J Glob Optim (2008) 40:41–50
DOI 10.1007/s10898-007-9171-z

Fields of extremals and sufficient conditions for the simplest problem of the calculus of variations

Dean A. Carlson · George Leitmann

Received: 24 May 2007 / Accepted: 28 May 2007 / Published online: 28 June 2007

Abstract In this paper we present an elementary proof of a classical sufficient condition for free problems in the calculus of variations using an extension of the direct sufficiency method, originating in the work of G. Leitmann (Int. J. NonLin. Mech. 2, 55–59 (1967)), found in D.A. Carlson (J. Optim. Theory Appl. 114, 345–362 (2002)).

Keywords Field of extremals · Sufficient conditions · Direct method

1 Introduction

Much of the theory of the calculus of variations, both past and present, focuses on the search for necessary conditions for optimality. For the most part these results have changed very little through the years, except for weakening smoothness conditions, differentiability notions, and the inclusions of varieties of constraints. Indeed, the seminal work on nonsmooth calculus of variations in the 1970s clearly indicates that the classical theory (properly interpreted of course) encompassed much of the modern theory, modulo the introduction of nonsmoothness. Necessary conditions however only provide candidates for solution and many examples exist where solutions of the necessary conditions are not optimal. The difficulty was known quite early in the history of the subject and it was well known to Lagrange and his contemporaries that the solution to the problem rested with the nonnegative definiteness of the second variation. Numerous attempts at characterizing the second variation are reported in the book by Todhunter [12], a historical treatise which discusses the major works in the calculus of variations in the first half of the nineteenth century. After the publication of Todhunter's treatise, the work of Weierstrass in the calculus of variations was presented. It is here that for the first time the role of weak and strong neighborhoods was recognized and the first

D. A. Carlson (✉)
American Mathematical Society, Mathematical Reviews, 416 Fourth Street, Ann Arbor, MI 48103, USA
e-mail: dac@ams.org

G. Leitmann
University of California at Berkeley, Berkeley, CA 94720, USA

distinctions between weak and strong minimizers was made. It is also here that the first general, systematic theory of sufficient conditions was initiated and fields of extremals were rigorously discussed. Of course, earlier work by Jacobi and Hamilton in the 1830s laid the groundwork for Weierstrass's contributions. Weierstrass's work was followed by the work of Hilbert, Bolza, Carathéodory, and others in the late nineteenth and early twentieth centuries.

Although there has been a number of excellent textbooks written for the calculus of variations that include all aspects of the theory, most university courses, due to the time constraints of one term, primarily discuss necessary conditions and give little or no information regarding sufficient conditions. In this paper our goal is to provide an elementary approach to these sufficient conditions and show that they may be proved via Leitmann's direct sufficiency method. This direct method, first presented in Leitmann [9], considers coordinate transformations to transform the original problem into one which can be solved more easily— often by inspection. The key feature of the method is a fundamental identity which relates the two problems. In the late 1990s, Leitmann returned to this method and presented a series of papers through which he and his colleagues were able to broaden the class of problems considered to infinite horizon optimization problems and to the study of open-loop games. In Carlson [3] (see also [11]) an extension of Leitmann's method was made that incorporated Carathéodory's method of equivalent problems and in the process expanded the utility of the method. Further updates and insights to the method subsequently followed in a series of papers due to both authors (see, e.g., [4–7]).

The outline of our paper is that we first introduce the problem we wish to consider and review the classical necessary conditions. In the second section we discuss fields of extremals. The third section briefly summarizes the direct sufficiency method we wish to use. Next we give a new proof of a classical sufficiency condition found in the calculus of variations. We conclude our paper by indicating further directions to explore in exploiting our method to broader classes of problems.

2 The free problem of lagrange and basic necessary conditions for optimality

We consider the simplest problem in the calculus of variations, namely that of minimizing an integral functional

$$J(x(\cdot)) = \int_a^b L(t, x(t), \dot{x}(t))\mathrm{d}t \tag{1}$$

over all admissible trajectories (see below) $x(\cdot)$: $[a, b] \to \mathbb{R}$ that are continuous on $[a, b]$ with piecewise continuous derivatives, (i.e., $x(\cdot) \in PC^1[a, b]$) satisfying a set of fixed end conditions

$$x(a) = x_a \quad \text{and} \quad x(b) = x_b. \tag{2}$$

Regarding the integrand $L(\cdot, \cdot, \cdot)$ assume that $A \subset [a, b] \times \mathbb{R}^2$ is a given open set and that $L(\cdot, \cdot, \cdot)$: $A \to \mathbb{R}$ is continuous and such that for each $t \in [a, b]$ we have $(x, p) \to L(t, x, p)$ is twice continuously differentiable on $A(t) = \{(x, p) \in \mathbb{R}^2 : (t, x, p) \in A\}$ and that for each $(t, x) \in [a, b] \times \mathbb{R}$ we have that $p \to L(t, x, p)$ is a convex function on $A(t, x) = \{p : (t, x, p) \in A\}$. With this notation we have the following definition.

