_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, λ_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 significance of this theory in the finite dimensional case, as pointed out by Hestenes [9, 10] and Rupp [20], has been recognized since the 70's particularly in the development of computational procedures (see, for example, [1, 2, 5, 6, 12] 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. One exception is that of convex programming [14] where the original method of multipliers for finding numerical solutions has been generalized.

More recently, this question has been studied in [15–19] for certain classes of optimal control problems involving mixed equality and inequality constraints. However, possible generalizations to other problems in optimization have not received the attention that this theory may deserve.

It is important to mention that the role of penalty functions in optimal control has been used to find solutions to the problem and in the derivation of necessary conditions (see [3] for a detailed explanation). To illustrate the technique used in [3], consider an optimal control problem where the cost is given by

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

and constraints in the state are given by inequalities of the type

$$h(t, x(t)) \leq 0 \quad \text{a.e. in } [t_0, t_1].$$

Then the constraints are removed by penalizing the cost with the integral

$$\int_{t_0}^{t_1} \max\{0, h(t, x(t))\}dt$$

thus obtaining a sequence of problems where one is interested in minimizing

$$\int_{t_0}^{t_1} L(t, x(t), u(t))dt + K \int_{t_0}^{t_1} \max\{0, h(t, x(t))\}dt$$

without constraints in the state functions. This technique produces a nonsmooth optimal control problem since the the cost with the penalty term is not differentiable.

The approach we shall follow in this paper for the Lagrange problem in the calculus of variations with equality constraints is entirely different, as it produces an augmentable integral with a penalty term which is differentiable and for which first and also second order conditions are obtainable.

In [10] one finds a sketch of how this theory can be applied to infinite dimensional problems such as a problem of Lagrange with differential constraints. 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.

Since the augmented problem is obtained by removing constraints, we shall also state the main results on first and second order necessary conditions for unconstrained problems which are used in the derivation of the corresponding conditions for the constrained problem. Finally, we shall state a crucial aspect of this theory, namely, that the well-known sufficient conditions for optimality imply augmentability so that this theory provides an alternative approach not only to the derivation of first and second order necessary conditions but also for sufficiency results. In particular, for a weak local solution to the problem, we provide a simple proof of the fact that the standard sufficient conditions imply weak augmentability.

The Problem

Suppose we are given an interval $T := [t_0, t_1]$ in $\mathbf{R}$, two points ξ_0, ξ_1 in $\mathbf{R}^n$, a set $\mathcal{A}$ of $T \times \mathbf{R}^n \times \mathbf{R}^n$, and a function L mapping $T \times \mathbf{R}^n \times \mathbf{R}^n$ to $\mathbf{R}$.

Let X be the space of all piecewise C^1 functions mapping T to $\mathbf{R}^n$, set

$$X(\mathcal{A}) := \{x \in X \mid (t, x(t), \dot{x}(t)) \in \mathcal{A} \ (t \in T)\},$$

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

and consider the functional $I \colon X \to \mathbf{R}$ 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 P($\mathcal{A}$), is that of minimizing I over $X_e(\mathcal{A})$.

Elements of X will be called *trajectories*, and a trajectory x *solves* P($\mathcal{A}$) if $x \in X_e(\mathcal{A})$ and

$$I(x) \leq I(y) \quad \text{for all } y \in X_e(\mathcal{A}).$$

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

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

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

For any $x \in X$ we shall use the notation $(\tilde{x}(t))$ to represent $(t, x(t), \dot{x}(t))$, and we assume that $L \in C^2(\mathcal{A})$.

II. The simple fixed endpoint problem

Let us consider problem P($\mathcal{A}$) with $\mathcal{A}$ a (relatively) open set in $T \times \mathbf{R}^n \times \mathbf{R}^n$. This problem is called the *simple fixed endpoint problem* in the calculus of variations.

The notion of "variations" in this context has the following meaning. 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 $(t, y, \dot{y}) \in T \times \mathbf{R}^n \times \mathbf{R}^n$,

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

Now, in order to express the best known first and second order necessary conditions for optimality in a succinct way, let us define the set of *admissible variations* by

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

and consider the following sets:

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

$$\mathcal{H} := \{x \in X \mid I''(x; y) \geq 0 \text{ for all } y \in Y\},$$

$$\mathcal{L} := \{x \in X \mid L_{\dot{x}\dot{x}}(\tilde{x}(t)) \geq 0 \text{ for all } t \in T\},$$

$$\mathcal{W}(\mathcal{A}) := \{x \in X(\mathcal{A}) \mid E(t, x(t), \dot{x}(t), u) \geq 0 \text{ for all } (t, u) \in T \times \mathbf{R}^n \text{ with } (t, x(t), u) \in \mathcal{A}\}$$

where $E: T \times \mathbf{R}^{3n} \to \mathbf{R}$, the Weierstrass "excess function," is given by

$$E(t, x, \dot{x}, u) := L(t, x, u) - L(t, x, \dot{x}) - L_{\dot{x}}(t, x, \dot{x})(u - \dot{x}).$$

Elements of $\mathcal{E} \cap C^1$ are usually called *extremals*, elements of $\mathcal{L}$ are said to satisfy the *condition of Legendre*, and elements of $\mathcal{W}(\mathcal{A})$ to satisfy the *condition of Weierstrass*.

The following theorem gives necessary conditions for a solution to P($\mathcal{A}$) where $\mathcal{A}$ is any relatively open set of $T \times \mathbf{R}^n \times \mathbf{R}^n$. We refer the reader to [8] where a full explanation of this theory can be found.

