COLLOQUIUM MATHEMATICUM

VOL. 101 2004 NO. 1

POINTWISE MINIMIZATION OF SUPPLEMENTED VARIATIONAL PROBLEMS

BY

PETER KOSMOL (Kiel) and DIETER MÜLLER-WICHARDS (Hamburg)

Abstract. We describe an approach to variational problems, where the solutions appear as pointwise (finite-dimensional) minima for fixed t of the supplemented Lagrangian. The minimization is performed simultaneously with respect to the state variable x and $\dot{x}$, as opposed to Pontryagin's maximum principle, where optimization is done only with respect to the $\dot{x}$-variable. We use the idea of the equivalent problems of Carathéodory employing suitable (and simple) supplements to the original minimization problem. Whereas Carathéodory considers equivalent problems by use of solutions of the Hamilton–Jacobi partial differential equations, we shall demonstrate that quadratic supplements can be constructed such that the supplemented Lagrangian is convex in the vicinity of the solution. In this way, the fundamental theorems of the calculus of variations are obtained. In particular, we avoid any employment of field theory.

1. Introduction. If a given function has to be minimized on a subset (restriction set) of a given set, then one can try to modify this function outside the restriction set by adding a supplement in such a way that the global minimal solution of the supplemented function lies in the restriction set. It turns out that this global minimal solution is a solution of the original (restricted) problem. The main task is to determine such a suitable supplement.

Supplement Method 1.1. *Let M be an arbitrary set, $f : M \to \mathbb{R}$ a function and T a subset of M. Let $\Lambda : M \to \mathbb{R}$ be a function that is constant on T. If $x_0 \in T$ is a minimal solution of the function $f + \Lambda$ on all of M, then x_0 is a minimal solution of f on T.*

Proof. For $x \in T$ arbitrary we have

$$f(x_0) + \Lambda(x_0) \le f(x) + \Lambda(x) = f(x) + \Lambda(x_0). \ \blacksquare$$

Definition 1.2 (Piecewise continuously differentiable functions). We consider functions $x \in C[a, b]^n$ for which there exists a partition $\{a = t_0 < t_1 < \cdots < t_m = b\}$ such that: x is continuously differentiable on $[t_{i-1}, t_i)$

2000 *Mathematics Subject Classification*: 49K15, 49K99.

Key words and phrases: variational problems, convexification, pointwise minimization, necessary and sufficient conditions, weak and strong local minima, fundamental theorems, extremals and extremaloids, equivalent problems, linear and quadratic supplements.

for each $i \in \{1, \ldots, m\}$ and the derivative $\dot{x}$ has a left-hand limit at t_i. The value of the derivative of x at b is then defined as the left-hand derivative at b. Such a function is called *piecewise continuously differentiable*. The set of these functions is denoted by $RCS^1[a,b]^n$.

DEFINITION 1.3. Let W be a subset of $\mathbb{R}^\nu \times [a,b]$. We call $G : W \to \mathbb{R}^k$ *piecewise continuous* if there is a partition $Z = \{a = t_0 < t_1 < \cdots < t_m = b\}$ of $[a,b]$ such that for each $i \in \{1, \ldots, m\}$ the function $G : W \cap (\mathbb{R}^\nu \times [t_{i-1}, t_i)) \to \mathbb{R}^k$ has a continuous extension $M_i : W \cap (\mathbb{R}^\nu \times [t_{i-1}, t_i]) \to \mathbb{R}^k$.

In what follows we shall consider variational problems in the following setting.

Let $U \subset \mathbb{R}^{2n+1}$ be such that $U_t := \{(p,q) \in \mathbb{R}^{2n} \mid (p,q,t) \in U\} \neq \emptyset$ for all $t \in [a,b]$, and let $L : U \to \mathbb{R}$ be piecewise continuous. The *restriction set* S for given $\alpha, \beta \in \mathbb{R}^n$ is a set of functions that is described by

$$S \subset \{x \in RCS^1[a,b]^n \mid (x(t), \dot{x}(t), t), (x(t), \dot{x}(t-), t) \in U \ \forall t \in [a,b], \\ x(a) = \alpha,\ x(b) = \beta\}.$$

The *variational functional* $f : S \to \mathbb{R}$ to be minimized is defined by

$$f(x) = \int_a^b L(x(t), \dot{x}(t), t)\, dt.$$

The *variational problem* with fixed endpoints is then given by:

$$\text{Minimize } f \text{ on } S.$$

The central idea of the subsequent discussion is to introduce a supplement in integral form that is constant on the restriction set. This leads to a new variational problem with a modified Lagrangian. The solutions of the original variational problem can now be found as minimal solutions of the modified variational functional. Because of the monotonicity of the integral, the variational problem is now solved by pointwise minimization of the Lagrangian with respect to the x- and $\dot{x}$-variables for every fixed t, employing the methods of finite-dimensional optimization.

This leads to sufficient conditions for a solution of the variational problem. This general approach does not even require differentiability of the integrand. Solutions of the pointwise minimization can even lie at the boundary of the restriction set so that the Euler–Lagrange equations do not have to be satisfied. For interior points the Euler–Lagrange equations will naturally appear by setting the partial derivatives to zero, using a linear supplement potential.