Definition 1 A function $x(\cdot)$: $[a, b] \to \mathbb{R}$ is called an admissible trajectory if $x(\cdot)$ is continuous with piecewise continuous derivatives on $[a, b]$, $(t, x(t), \dot{x}(t)) \in A$, it satisfies the end conditions (2), and $J(x(\cdot))$ is finite. The set of all admissible trajectories will be denoted by Ω.

For brevity, we refer to the problem of minimizing the functional (1) over the set of admissible trajectories Ω as problem (P). We now recall the following classical first order necessary conditions for (P), which are readily found in almost any textbook on the calculus of variations (see for example Ewing [8]).

Theorem 1 *(*Euler–Lagrange*) If $x^*(\cdot) \in \Omega$ is at least a weak local minimizer of (1), then the Euler-Lagrange equation*

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial p}\left(t, x^*(t), \dot{x}^*(t)\right) = \frac{\partial L}{\partial x}\left(t, x^*(t), \dot{x}^*(t)\right), \quad t \in [a, b] \tag{3}$$

holds, where if $t \in [a, b]$ is a point of discontinuity for $\dot{x}^*(\cdot)$ the above equation holds in the usual sense of one-sided limits.

In this paper we show how to obtain elementary proofs of some of a classical sufficient condition for free problems in the calculus of variations using an extension of the direct sufficiency method, originating in the work of Leitmann [9], found in Carlson [3].

Theorem 2 (Legendre) *If $x^*(\cdot) \in \Omega$ is at least a weak local minimizer of (1), then*

$$\frac{\partial^2 L}{\partial p^2}(t, x^*(t), \dot{x}^*(t)) \geq 0, \quad t \in [a, b]. \tag{4}$$

Theorem 3 (Weierstrass) *If $x^*(\cdot) \in \Omega$ is either a strong local or a global minimizer of (1), then*

$$\begin{aligned} E(t, x^*(t), \dot{x}^*(t), q) &= L(t, x^*(t), q) - L(t, x^*(t), \dot{x}^*(t)) \\ &\quad -(q - \dot{x}^*(t))\frac{\partial L}{\partial p}(t, x^*(t), \dot{x}^*(t)) \geq 0 \end{aligned} \tag{5}$$

for all $t \in [a, b]$ and all $q \in \mathbb{R}$ such that $(t, x^*(t), q) \in A$.

To present the last necessary condition we introduce the *accessory minimum problem*. To do this suppose that $x^*(\cdot) \in \Omega$ is a smooth (i.e., continuously differentiable) minimizer of (1) and consider the problem of minimizing the functional

$$\begin{aligned} I(\eta(\cdot)) = \int_a^b \Bigg[&\frac{\partial^2 L}{\partial x^2}(t, x^*(t), \dot{x}^*(t))\eta(t)^2 + 2\frac{\partial^2 L}{\partial x \partial p}((t, x^*(t), \dot{x}^*(t))\eta(t)\dot{\eta}(t) \\ &+ \frac{\partial^2 L}{\partial p^2}(t, x^*(t), \dot{x}^*(t))\dot{\eta}(t)^2 \Bigg] \mathrm{d}t \end{aligned} \tag{6}$$

over all $\eta(\cdot) \in PC^1[a, b]$ satisfying the end conditions

$$\eta(a) = 0 \quad \text{and} \quad \eta(b) = 0. \tag{7}$$

With regards to this accessory minimum problem we have the following definition.

Definition 2 If u: $[a, c] \to \mathbb{R}$ satisfies the Euler–Lagrange equation for the accessory minimum problem and the end conditions $u(a) = u(c) = 0$ with $u(t) \neq 0$ for $t \in (a, c)$ we say c is a conjugate value to a and that the point $(c, x^*(c))$ is called a conjugate point to the initial point $(a, x^*(a))$ of $x^*(\cdot)$.

Theorem 4 (Jacobi) *If $x^*(\cdot) \in \Omega$ is a smooth minimizer of (1) and if the Legendre necessary condition holds with strict inequality (i.e., the so-called strengthened Legendre condition) on $[a, b]$ then there is no value $c \in (a, b)$ that is conjugate to a.*

This concludes our brief reminder of the classical necessary conditions. In the next section begin our discussion of sufficient conditions.