It should be noted that, in particular, if x is a weak minimum of P($\mathcal{A}$) then x belongs to $\mathcal{E}$, $\mathcal{H}$, $\mathcal{L}$ and $\mathcal{W}(T_1(x;\epsilon) \cap \mathcal{A})$ for some $\epsilon > 0$. On the other hand, if x is a strong minimum of P($\mathcal{A}$) then x belongs to $\mathcal{E}$, $\mathcal{H}$, $\mathcal{L}$ and $\mathcal{W}(\mathcal{A})$ since, for any $\epsilon > 0$,

$$x \in \mathcal{W}(T_0(x;\epsilon) \cap \mathcal{A}) \Leftrightarrow x \in \mathcal{W}(\mathcal{A}).$$

2.1 Theorem: *If x solves* P($\mathcal{A}$) *then x belongs to $\mathcal{E}$, $\mathcal{H}$, $\mathcal{L}$ and $\mathcal{W}(\mathcal{A})$.*

For sufficiency, let us consider the following sets obtained by slightly strengthening those defined for necessary conditions:

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

$$\mathcal{L}' := \{x \in X \mid L_{\dot{x}\dot{x}}(\tilde{x}(t)) > 0 \text{ for all } t \in T\},$$

$$\mathcal{W}(\mathcal{A}, \epsilon) := \{x_0 \in X(\mathcal{A}) \mid E(t, x, \dot{x}, u) \geq 0 \text{ for all } (t, x, \dot{x}, u) \in T \times \mathbf{R}^{3n} \text{ with } (t, x, \dot{x}) \in T_1(x_0;\epsilon) \text{ and } (t, x, u) \in \mathcal{A}\}.$$

The following theorem gives sufficient conditions for local minima. It is worth mentioning that, in [8], this result is proved directly without referring to Mayer fields, Hamilton-Jacobi theory, Riccati equations or conjugate points.

2.2 Theorem: *Suppose $x \in X_e(\mathcal{A}) \cap C^1$. Then:*

a. *$x \in \mathcal{E} \cap \mathcal{H}' \cap \mathcal{L}' \Rightarrow x$ is a strict weak minimum of* P($\mathcal{A}$).

b. *$x \in \mathcal{E} \cap \mathcal{H}' \cap \mathcal{L}' \cap \mathcal{W}(\mathcal{A};\epsilon)$ for some $\epsilon > 0 \Rightarrow x$ is a strict strong minimum of* P($\mathcal{A}$).

As it is well-known, if $x \in X(\mathcal{A})$, then $x \in \mathcal{E}$ if and only if there exists $c \in \mathbf{R}^n$ such that

$$L_{\dot{x}}(\tilde{x}(t)) = \int_{t_0}^{t} L_x(\tilde{x}(s))ds + c \qquad (t \in T).$$

This equation is the integral form of *Euler's equation*

$$\frac{d}{dt}L_{\dot{x}}(\tilde{x}(t)) = L_x(\tilde{x}(t)) \qquad (t \in T).$$

If $\dot{x}$ has a discontinuity, the derivative d/dt is to be interpreted as a left- or a right-hand derivative, and it holds even if x fails to have a second derivative.

It is important, in the theory to follow, to bear in mind that if $x \in \mathcal{E}$ then x satisfies Euler's equation, but the converse may not hold. This is easily illustrated by, for example, setting

$$L(t, x, \dot{x}) = (\dot{x}^2 - x^2)/2,\ T = [0, 2\pi],$$

$$x_0(t) := \sin t \text{ if } t \in [0, \pi],\ x_0(t) := 0 \text{ if } t \in [\pi, 2\pi].$$

Then x_0 satisfies Euler's equation, but $x_0 \notin \mathcal{E}$ since there is no $c \in \mathbf{R}$ satisfying $\dot{x}_0(t) = \int_0^t -x_0(s)ds + c$ for all $t \in T$. For the converse note that, if x satisfies Euler's equation, then $x \in \mathcal{E}$ if the function $t \mapsto L_{\dot{x}}(\tilde{x}(t))$ $(t \in T)$ belongs to X.

We shall find convenient to restate this characterization of $\mathcal{E}$ as follows. For all $(t, x, \dot{x}, p) \in T \times \mathbf{R}^{3n}$ let

$$H(t, x, \dot{x}, p) := \langle p, \dot{x} \rangle - L(t, x, \dot{x})$$

and set ('$*$' denotes transpose)

$$M(x) := \{p \in X \mid \dot{p}(t) = -H_x^*(\tilde{x}(t), p(t))$$

$$\text{and } H_{\dot{x}}(\tilde{x}(t), p(t)) = 0\ (t \in T)\} \quad (x \in X).$$

Then, if $x \in X(\mathcal{A})$, we have $x \in \mathcal{E} \Leftrightarrow M(x) \neq \emptyset$.

III. The normality approach

Suppose the data are as before but we are also given a function φ mapping $T \times \mathbf{R}^n \times \mathbf{R}^n$ to $\mathbf{R}^q$. Let

$$\mathcal{B} := \{(t, x, \dot{x}) \in \mathcal{A} \mid \varphi(t, x, \dot{x}) = 0\}$$

and consider problem P($\mathcal{B}$), *the problem of Lagrange with equality constraints*.

The problem is thus that of minimizing

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

subject to

a. $x\colon T \to \mathbf{R}^n$ is piecewise C^1;
b. $x(t_0) = \xi_0$, $x(t_1) = \xi_1$;
c. $(t, x(t), \dot{x}(t)) \in \mathcal{A}$ $(t \in T)$;
d. $\varphi(t, x(t), \dot{x}(t)) = 0$ $(t \in T)$.

We assume that $\varphi \in C^2(\mathcal{A})$ and the $q \times n$-matrix $\varphi_{\dot{x}}(t, x, \dot{x})$ has rank q on $\mathcal{B}$.

Let us begin by stating first order necessary conditions for P($\mathcal{B}$) in the form of a maximum principle. A proof of this result and its corollary can be found in [8].

The Hamiltonian formulation

For all $(t, x, \dot{x}, p, \mu, \lambda) \in T \times \mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R}^q \times \mathbf{R}$ define the *Hamiltonian* as

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

and denote by $\mathcal{U}_q$ the space of piecewise continuous functions mapping T to $\mathbf{R}^q$.

3.1 Theorem: *Suppose x_0 solves P($\mathcal{B}$). Then there exist $\lambda_0 \geq 0$, $p \in X$, and $\mu \in \mathcal{U}_q$ continuous on each interval of continuity of $\dot{x}_0$, not vanishing simultaneously on T, such that*

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

b. *$H(t, x_0(t), u, p(t), \mu(t), \lambda_0) \leq H(t, \lambda_0)$ for all $(t, u) \in T \times \mathbf{R}^n$ with $(t, x_0(t), u) \in \mathcal{B}$.*

where $H(t, \lambda_0)$ denotes $H(\tilde{x}_0(t), p(t), \mu(t), \lambda_0)$.

3.2 Corollary: *Suppose x_0 solves P($\mathcal{B}$). Let (p, μ, λ_0) be as in Theorem 3.1. Then*

$$\langle h, H_{\dot{x}\dot{x}}(\tilde{x}_0(t), p(t), \mu(t), \lambda_0)h \rangle \leq 0$$

for all $h \in \mathbf{R}^n$ such that $\varphi_{\dot{x}}(\tilde{x}_0(t))h = 0$.

In general we are interested in deriving those necessary conditions in such a way that the cost multiplier will not vanish. The notion of "normality" is introduced for that purpose.

3.3 Definition: A trajectory $x \in X(\mathcal{B})$ will be said to be *normal* to P($\mathcal{B}$) if, given $(p, \mu) \in X \times \mathcal{U}_q$ such that, for all $t \in T$,

$$\dot{p}(t) = \varphi_x^*(\tilde{x}(t))\mu(t)\ \ [= -H_x^*(\tilde{x}(t), p(t), \mu(t), 0)]$$

$$0 = p(t) - \varphi_{\dot{x}}^*(\tilde{x}(t))\mu(t)\ \ [= H_{\dot{x}}^*(\tilde{x}(t), p(t), \mu(t), 0)]$$

then $p \equiv 0$. In this event, clearly, also $\mu \equiv 0$, since

$$\mu^*(t) = p^*(t)\varphi_{\dot{x}}^*(\tilde{x}(t))D(t)$$

where $D(t) = [\varphi_{\dot{x}}(\tilde{x}(t))\varphi_{\dot{x}}^*(\tilde{x}(t))]^{-1}$. Thus x is normal to P($\mathcal{B}$) if there is no nonnull solution to

$$\dot{p}^*(t) = p^*(t)\varphi_{\dot{x}}^*(\tilde{x}(t))D(t)\varphi_x(\tilde{x}(t)).$$

Note that, if x_0 is a normal solution to P($\mathcal{B}$) then, in Theorem 3.1, $\lambda_0 > 0$. In this event, the multipliers

λ_0, p, μ can be chosen so that $\lambda_0 = 1$ and, when so chosen, they are unique.

Let us turn now to second order necessary conditions for problem P($\mathcal{B}$) under normality assumptions. We refer to [8] for a proof of the conditions stated below.

3.4 Definition: Given $(x, p, \mu) \in X \times X \times \mathcal{U}_q$, define the *second variation* (with respect to $H(t, 1)$) by

$$K(x, p, \mu; y) = \int_{t_0}^{t_1} 2\tilde{\Omega}(t, y(t), \dot{y}(t))dt \quad (y \in X)$$

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

$$2\tilde{\Omega}(t, y, \dot{y}) := -[\langle y, H_{xx}(t, 1)y\rangle + 2\langle y, H_{x\dot{x}}(t, 1)\dot{y}\rangle + \langle \dot{y}, H_{\dot{x}\dot{x}}(t, 1)\dot{y}\rangle]$$

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

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

$$\varphi_x(\tilde{x}(t))y(t) + \varphi_{\dot{x}}(\tilde{x}(t))\dot{y}(t) = 0 \quad (t \in T).$$

3.6 Theorem: *Suppose x_0 solves* P($\mathcal{B}$) *and is normal to* P($\mathcal{B}$)*. Then there exist a unique pair $(p, \mu) \in X \times \mathcal{U}_q$ continuous on each interval of continuity of $\dot{x}_0$ such that, if $H(t, 1)$ denotes $H(\tilde{x}_0(t), p(t), \mu(t), 1)$, then*

a. $\dot{p}(t) = -H_x^*(t, 1)$ *and* $H_{\dot{x}}(t, 1) = 0$ *on each interval of continuity of* $\dot{x}_0$*;*

b. $H(t, x_0(t), \dot{x}, p(t), \mu(t), 1) \leq H(t, 1)$ *for all* $(t, \dot{x}) \in T \times \mathbf{R}^n$ *with* $(t, x_0(t), \dot{x}) \in \mathcal{B}$*;*

c. $\langle h, H_{\dot{x}\dot{x}}(t, 1)h\rangle \leq 0$ *for all* $h \in \mathbf{R}^n$ *such that* $\varphi_{\dot{x}}(\tilde{x}_0(t))h = 0$*;*

d. $K(x_0, p, \mu; y) \geq 0$ *for all* $y \in Y(\mathcal{B}, x_0)$.

This result gives us first and second order conditions for a normal solution to the problem of Lagrange with equality constraints. A proof different from the one given in [8], using a reduction approach based on the implicit function theorem, can be found in [4] where the optimal control problem studied is that of minimizing the functional

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

subject to

a. $x: T \to \mathbf{R}^n$ piecewise C^1; $u: T \to \mathbf{R}^m$ 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. $(t, x(t), u(t)) \in \mathcal{A}$ $(t \in T)$,

where $T = [t_0, t_1]$,

$$\mathcal{A} = \{(t, x, u) \in T \times \mathbf{R}^n \times \mathbf{R}^m \mid \varphi(t, x, u) = 0\}$$

and $\varphi: T \times \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R}^q$ is a given function. Under mild assumptions on the data of the problem, an immediate consequence of the results obtained in [4] is precisely Theorem 3.6.