3 Fields of extremals

The notation and terminology for a field of extremals is not standard in the classical theory. The approach we adopt here is found in Bolza [1] beginning with the following definition.

Definition 3 Let $\xi(\cdot, \cdot)$: $[a, b] \times \mathbb{R} \to \mathbb{R}$ be a twice continuously differentiable function. We say that $\xi(\cdot, \cdot)$ is a family of extremals for problem (P) if for each constant $\beta \in \mathbb{R}$ one has that the function $t \to \xi(t, \beta)$ satisfies the Euler–Lagrange equations (3), namely

$$\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial p} \left(t, \xi(t, \beta), \frac{\partial \xi}{\partial t}(t, \beta) \right) = \frac{\partial L}{\partial x} \left(t, \xi(t, \beta), \frac{\partial \xi}{\partial t}(t, \beta) \right), \quad t \in [a, b], \tag{8}$$

and that $(t, \xi(t, \beta), \frac{\partial \xi}{\partial t}(t, \beta)) \in A$ for $t \in [a, b]$.

Remark 1 In the above definition, the boundary conditions of problem (P) are not necessarily satisfied.

To solve (P), one naturally makes an effort to solve one or more of the corresponding necessary conditions to obtain candidates of optimality. For the case considered here, these necessary conditions are described in the previous section. Therefore, we assume from now on that $x^*(\cdot)$ satisfies the Euler–Lagrange equations (3) as well as the fixed end conditions (2). Now suppose we are given a family of extremals for which $x^*(\cdot) = \xi(\cdot, \beta^*)$ for some value $\beta = \beta^*$. For $k > 0$ let $S_k = \{(t, x) = (t, \xi(t, \beta)) \colon (t, \beta) \in [a, b] \times [\beta^* - k, \beta^* + k]\}$. With this notation we have the following definition.

Definition 4 The set S_k is called a field of extremals about the trajectory $x^*(\cdot)$ if for every point $(t, x) \in S_k$ there exists exactly one $\beta \in [\beta^* - k, \beta^* + k]$ such that $(t, x) = (t, \xi(t, \beta))$.

Remark 2 Observe that the set S_k can also be viewed as a "strip" in the (t, x)-plane swept out by the extremals $t \to \xi(t, \beta)$ as β increases from $\beta^* - k$ to $\beta^* + k$. Moreover, the condition that S_k be a field means that there exists a single-valued function $\psi(\cdot, \cdot)$ such that for each $(t, x) \in S_k$ one has

$$\begin{aligned} \beta &= \psi(t, x), \\ x &= \xi(t, \psi(t, x)) \end{aligned}$$

and

$$|\psi(t, x) - \beta^*| \le k.$$

With regards to the above remark Bolza presents the following theorem which is a direct consequence of the implicit function theorem.

Theorem 5 *For a given family of extremals suppose there exists β^* such that $x^*(\cdot) = \xi(\cdot, \beta^*)$. Then if*

$$\frac{\partial \xi}{\partial \beta}(t, \beta^*) \neq 0 \quad \text{for all} \quad t \in [a, b]$$

there exists $k > 0$ sufficiently small so that S_k is a field of extremals about $x^*(\cdot)$.

4 Leitmann's direct method

In 1967 Leitmann [9] introduced a direct sufficiency method that involved a transformation of coordinates which permitted him to deduce the solution of a calculus of variations problem by inspection. During the 1990s, he revived this method extending it to variational games and applying it to a number of problems arising in economic modeling. In Carlson [3], Leitmann's method was compared and contrasted with Carathéodory's notion of equivalent variational problem and finally combined into a new and improved version of Leitmann's direct method. It is this improved version which we now describe. To do this we let $x = z(t, \tilde{x})$ be a transformation of class C^1 having a unique inverse $\tilde{x} = \tilde{z}(t, x)$ for all $t \in [a, b]$ such that there is a one-to-one correspondence $x(t) \Leftrightarrow \tilde{x}(t)$ for all admissible trajectories $x(\cdot)$ satisfying the boundary conditions (2) and for all $\tilde{x}(\cdot)$ satisfying

$$\tilde{x}(a) = \tilde{z}(a, x_a) \quad \text{and} \quad \tilde{x}(b) = \tilde{z}(b, x_b). \tag{9}$$

Observe that if $x(\cdot) \in \Omega$ then we have that there exists a trajectory $\tilde{x}(\cdot)$ satisfying (9) and such that

$$(t, x(t), \dot{x}(t)) = \left(t, z(t, \tilde{x}(t)), \frac{\partial z}{\partial t}(t, \tilde{x}(t)) + \frac{\partial z}{\partial \tilde{x}}(t, \tilde{x}(t))\dot{\tilde{x}}(t)\right) \quad \text{for} \quad t \in [a, b].$$