Now, these conditions are expressed in terms of the Hamiltonian and a maximum principle but, in order to express them as in Theorem 2.1, that is, in terms of the classical conditions in the calculus of variations, let us now introduce the Lagrangian and derive the corresponding conditions from the previous results.

The Lagrangian formulation

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

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

so that

$$H(t, x, \dot{x}, p, \mu, \lambda) = \langle p, \dot{x}\rangle - F(t, x, \dot{x}, \mu, \lambda).$$

Note that, if x_0 solves P($\mathcal{B}$) and λ_0, p, μ are as in Theorem 3.1, then

$$\begin{aligned} 0 &= H_{\dot{x}}^*(\tilde{x}_0(t), p(t), \mu(t), \lambda_0) \\ &= p(t) - F_{\dot{x}}^*(\tilde{x}_0(t), \mu(t), \lambda_0) \end{aligned}$$

and therefore the function $t \mapsto F_{\dot{x}}^*(\tilde{x}_0(t), \mu(t), \lambda_0)$ $(t \in T)$ belongs to X. Since $H_x = -F_x$, we obtain

$$\frac{d}{dt}F_{\dot{x}}(\tilde{x}_0(t), \mu(t), \lambda_0) = F_x(\tilde{x}_0(t), \mu(t), \lambda_0).$$

Observe also that if E_F is the Weierstrass E-function

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

then we have

$$E_F(\tilde{x}_0(t), u, \mu(t), \lambda) = H(\tilde{x}_0(t), p(t), \mu(t), \lambda) - H(t, x_0(t), u, p(t), \mu(t), \lambda).$$

These remarks imply, by Theorem 3.1, the following result.

3.7 Theorem: *If x_0 solves* P($\mathcal{B}$) *then there exist $\lambda_0 \geq 0$ and $\mu \in \mathcal{U}_q$ continuous on each interval of continuity of $\dot{x}_0$, not vanishing simultaneously on T, such that*

a. *There exists* $c \in \mathbf{R}^n$ *such that, for all* $t \in T$,

$$F_{\dot{x}}(\tilde{x}_0(t), \mu(t), \lambda_0) = \int_{t_0}^{t} F_x(\tilde{x}_0(s), \mu(s), \lambda_0) ds + c;$$

b. $E_F(\tilde{x}_0(t), u, \mu(t), \lambda_0) \geq 0$ *for all* $(t, u) \in T \times \mathbf{R}^n$ *with* $(t, x_0(t), u) \in \mathcal{B}$.

From the definition we clearly have that

$$H_{xx} = -F_{xx}, \quad H_{x\dot{x}} = -F_{x\dot{x}}, \quad H_{\dot{x}\dot{x}} = -F_{\dot{x}\dot{x}}.$$

By Corollary 3.2, if x_0 solves P($\mathcal{B}$) and (μ, λ_0) is as in Theorem 3.7, then

$$\langle h, F_{\dot{x}\dot{x}}(\tilde{x}_0(t), \mu(t), \lambda_0)h \rangle \geq 0$$

for all $h \in \mathbf{R}^n$ such that $\varphi_{\dot{x}}(\tilde{x}_0(t))h = 0$ $(t \in T)$. Moreover, $K(x, p, \mu; y)$ coincides with $J''(x, \mu; y)$, the second variation of the functional

$$J(x, \mu) := \int_{t_0}^{t_1} F(t, x(t), \dot{x}(t), \mu(t), 1) dt$$

given by

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

where

$$2\Omega_\mu(t, y, \dot{y}) := \langle y, F_{xx}(\tilde{x}(t), \mu(t), 1)y \rangle$$
$$+ 2\langle y, F_{x\dot{x}}(\tilde{x}(t), \mu(t), 1)\dot{y} \rangle + \langle \dot{y}, F_{\dot{x}\dot{x}}(\tilde{x}(t), \mu(t), 1)\dot{y} \rangle.$$

To express the corresponding Theorem 3.6 for normal solutions in a succinct way as in Theorem 2.1 let us define, for all $\mu \in \mathcal{U}_q$,

$$\mathcal{E}(\mu) := \{x \in X \mid \text{there exists } c \in \mathbf{R}^n \text{ such that}$$
$$F_{\dot{x}}(\tilde{x}(t), \mu(t), 1) = \int_{t_0}^{t} F_x(\tilde{x}(s), \mu(s), 1) ds + c \quad (t \in T)\},$$

$$\mathcal{H}(\mu) := \{x \in X \mid J''(x, \mu; y) \geq 0 \text{ for all } y \in Y(\mathcal{B}, x)\},$$

$$\mathcal{L}(\mu) := \{x \in X \mid \langle h, F_{\dot{x}\dot{x}}(\tilde{x}(t), \mu(t), 1)h \rangle \geq 0$$
$$\text{for all } h \in \mathbf{R}^n \text{ such that } \varphi_{\dot{x}}(\tilde{x}(t))h = 0 \ (t \in T)\},$$

$$\mathcal{W}(\mathcal{B}, \mu) := \{x \in X(\mathcal{B}) \mid E_F(\tilde{x}(t), u, \mu(t), 1) \geq 0$$
$$\text{for all } (t, u) \in T \times \mathbf{R}^n \text{ with } (t, x(t), u) \in \mathcal{B}\}.$$

3.8 Theorem: *If* x_0 *is a normal solution to* P($\mathcal{B}$) *then there exists a unique* $\mu \in \mathcal{U}_q$ *such that* x_0 *belongs to* $\mathcal{E}(\mu)$, $\mathcal{H}(\mu)$, $\mathcal{L}(\mu)$ *and* $\mathcal{W}(\mathcal{B}, \mu)$.

As pointed out in [10], "usually these conditions are derived under normality (controllability) assumptions by means of a very complicated argument." On the other hand, as we shall show next, the augmentability approach yields these standard necessary conditions for a minimum for the problem of Lagrange in a much simpler way.

IV. THE AUGMENTABILITY APPROACH

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

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

where

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

Note that

$$\tilde{F}(t, x, \dot{x}, \mu) = F(t, x, \dot{x}, \mu, 1)$$
$$+ \frac{\sigma(t, x, \dot{x})}{2} |\varphi(t, x, \dot{x})|^2.$$

Associated with the integral I, consider the augmented integral, for all $((x, \mu) \in X \times \mathcal{U}_q$,

$$\tilde{J}(x, \mu) := \int_{t_0}^{t_1} \tilde{F}(t, x(t), \dot{x}(t), \mu(t)) dt$$

and denote by Q($\mathcal{A}, \mu, \sigma$) the unconstrained problem of minimizing $\tilde{J}(\cdot, \mu)$ over $X_e(\mathcal{A})$.

4.1 Definition: For any $x_0 \in X_e(\mathcal{B})$ we shall say that P($\mathcal{B}$) is *augmentable at* x_0 if there exist σ and μ such that x_0 solves Q($\mathcal{A}, \mu, \sigma$). Note that, in this event, x_0 solves P($\mathcal{B}$) since, for any $x \in X_e(\mathcal{B})$, we have

$$I(x_0) = \tilde{J}(x_0, \mu) \leq \tilde{J}(x, \mu) = I(x).$$

Let us now prove that an augmentability assumption allows us to derive the necessary conditions for a minimum given in Theorem 3.8.

4.2 Theorem: *Let* $x_0 \in X_e(\mathcal{B})$ *and suppose* P($\mathcal{B}$) *is augmentable at* x_0. *Then there exists* $\mu \in \mathcal{U}_q$ *such* *that x_0 belongs to $\mathcal{E}(\mu)$, $\mathcal{H}(\mu)$, $\mathcal{L}(\mu)$ and $\mathcal{W}(\mathcal{B},\mu)$. Moreover, x_0 solves* P($\mathcal{B}$).

Proof: Since P($\mathcal{B}$) is augmentable at x_0, there exist σ and μ such that x_0 solves Q($\mathcal{A},\mu,\sigma$). We apply Theorem 2.1 which states that x_0 belongs to $\mathcal{E}$, $\mathcal{H}$, $\mathcal{L}$ and $\mathcal{W}(\mathcal{A})$ with respect to the integral $\tilde{J}(\cdot,\mu)$, that is, x_0 belongs to the sets

$$\tilde{\mathcal{E}}(\mu,\sigma) := \{x \in X \mid \text{there exists } c \in \mathbf{R}^n \text{ such that}$$

$$\tilde{F}_{\dot{x}}(\tilde{x}(t),\mu(t)) = \int_{t_0}^{t} \tilde{F}_x(\tilde{x}(s),\mu(s))ds + c \quad (t \in T)\},$$

$$\tilde{\mathcal{H}}(\mu,\sigma) := \{x \in X \mid \tilde{J}''(x,\mu;y) \geq 0 \text{ for all } y \in Y\},$$

$$\tilde{\mathcal{L}}(\mu,\sigma) := \{x \in X \mid \tilde{F}_{\dot{x}\dot{x}}(\tilde{x}(t),\mu(t)) \geq 0 \ (t \in T)\},$$

$$\tilde{\mathcal{W}}(\mathcal{A},\mu,\sigma) := \{x \in X(\mathcal{A}) \mid E_{\tilde{F}}(t,x(t),\dot{x}(t),u,\mu(t)) \geq 0$$
$$\text{for all } (t,u) \in T \times \mathbf{R}^n \text{ with } (t,x(t),u) \in \mathcal{A}\}.$$

By the first contention, there exists $c \in \mathbf{R}^n$ such that

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

and therefore

$$F_{\dot{x}}(\tilde{x}_0(t),\mu(t),1) = \int_{t_0}^{t} F_x(\tilde{x}_0(s),\mu(s),1)ds + c$$

showing that $x_0 \in \mathcal{E}(\mu)$.

By the second contention, we have

$$\tilde{J}''(x_0,\mu;y) \geq 0$$

for all $y \in X$ such that $y(t_0) = y(t_1) = 0$ where

$$\tilde{J}''(x_0,\mu;y) = \int_{t_0}^{t_1} \{\langle y(t), \tilde{F}_{xx}(t)y(t)\rangle$$
$$+ 2\langle y(t), \tilde{F}_{x\dot{x}}(t)\dot{y}(t)\rangle + \langle \dot{y}(t), \tilde{F}_{\dot{x}\dot{x}}(t)\dot{y}(t)\rangle\}dt$$

and $\tilde{F}(t)$ denotes $\tilde{F}(\tilde{x}_0(t),\mu(t))$. As one readily verifies,

$$\tilde{J}''(x_0,\mu;y) = J''(x_0,\mu;y)$$
$$+ \int_{t_0}^{t_1} \sigma(\tilde{x}_0(t))|\varphi_x(\tilde{x}_0(t)y(t) + \varphi_{\dot{x}}(\tilde{x}_0(t))\dot{y}(t)|^2 dt$$

and therefore $J''(x_0,\mu;y) \geq 0$ for all $y \in Y(\mathcal{B},x_0)$, showing that $x_0 \in \mathcal{H}(\mu)$.