This means that the trajectories $\tilde{x}(\cdot)$ are such that $(t, \tilde{x}(t), \dot{\tilde{x}}(t)) \in \tilde{A}$ holds for all $t \in [a, b]$ in which

$$\tilde{A} = \left\{(t, \tilde{x}, \tilde{p}) : \left(t, z(t, \tilde{x}), \frac{\partial z}{\partial t}(t, \tilde{x}) + \frac{\partial z}{\partial \tilde{x}}(t, \tilde{x})\tilde{p}\right) \in A\right\}. \tag{10}$$

Now let $\tilde{L}(\cdot, \cdot, \cdot)\colon \tilde{A} \to \mathbb{R}$ be another integrand that enjoys the same properties as $L(\cdot, \cdot, \cdot)$ and consider the problem ($\tilde{\text{P}}$) of minimizing the integral

$$\tilde{J}(\tilde{x}(\cdot)) = \int_a^b \tilde{L}(t, \tilde{x}(t), \dot{\tilde{x}}(t))\, \mathrm{d}t$$

over all trajectories $\tilde{x}(\cdot) \in PC^1[a, b]$ satisfying the end conditions (9) and satisfying

$$(t, \tilde{x}(t), \dot{\tilde{x}}(t)) \in \tilde{A}$$

for $t \in [a, b]$. With this notation we have the following definition.

Definition 5 We say the problems (P) and ($\tilde{\text{P}}$) are equivalent if whenever $x^*(\cdot)$ is a minimizer of (P) the function $\tilde{x}^*(\cdot) = \tilde{z}(\cdot, x^*(\cdot))$ is a minimizer of ($\tilde{\text{P}}$), and conversely.

We now give Leitmann's direct sufficiency method.

Theorem 6 *Let $z(\cdot, \cdot)$, $L(\cdot, \cdot, \cdot)$ and $\tilde{L}(\cdot, \cdot, \cdot)$ satisfy the hypotheses described above. If there exists a C^1 function $G(\cdot, \cdot)\colon [a, b] \times \mathbb{R} \to \mathbb{R}$ such that for all admissible trajectories $x(\cdot)\colon [a, b] \to \mathbb{R}$ for (P) the functional identity*

$$L(t, x(t), \dot{x}(t)) - \tilde{L}(t, \tilde{x}(t), \dot{\tilde{x}}(t)) = \frac{\mathrm{d}}{\mathrm{d}t} G(t, \tilde{x}(t)) \tag{11}$$

holds for all $t \in [a, b]$, then the problems (P) and ($\tilde{P}$) are equivalent.

Proof Leitmann [11]. □

Two immediate and useful corollaries are the following.

Corollary 1 *The existence of $G(\cdot,\cdot)$ in (11) implies that the following identity holds for $(t,\tilde{x},\tilde{p}) \in (a,b)\times\mathbb{R}\times\mathbb{R}$:*

$$L\left(t, z(t,\tilde{x}), \frac{\partial z(t,\tilde{x})}{\partial t} + \frac{\partial z(t,\tilde{x})}{\partial \tilde{x}}\tilde{p}\right) - \tilde{L}(t,\tilde{x},\tilde{p}) \equiv \frac{\partial G(t,\tilde{x})}{\partial t} + \frac{\partial G(t,\tilde{x})}{\partial \tilde{x}}\tilde{p}. \tag{12}$$

Corollary 2 *The left-hand side of the identity, (12) is linear in $\tilde{p}$, that is, it is of the form,*

$$\theta(t,\tilde{x}) + \psi(t,\tilde{x})\tilde{p}$$

and

$$\frac{\partial G(t,\tilde{x})}{\partial t} = \theta(t,\tilde{x}) \quad \textit{and} \quad \frac{\partial G(t,\tilde{x})}{\partial \tilde{x}} = \psi(t,\tilde{x})$$

on $[a,b]\times\mathbb{R}$.

Remark 3 The utility of the above theorem rests in being able to choose not only the transformation $z(\cdot,\cdot)$ but also the integrand $\tilde{L}(\cdot,\cdot,\cdot)$. In the next section we use this direct method to give an elementary proof of a classical sufficient condition for the problems considered. We further observe that the condition that $p \to L(t,x,p)$ is convex is not required to obtain the above results.