By the third contention, we have

$$\tilde{F}_{\dot{x}\dot{x}}(\tilde{x}_0(t),\mu(t)) \geq 0 \quad (t \in T).$$

Observe that

$$\tilde{F}_{\dot{x}\dot{x}}(\tilde{x}_0(t),\mu(t)) = F_{\dot{x}\dot{x}}(\tilde{x}_0(t),\mu(t),1)$$
$$+ \sigma(\tilde{x}_0(t))G_{\dot{x}\dot{x}}(\tilde{x}_0(t)).$$

Now, for any $x \in X$,

$$G_{\dot{x}}(\tilde{x}(t)) = \sum_1^q \varphi_\alpha(\tilde{x}(t))\varphi_{\alpha\dot{x}}(\tilde{x}(t))$$

and therefore, for any $h \in \mathbf{R}^n$,

$$\langle h, G_{\dot{x}\dot{x}}(\tilde{x}_0(t))h\rangle =$$
$$h^*\Big(\sum_1^q \varphi^*_{\alpha\dot{x}}(\tilde{x}_0(t))\varphi_{\alpha\dot{x}}(\tilde{x}_0(t))\Big)h =$$
$$\sum_1^q (\varphi_{\alpha\dot{x}}(\tilde{x}_0(t))h)^2 = |\varphi_{\dot{x}}(\tilde{x}_0(t))h|^2.$$

Consequently

$$\langle h, F_{\dot{x}\dot{x}}(\tilde{x}_0(t),\mu(t),1)h\rangle$$
$$+ \sigma(\tilde{x}_0(t))|\varphi_{\dot{x}}(\tilde{x}_0(t))h|^2 \geq 0 \quad (h \in \mathbf{R}^n,\ t \in T)$$

implying that $x_0 \in \mathcal{L}(\mu)$.

Finally, by the fourth contention, we have

$$\tilde{F}(t,x_0(t),u,\mu(t)) - \tilde{F}(\tilde{x}_0(t),\mu(t))$$
$$- \tilde{F}_{\dot{x}}(\tilde{x}_0(t)),\mu(t))(u - \dot{x}_0(t)) \geq 0$$

for all $(t,u) \in T \times \mathbf{R}^n$ with $(t,x_0(t),u) \in \mathcal{A}$. This implies that

$$E_F(\tilde{x}_0(t),u,\mu(t),1)$$
$$+ \frac{\sigma(t,x_0(t),u)}{2}|\varphi(t,x_0(t),u)|^2 \geq 0$$

and therefore x_0 belongs to $\mathcal{W}(\mathcal{B},\mu)$. ∎

V. SUFFICIENCY THROUGH AUGMENTABILITY

For sufficiency, let us enlarge the sets we are dealing with and denote by X' the space of all absolutely continuous functions mapping T to $\mathbf{R}^n$. The notations $X'(\mathcal{B})$ and $X'_e(\mathcal{B})$ have the obvious meanings. The subclass of X' of all arcs in X' having square integrable derivatives on T will be denoted by X''.

5.1 Definition: For any $x_0 \in X_e(\mathcal{B})$ we shall say that P($\mathcal{B}$) is *strongly augmentable at* x_0 if there exist $\epsilon > 0$, σ and μ such that x_0 solves

$$\text{Q}(T_0(x_0;\epsilon) \cap \mathcal{A},\mu,\sigma).$$

Clearly, in this event, x_0 is a strong minimum of P($\mathcal{B}$). Similarly, P($\mathcal{B}$) is *weakly augmentable at* x_0 if there exist $\epsilon > 0$, σ and μ such that x_0 solves

$$\mathrm{Q}(T_1(x_0;\epsilon) \cap \mathcal{A}, \mu, \sigma),$$

implying that x_0 is a weak minimum of P($\mathcal{B}$).

5.2 Definition: For all $x \in X'$ define $Y'(\mathcal{B}, x)$ as the set of all $y \in X''$ satisfying $y(t_0) = y(t_1) = 0$ and

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

Consider now the following sets (in terms of F):

$$\mathcal{H}'(\mu) := \{x \in X' \mid J''(x,\mu;y) > 0$$
$$\text{for all } y \in Y'(\mathcal{B},x),\ y \neq 0\},$$
$$\mathcal{L}'(\mu) := \{x \in X' \mid \langle h, F_{\dot{x}\dot{x}}(\tilde{x}(t),\mu(t),1)h\rangle > 0 \text{ for}$$
$$\text{all } h \in \mathbf{R}^n,\ h \neq 0 \text{ such that } \varphi_{\dot{x}}(\tilde{x}(t))h = 0\ (t \in T)\},$$
$$\mathcal{W}(\mathcal{B},\mu;\epsilon) := \{x_0 \in X'(\mathcal{B}) \mid$$
$$E_F(t,x,\dot{x},u,\mu(t),1) \geq 0$$
$$\text{for all } (t,x,\dot{x},u) \in T \times \mathbf{R}^{3n} \text{ with}$$
$$(t,x,\dot{x}) \in T_1(x_0;\epsilon) \cap \mathcal{B} \text{ and } (t,x,u) \in \mathcal{B}\}$$

together with the respective sets (in terms of $\tilde{F}$):

$$\tilde{\mathcal{H}}'(\mu,\sigma) := \{x \in X' \mid \tilde{J}''(x,\mu;y) > 0$$
$$\text{for all } y \in X'',\ y \neq 0 \text{ with } y(t_0) = y(t_1) = 0\},$$
$$\tilde{\mathcal{L}}'(\mu,\sigma) := \{x \in X' \mid \langle h, \tilde{F}_{\dot{x}\dot{x}}(\tilde{x}(t),\mu(t))h\rangle > 0$$
$$\text{for all } h \in \mathbf{R}^n,\ h \neq 0,\ (t \in T)\},$$
$$\tilde{\mathcal{W}}(\mathcal{A},\mu,\sigma;\epsilon) := \{x_0 \in X'(\mathcal{A}) \mid$$
$$E_{\tilde{F}}(t,x,\dot{x},u,\mu(t)) \geq 0$$
$$\text{for all } (t,x,\dot{x},u) \in T \times \mathbf{R}^{3n} \text{ with}$$
$$(t,x,\dot{x}) \in T_1(x_0;\epsilon) \text{ and } (t,x,u) \in \mathcal{A}\}.$$

Let us now state an auxiliary result which will be used to prove that the standard sufficient conditions for a weak minimum imply weak augmentability. The first statement of this result has been established by Reid [13] and the second follows from the fact that, under the assumptions of the lemma, the function $\tilde{J}''(x_0,\mu;\cdot)$ is lower semicontinuous (see [11]).