Remark 4 The above theorem is an extension of Leitmann's original idea and includes it as a special case by taking $\tilde{L}(\cdot,\cdot,\cdot) = L(\cdot,\cdot,\cdot)$; see [10]. In addition, it includes the notion of equivalent variational problem found in Carathéodory's work (e.g., see Carathéodory's book [2]) by taking $z(\cdot,\cdot)$ to be the identity transformation (i.e., $z(t,\tilde{x}) \equiv \tilde{x}$). We further remark that these ideas are also closely related to the canonical transformations found in classical mechanics originally investigated by Hamilton and Jacobi in the 1830s.

5 The classical sufficient conditions

We begin by letting $\xi(\cdot,\cdot)$ be a family of extremals as defined in Definition 3 and letting $x^*(\cdot)$ be an extremal satisfying the conditions:

(a) The transformation $x = \xi(t,\beta)$ is of class C^2 with a unique inverse $\beta = \tilde{\xi}(t,x)$.
(b) There exists a constant number $\beta^* \in \mathbb{R}$ such that $x^*(t) = \xi(t,\beta^*)$ for $t \in [a,b]$.

Remark 5 A family of extremals satisfying the conditions indicated above defines a one-to-one correspondence between the admissible trajectories of (P) and a set of functions $\tilde{x}(\cdot)$ satisfying the boundary conditions

$$\tilde{x}(a) = \tilde{\xi}(a,x_a) = \beta^* \quad \text{and} \quad \tilde{x}(b) = \tilde{\xi}(b,x_b) = \beta^*.$$

To see this we notice that if $x(\cdot)$ is admissible for the original problem then we have that the function $\tilde{x}(\cdot)\colon [a,b] \to \mathbb{R}$ defined by $\tilde{x}(t) = \tilde{\xi}(t,x(t))$ satisfies the fixed end conditions

$$\tilde{x}(a) = \tilde{\xi}(a,x_a) \quad \text{and} \quad \tilde{x}(b) = \tilde{\xi}(b,x_b).$$

Further, since $\beta^* = \tilde{\xi}(t, x^*(t))$ for all $t \in [a, b]$ it follows immediately that $\beta^* = \tilde{\xi}(a, x_a) = \tilde{\xi}(b, x_b)$ as desired. Thus, as we will see below, a family of extremals satisfying the above makes a good candidate for a transformation to which the direct method may applied.

We begin with the following technical lemma.

Lemma 1 *Let $L(\cdot, \cdot, \cdot)$ satisfy the hypotheses of Sect. 2 and let $\xi(\cdot, \cdot)$ be a family of extremals as defined in Definition 3. Then there exists a continuously differentiable function $G(\cdot, \cdot)$: $[a, b] \times \mathbb{R} \to \mathbb{R}$ and a continuous function $\gamma(\cdot, \cdot, \cdot)$: $[a, b] \times \mathbb{R} \times \mathbb{R} \to [0, 1]$ such that*

$$L\left(t, \xi(t, \beta), \frac{\partial \xi}{\partial t}(t, \beta) + \frac{\partial \xi}{\partial \beta}(t, \beta)\tilde{p}\right) - \tilde{L}(t, \beta, \tilde{p}) = \frac{\partial G}{\partial t}(t, \beta) + \frac{\partial G}{\partial \beta}(t, \beta)\tilde{p} \quad (13)$$

in which $\tilde{L}(\cdot, \cdot, \cdot)$: $[a, b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is given by

$$\tilde{L}(t, \beta, \tilde{p}) = \frac{1}{2}\frac{\partial^2 L}{\partial p^2}\left(t, \xi(t, \beta), \frac{\partial \xi}{\partial t}(t, \beta) + \gamma(t, \beta, \tilde{p})\frac{\partial \xi}{\partial \beta}(t, \beta)\tilde{p}\right)\left(\frac{\partial \xi}{\partial \beta}(t, \beta)\right)^2 \tilde{p}^2. \quad (14)$$

Proof We begin by applying Taylor's theorem with remainder to $p \to L(t, \xi(t, \beta), p)$.

Specifically we have that for each $t \in [a, b]$, $\beta \in \mathbb{R}$, and $\tilde{p} \in \mathbb{R}$ there exists a constant $\gamma(t, \beta, \tilde{p}) \in [0, 1]$ such that