5.3 Lemma: *If* $x_0 \in \mathcal{L}'(\mu)$, *there exists* $\theta > 0$ *such that, if* $\sigma(t,x,\dot{x}) \geq \theta$, *then*

a. $x_0 \in \tilde{\mathcal{L}}'(\mu,\sigma)$.

b. *If* $\{y_q\} \subset X''$ *converges uniformly on* T *to* y_0 *then*

$$\liminf_{q\to\infty} \tilde{J}''(x_0,\mu;y_q) \geq \tilde{J}''(x_0,\mu;y_0).$$

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 σ for which the second variation with respect to $\tilde{F}$ is also positive. Some of the ideas of the proof are based on the theory developed in [7].

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

Proof: Define

$$\Phi(t,y,\dot{y}) := \varphi_x(\tilde{x}_0(t))y + \varphi_{\dot{x}}(\tilde{x}_0(t))\dot{y}$$
$$P(y) := J''(x_0,\mu;y),$$
$$Q(y) := \int_{t_0}^{t_1} |\Phi(t,y(t),\dot{y}(t))|^2 dt.$$

Since $x_0 \in \mathcal{H}'(\mu)$ 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 seen before,

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

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 $q \in \mathbf{N}$, if $\sigma(t,x,\dot{x}) \geq q$ we have $x_0 \notin \tilde{\mathcal{H}}'(\mu,\sigma)$. That is, for all $q \in \mathbf{N}$ there exists $y_q \in X''$ nonnull with $y_q(t_0) = y_q(t_1) = 0$ such that

$$P(y_q) + qQ(y_q) \leq \tilde{J}''(x_0,\mu;y_q) \leq 0 \qquad (1)$$

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

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

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

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

Obviously $y_0(t_0) = y_0(t_1) = 0$. By Lemma 5.3 there exists $\theta > 0$ such that

$$\liminf_{q\to\infty}\{P(y_q) + \theta Q(y_q)\} \geq P(y_0) + \theta Q(y_0). \quad (4)$$

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

$$\liminf_{q\to\infty} Q(y_q) \leq 0.$$

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

$$\liminf_{q\to\infty} Q(y_q) \geq Q(y_0) \geq 0.$$

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 $y_0 \not\equiv 0$. 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 5.3. Then, by (4), we have

$$\liminf_{q\to\infty} \tilde{J}''(x_0, \mu; y_q) =$$

$$\liminf_{q\to\infty}\{P(y_q) + \theta Q(y_q)\} \geq 0$$

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

$$0 = \liminf_{q\to\infty} \tilde{J}''(x_0, \mu; y_q) =$$

$$\liminf_{q\to\infty} \int_{t_0}^{t_1} \langle \dot{y}_q(t), \tilde{F}_{\dot{x}\dot{x}}(\tilde{x}_0(t), \mu(t))\dot{y}_q(t)\rangle dt. \quad (5)$$

Since, by Lemma 5.3, the last integrand is a positive definite form, there is a constant $c > 0$ such that

$$\langle h, \tilde{F}_{\dot{x}\dot{x}}(\tilde{x}_0(t), \mu(t))h\rangle \geq \langle h, ch\rangle \geq c|h|^2.$$

Consequently equation (5) implies that

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

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. ▮

5.5 Theorem: *Suppose* $x_0 \in X'_e(\mathcal{B}) \cap C^1$ *and* $\mu: T \to \mathbf{R}^q$ *is absolutely continuous. If* x_0 *belongs to*

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

then P($\mathcal{B}$) *is weakly augmentable at* x_0.

Proof: By Theorem 2.2, the conclusion that P($\mathcal{B}$) is weakly augmentable at x_0 will follow if we show that, for some function $\sigma(t, x, \dot{x})$, x_0 belongs to

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

Let us begin by showing that, for any σ, $x_0 \in \tilde{\mathcal{E}}(\mu, \sigma)$. Indeed, since $x_0 \in \mathcal{E}(\mu)$, there exists $c \in \mathbf{R}^n$ such that, for all $t \in T$,

$$F_{\dot{x}}(\tilde{x}_0(t), \mu(t), 1) = \int_{t_0}^{t} F_x(\tilde{x}_0(s), \mu(s), 1)ds + c$$

and therefore, for all $t \in T$,

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

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

As a second step, let us show that there exists $c > 0$ such that $\sigma(\tilde{x}_0(t)) > c$ $(t \in T) \Rightarrow x_0 \in \tilde{\mathcal{L}}'(\mu, \sigma)$. For all $t \in T$ and $h \in \mathbf{R}^n$, define

$$P(t, h) := \langle h, F_{\dot{x}\dot{x}}(\tilde{x}_0(t), \mu(t), 1)h\rangle,$$

$$Q(t, h) := |\varphi_{\dot{x}}(\tilde{x}_0(t))h|^2.$$

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 $c > 0$,

$$P(t, h) + cQ(t, h) > 0 \text{ for all } t \in T \text{ and } h \neq 0.$$

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

$$P(t_q, h_q) + qQ(t_q, h_q) \leq 0.$$

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 $x_0 \in \mathcal{L}'(\mu)$. Now, let $\sigma(t, x, \dot{x})$ be such that $\sigma(\tilde{x}_0(t)) > c$ $(t \in T)$. Hence

$$\langle h, \tilde{F}_{\dot{x}\dot{x}}(\tilde{x}_0(t), \mu(t))h\rangle =$$

$$P(t,h) + \sigma(\tilde{x}_0(t))Q(t,h) > 0$$

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

Finally the statement that there exists $\theta_0 > 0$ such that $\sigma(\tilde{x}_0(t)) \geq \theta_0$ $(t \in T) \Rightarrow x_0 \in \tilde{\mathcal{H}}'(\mu,\sigma)$ follows by Lemma 5.4. ▮