$$\begin{aligned} &L\left(t, \xi(t, \beta), \frac{\partial \xi}{\partial t}(t, \beta) + \frac{\partial \xi}{\partial \beta}(t, \beta)\tilde{p}\right) \\ &\quad = L\left(t, \xi(t, \beta), \frac{\partial \xi}{\partial t}(t, \beta)\right) + \frac{\partial L}{\partial p}\left(t, \xi(t, \beta), \frac{\partial \xi}{\partial t}(t, \beta)\right)\frac{\partial \xi}{\partial \beta}(t, \beta)\tilde{p} \qquad (15) \\ &\qquad + \frac{1}{2}\frac{\partial^2 L}{\partial p^2}\left(t, \xi(t, \beta), \frac{\partial \xi}{\partial t}(t, \beta) + \gamma(t, \beta, \tilde{p})\frac{\partial \xi}{\partial \beta}(t, \beta)\tilde{p}\right)\left(\frac{\partial \xi}{\partial \beta}(t, \beta)\right)^2 \tilde{p}^2 \\ &\quad = L\left(t, \xi(t, \beta), \frac{\partial \xi}{\partial t}(t, \beta)\right) + \frac{\partial L}{\partial p}\left(t, \xi(t, \beta), \frac{\partial \xi}{\partial t}(t, \beta)\right)\frac{\partial \xi}{\partial \beta}(t, \beta)\tilde{p} + \tilde{L}(t, \beta, \tilde{p}). \end{aligned}$$

That is we have,

$$\begin{aligned} &L\left(t, \xi(t, \beta), \frac{\partial \xi}{\partial t}(t, \beta) + \frac{\partial \xi}{\partial \beta}(t, \beta)\tilde{p}\right) - \tilde{L}(t, \beta, \tilde{p}) \\ &\quad = L\left(t, \xi(t, \beta), \frac{\partial \xi}{\partial t}(t, \beta)\right) + \frac{\partial L}{\partial p}\left(t, \xi(t, \beta), \frac{\partial \xi}{\partial t}(t, \beta)\right)\frac{\partial \xi}{\partial \beta}(t, \beta)\tilde{p}. \qquad (16) \end{aligned}$$

Now define the two functions $\theta(\cdot, \cdot)$: $[a, b] \times \mathbb{R} \to \mathbb{R}$ and $\psi(\cdot, \cdot)$: $[a, b] \times \mathbb{R} \to \mathbb{R}$ by the formulas,

$$\theta(t, \beta) = L\left(t, \xi(t, \beta), \frac{\partial \xi}{\partial t}(t, \beta)\right) \quad \text{and} \quad \psi(t, \beta) = \frac{\partial}{\partial p}L\left(t, \xi(t, \beta), \frac{\partial \xi}{\partial t}(t, \beta)\right)\frac{\partial \xi}{\partial \beta}(t, \beta)$$

and notice that we may write (16) as

$$L\left(t, \xi(t, \beta), \frac{\partial \xi}{\partial t}(t, \beta) + \frac{\partial \xi}{\partial \beta}(t, \beta)\tilde{p}\right) - \tilde{L}(t, \beta, \tilde{p}) = \theta(t, \beta) + \psi(t, \beta)\tilde{p}.$$

Therefore to achieve our result it suffices to prove that there exists a function $G(\cdot, \cdot)$: $[a, b] \times \mathbb{R} \to \mathbb{R}$ such that

$$\theta(t, \beta) = \frac{\partial G}{\partial t}(t, \beta) \quad \text{and} \quad \psi(t, \beta) = \frac{\partial G}{\partial \beta}(t, \beta).$$

From a well known result of classical analysis a necessary and sufficient condition for this to occur is to show that

$$\frac{\partial}{\partial \beta}\theta(t,\beta) = \frac{\partial^2 G}{\partial \beta \partial t}(t,\beta) = \frac{\partial^2 G}{\partial t \partial \beta}(t,\beta) = \frac{\partial}{\partial t}\psi(t,\beta).$$

To see that this is the case we observe that since $\xi(\cdot,\cdot)$ is a family of extremals we have

$$\begin{aligned}
\frac{\partial}{\partial \beta}\theta(t,\beta) &= \frac{\partial L}{\partial x}\left(t,\xi(t,\beta),\frac{\partial \xi}{\partial t}(t,\beta)\right)\frac{\partial \xi}{\partial \beta}(t,\beta)\\
&\quad + \frac{\partial L}{\partial p}\left(t,\xi(t,\beta),\frac{\partial \xi}{\partial t}(t,\beta)\right)\frac{\partial^2 \xi}{\partial \beta \partial t}(t,\beta)\\
&= \frac{\partial}{\partial t}\left[\frac{\partial L}{\partial p}\left(t,\xi(t,\beta),\frac{\partial \xi}{\partial t}(t,\beta)\right)\right]\frac{\partial \xi}{\partial \beta}(t,\beta)\\
&\quad + \frac{\partial L}{\partial p}\left(t,\xi(t,\beta),\frac{\partial \xi}{\partial t}(t,\beta)\right)\frac{\partial^2 \xi}{\partial \beta \partial t}(t,\beta)\\
&= \frac{\partial}{\partial t}\left[\frac{\partial}{\partial p}L\left(t,\xi(t,\beta),\frac{\partial \xi}{\partial t}(t,\beta)\right)\right]\frac{\partial \xi}{\partial \beta}(t,\beta)\\
&\quad + \frac{\partial}{\partial p}L\left(t,\xi(t,\beta),\frac{\partial \xi}{\partial t}(t,\beta)\right)\frac{\partial^2 \xi}{\partial t \partial \beta}(t,\beta)\\
&= \frac{\partial}{\partial t}\psi(t,\beta),
\end{aligned}$$

where we have used the fact that the family of extremals is twice continuously differentiable so that we know

$$\frac{\partial^2 \xi}{\partial t \partial \beta}(t,\beta) = \frac{\partial^2 \xi}{\partial \xi \partial t}(t,\beta).$$

From this we can conclude that there exists a function $G(\cdot,\cdot)\colon [a,b]\times\mathbb{R}\to\mathbb{R}$ such that

$$L\left(t,\xi(t,\beta),\frac{\partial \xi}{\partial t}(t,\beta)+\frac{\partial \xi}{\partial \beta}(t,\beta)\tilde{p}\right) - \tilde{L}(t,\beta,\tilde{p}) = \frac{\partial G}{\partial t}(t,\beta) + \frac{\partial G}{\partial \beta}(t,\beta)\tilde{p}.$$

That is, Eq. 13 holds. □

Theorem 7 *In addition to the hypotheses given in Sect. 2, assume that the function $p \to L(t,x,p)$ is convex on $A(t,x) = \{p \in \mathbb{R}\colon (x,p) \in A(t)\}$. Further let $x^*(\cdot)$ be an admissible trajectory for problem (P) which satisfies the Euler–Lagrange equation (3) and assume that there exists a family of extremals, $\xi(\cdot,\cdot)$ satisfying (a) and (b). Then $x^*(\cdot)$ is a solution of (P).*

Proof Using the field of extremals, $\xi(\cdot,\cdot)$ define the transformations $z(\cdot,\cdot)\colon S_k \to \mathbb{R}$ by $z(t,\tilde{x}) = \xi(t,\tilde{x})$. Observe that by (a) this transformation is C^2 and invertible with inverse $\tilde{z}(\cdot,\cdot)$ given by $\tilde{z}(t,x) = \tilde{\xi}(t,x)$. Thus, this transformation defines a one-to-one correspondence between the admissible trajectories of problem (P) and the trajectories $\tilde{x}(\cdot)\colon [a,b]\to\mathbb{R}$ defined, for each admissible trajectory $x(\cdot)$ for (P), by $\tilde{x}(t) = \tilde{z}(t,x(t))$. Moreover, as observed in Remark 5, we also have that each such trajectory $\tilde{x}(\cdot)$ satisfies the fixed endpoint conditions

$$\tilde{x}(a) = \beta^* \quad \text{and} \quad \tilde{x}(b) = \beta^*.$$

Moreover, as a result of Lemma 1 we have for each admissible trajectory $x(\cdot)$ for (P) and corresponding trajectory $\tilde{x}(\cdot) = \tilde{z}(\cdot, x(\cdot))$ that

$$\begin{aligned} L(t, x(t), \dot{x}(t)) - \tilde{L}(t, \tilde{x}(t), \dot{\tilde{x}}(t)) &= L\left(t, z(t, \tilde{x}(t)), \frac{\partial z}{\partial t}(t, \tilde{x}(t)) + \frac{\partial z}{\partial \tilde{x}}(t, \tilde{x}(t))\dot{\tilde{x}}(t)\right) \\ &\quad - \tilde{L}(t, \tilde{x}(t), \dot{\tilde{x}}(t)) \\ &= \frac{\partial G}{\partial t}(t, \tilde{x}(t)) + \frac{\partial G}{\partial \tilde{x}}(t, \tilde{x})\dot{\tilde{x}}(t) \\ &= \frac{d}{dt} G(t, \tilde{x}(t)), \end{aligned}$$

which is precisely the condition required in the Theorem 6. Therefore, by applying Theorem 6 we know that the minimizers of problem (P) are in one-to-one correspondence with the minimizers of the problem ($\tilde{\mathrm{P}}$) consisting of minimizing the integral functional

$$\tilde{J}(\tilde{x}(\cdot)) = \int_a^b \tilde{L}(t, \tilde{x}(t), \dot{\tilde{x}}(t))\, \mathrm{d}t,$$