We end by stating the corresponding result for strong minima related to strong augmentability. It can be shown that the notion of strong augmentability, like that of weak augmentability, is implied by the standard sufficient conditions for a strong minimum and, therefore, in both cases, the notion of augmentability can also be seen as an alternative approach to establish sufficiency results.

5.6 Theorem: *Suppose* $x_0 \in X'_e(\mathcal{B}) \cap C^1$ *and* $\mu: T \to \mathbf{R}^q$ *is absolutely continuous. If, for some* $\epsilon > 0$, x_0 *belongs to*

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

then P($\mathcal{B}$) *is strongly augmentable at* x_0.

VI. Conclusions

In this paper we introduce the notion of *generalized Lagrangians*, based on the concept of augmentability, applicable to a Lagrange problem in the calculus of variations involving equality constraints.

Though it is well-known that, by adding penalty functions to constrained minimum problems in finite dimensional spaces, one deals with unconstrained problems for which the derivation of necessary and sufficient optimality conditions can be easily obtained and methods of multipliers for finding numerical solutions can be derived, this concept of augmentability has received little attention to other problems that lie beyond the finite dimensional case.

For the infinite dimensional problem studied in this paper, we show that the concept of augmentability can be successfully generalized. We introduce the notions of weak and strong augmentability for the Lagrange problem and, without the usual assumption of normality, we derive first and second order necessary conditions for local optimality. Moreover, we provide a simple proof of the fact that the standard sufficient conditions for a weak local minimum imply weak augmentability of the problem at the corresponding extremal. We also state the fact that the standard sufficiency conditions for a strong local minimum imply strong augmentability.

It is of interest to see if the main ideas of this paper can be generalized to optimal control problems, not only for the derivation of necessary and sufficient conditions, but also to obtain a method of multipliers for finding numerical solutions of such problems.

References

[1] D.P. Bertsekas, On penalty and multiplier methods for constrained minimization, *SIAM Journal on Control and Optimization* 14, 1976, pp. 216-235.

[2] D.P. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods*, Computer Science and Applied Mathematics, Academic Press, New York 1982.

[3] H.A. Biswas, Necessary conditions for optimal control problems with and without state constraints: a comparative study, *WSEAS Transactions on Systems and Control* 6, 2011, pp. 217-228.

[4] M.R. de Pinho and J.F. Rosenblueth, Mixed constraints in optimal control: an implicit function theorem approach, *IMA Journal of Mathematical Control and Information* 24, 2007, pp. 197-218.

[5] M. Fortin and R. Glowinski, *Augmented Lagrangian Methods. Applications to the Numerical Solution of Boundary-Value Problems*, North-Holland, Amsterdam 1983.

[6] R. Glowinski, *Numerical Methods for Nonlinear Variational Problems*, Springer Series in Computational Physics, Springer-Verlag, New York 1984.

[7] M.R. Hestenes, An indirect sufficiency proof for the problem of Bolza in nonparametric form, *Transactions of the American Mathematical Society* 62, 1947, pp. 509–535.

[8] M.R. Hestenes, *Calculus of Variations and Optimal Control Theory,* John Wiley & Sons, New York 1966.

[9] M.R. Hestenes, *Optimization Theory, The Finite Dimensional Case,* John Wiley & Sons, New York 1975.

[10] M.R. Hestenes, Augmentability in optimization theory, *Journal of Optimization Theory and Applications* 32, 1980, pp. 427–440.

[11] E.J. McShane, Sufficient conditions for a weak relative minimum in the problem of Bolza, *Transactions of the American Mathematical Society* 52, 1942, pp. 344-379.

[12] A. Miele, P.E. Moseley, A.V. Levy and G.M. Coggins, On the method of multipliers for mathematical programming problems, *Journal of Optimization Theory and Applications* 10, 1972, pp. 1-33.

[13] W.T. Reid, Isoperimetric problems of Bolza in nonparametric form, *Duke Mathematical Journal* 5, 1939, pp. 675–691.

[14] R.T. Rockafellar, The multiplier method of Hestenes and Powell applied to convex programming, *Journal of Optimization Theory and Applications* 12, 1973, pp. 555–562.

[15] J.F. Rosenblueth, Admissible variations for optimal control problems with mixed constraints, *WSEAS Transactions on Systems* 12, 2005, pp. 2204–2211.

[16] J.F. Rosenblueth, Augmentability in optimal control, *International Journal of Mathematics and Statistics* 5, 2009, pp. 102–109.

[17] J.F. Rosenblueth, Constrained optimization through augmentability, *10th WSEAS International Conference on Dynamical Systems and Control*, Iasi, Romania, WSEAS Press 2011, pp. 105-112.

[18] J.F. Rosenblueth, Augmented integrals for optimal control problems, *International Journal of Applied Mathematics and Informatics* 7, 2013, pp. 44–54.

[19] J.F. Rosenblueth, Augmentability for a constrained problem in the calculus of variations, *4th European Conference of Control*, Paris, France, WSEAS Press 2013, Recent Advances in Circuits, Systems, Telecommunications and Control, pp. 17-23.

[20] R.D. Rupp, A nonlinear optimal control minimization technique, *Transactions of the American Mathematical Society* 178, 1973, pp. 357–381.