in which $\tilde{L}(\cdot, \cdot, \cdot)$ is given by (14), over all piecewise continuously differentiable trajectories $\tilde{x}(\cdot)$ satisfying $(t, \tilde{x}(t), \dot{\tilde{x}}(t)) \in \tilde{A}$ for all $t \in [a, b]$ (where $\tilde{A}$ is given by (10)), satisfying the fixed endpoint conditions

$$\tilde{x}(a) = \beta^* \quad \text{and} \quad \tilde{x}(b) = \beta^*.$$

Observe that since $L(\cdot, \cdot, \cdot)$ is convex in its last argument we have $(t, \tilde{x}, \tilde{p}) \to \tilde{L}(t, \tilde{x}, \tilde{p})$ is nonnegative and assumes the value zero whenever $\tilde{p} = 0$. This means that $\tilde{x}^*(t) \equiv \beta^*$ is a minimizer of ($\tilde{\mathrm{P}}$). Thus an application of the direct method gives us that the trajectory $x^*(\cdot) \equiv \xi(t, \beta^*)$ is an optimal solution for (P) as desired. □

Remark 6 The minimum in the above theorem is a strong local minimum and becomes a global (absolute) minimum whenever S_k is the infinite strip $[a, b] \times \mathbb{R}$.

Remark 7 It is possible that the auxiliary problem has additional minimizers. In particular, it is possible that there exists a nonconstant trajectory $\tilde{x}(\cdot)$ that is admissible for ($\tilde{\mathrm{P}}$) such that $\tilde{L}(t, \tilde{x}(t), \dot{\tilde{x}}(t)) \equiv 0$ on $[a, b]$. For example, this would be true if one had

$$\frac{\partial z}{\partial \tilde{x}}(t, \tilde{x}(t)) = 0 \quad \text{for all } t \in [a, b].$$

What is important is that the constant trajectory $\tilde{x}^*(t) \equiv \beta^*$ is indeed a minimizer.

6 Conclusions

In this paper we have used Leitmann's direct sufficiency method to give a new elementary proof of a classical sufficient condition in the Calculus of Variations. Additionally, it is clear that the fundamental result, Theorem 6, includes the basic ideas of canonical transformations and the notion of equivalent variational problems as defined in Carathéodory [2]. Some interesting questions remain open. The first is to determine if this approach can be applied to an example in which the $\tilde{L}(\cdot, \cdot, \cdot)$ is not "quadratic-like." This would genuinely show that this approach leads to something new and extends further than the classical field theory. A second question would be that of generalizing the approach to include nonsmoothness either in the transformations or in the integrands. Other questions would be those of including constraints, extensions to optimal control, etc.

References

1. Bolza, O.: Calculus of Variations. 3rd edn. Chelsea Publishing Company, New York, NY (1973)
2. Carathéodory, C.: Calculus of Variations and Partial Differential Equations. Chelsea, New York, NY (1982)
3. Carlson, D.A.: An observation on two methods of obtaining solutions to variational problems. J. Optim. Theory Appl. **114**, 345–362 (2002)
4. Carlson, D.A., Leitmann, G.: An extension of the coordinate transformation method for open-loop Nash equilibria. J. Optim. Theory Appl. **123**, 27–47 (2004)
5. Carlson, D.A., Leitmann, G.: A coordinate transformation method for the extremization of multiple integrals. J. Optim. Theory Appl. **127**, 523–533 (2005)
6. Carlson, D.A., Leitmann, G.: The direct method for a class of infinite Horizon Dynamic games. Adv. Comput. Manage. Sci. **7**(19), 319–334. Springer, New York, NY (2005)
7. Carlson, D.A., Leitmann, G.: A direct method for open-loop Dynamic games for affine control systems. GERAD 25th Anniv. Ser. **10**(2), 37–55. Springer, New York, NY (2005)
8. Ewing, G.M.: Calculus of Variations with Applications. Dover Publications, Inc., 31 East 2nd Street, Mineola, NY, P 11501 (1985). Reprint of work published by W. W. Norton & Co., New York (1969)
9. Leitmann, G.: A note on absolute extrema of certain integrals. Int. J. NonLin. Mech. **2**, 55–59 (1967)
10. Leitmann, G.: On a class of direct optimization problems. J. Optim. Theory Appl. **108**, 467–481 (2001)
11. Leitmann, G.: A direct method of optimization and its applications to a class of differential games. Dyn. Contin. Discrete Impulsive Syst. Ser. A **11**, 191–204 (2004)
12. Todhunter, I.: A history of the calculus of variations during the nineteenth century. Chelsea Publishing Company, New York, NY Reprint of original 1861 edition