1.1. *Equivalent variational problems.* We now attempt to describe an approach to variational problems that uses the idea of equivalent problems of Carathéodory (see [4], also compare Krotov [19]) employing suitable supplements to the original minimization problem. Carathéodory constructs equivalent problems by use of solutions of the Hamilton–Jacobi partial differential equations. In the context of Bellman's dynamic programming (see [1]) this supplement can be interpreted as the so-called value function. The technique to modify the integrand of the variational problem already appears in the works of Legendre in the context of the second variation (accessory problem).

In this paper we shall demonstrate that explicitly given quadratic supplements are sufficient to yield the main results.

Definition 1.4. Let $F : [a,b] \times \mathbb{R}^n \to \mathbb{R}$ with $(t,x) \mapsto F(t,x)$ be continuous and have the partial derivative F_x continuous, and F_t piecewise continuous. Moreover, we require that the partial derivative F_{xx} exists and is continuous, and that F_{tx}, F_{xt} exist in the piecewise sense, and are piecewise continuous and equal. Then we call F a *supplement potential.*

Lemma 1.5. *Let $F : [a,b] \times \mathbb{R}^n \to \mathbb{R}$ be a supplement potential. Then the integral*

$$\int_a^b [\langle F_x(t,x(t)), \dot{x}(t)\rangle + F_t(t,x(t))]\,dt$$

is constant on S.

Proof. Let Z_1 be a common partition of $[a,b]$ for F and x, i.e. $Z_1 = \{a = \widetilde{t}_0 < \widetilde{t}_1 < \cdots < \widetilde{t}_j = b\}$, such that the requirements of piecewise continuity for $\dot{x}$ and F_t are satisfied with respect to Z_1. Then

$$\begin{aligned}
&\int_a^b [\langle F_x(t,x(t)), \dot{x}(t)\rangle + F_t(t,x(t))]\,dt \\
&\qquad = \sum_{i=1}^{j} \int_{\widetilde{t}_{i-1}}^{\widetilde{t}_i} [\langle F_x(t,x(t)), \dot{x}(t)\rangle + F_t(t,x(t))]\,dt \\
&\qquad = \sum_{i=1}^{j} (F(\widetilde{t}_i, x(\widetilde{t}_i)) - F(\widetilde{t}_{i-1}, x(\widetilde{t}_{i-1}))) \\
&\qquad = F(b,\beta) - F(a,\alpha)
\end{aligned}$$

as F is continuous. ∎

An *equivalent problem* is then given through the supplemented Lagrangian $\widetilde{L}$:

$$\widetilde{L} := L - \langle F_x, \dot{x}\rangle - F_t.$$

1.2. *Principle of pointwise minimization.* The aim is to develop sufficient criteria for minimal solutions of the variational problem by replacing the minimization of the variational functional on subsets of a function space by finite-dimensional minimization. This can be accomplished by pointwise minimization of an explicitly given supplemented integrand for fixed t using the monotonicity of the integral (this method was applied to general control problems in [16]). The minimization is done simultaneously with respect to the x- and $\dot{x}$-variables in $\mathbb{R}^{2n}$. This is the main difference as compared to Hamilton–Pontryagin methods, where minimization is done solely with respect to the $\dot{x}$-variables, leading to necessary conditions in the first place.

The principle of pointwise minimization is demonstrated in [15], where a complete treatment of the brachistochrone problem is presented using only elementary minimization in $\mathbb{R}$ (compare also [13, p. 120] or [17]). For a treatment of this problem using fields of extremals see [8, p. 367].

Our approach is based on the following obvious

LEMMA 1.6. *Let A be a set of integrable real functions on $[a,b]$ and let $l^* \in A$. If for all $l \in A$ and all $t \in [a,b]$ we have $l^*(t) \le l(t)$ then*

$$\int_a^b l^*(t)\,dt \le \int_a^b l(t)\,dt \qquad \text{for all } l \in A. \ \blacksquare$$

THEOREM 1.7 (Principle of Pointwise Minimization). *Let a variational problem with Lagrangian L and restriction set S be given. Suppose that for an equivalent variational problem*

$$\text{Minimize } g(x) := \int_a^b \widetilde{L}(x(t), \dot{x}(t), t)\,dt,$$

where

$$\widetilde{L} = L - \langle F_x, \dot{x}\rangle - F_t,$$

there exists an $x^* \in S$ such that for all $t \in [a,b]$ the point $(p_t, q_t) := (x^*(t), \dot{x}^*(t))$ is a minimal solution of the function $(p,q) \mapsto \widetilde{L}(p,q,t) =: \varphi_t(p,q)$ on $U_t := \{(p,q) \in \mathbb{R}^{2n} \mid (p,q,t) \in U\}$. Then x^* is a solution of the original variational problem.

Proof. For the application of Lemma 1.6, set $A = \{l_x : [a,b] \to \mathbb{R} \mid t \mapsto l_x(t) = \widetilde{L}(x(t), \dot{x}(t), t),\ x \in S\}$ and $l^* := l_{x^*}$. According to Lemma 1.5 the integral over the supplement is constant. $\blacksquare$

It turns out (see below) that taking a linear supplement (with respect to x) already leads to the Euler–Lagrange equation by setting the partial derivatives of $\widetilde{L}$ (with respect to p and q) to zero.

1.3. *Linear supplement.* In the classical theory a linear supplement potential F has the structure

$$(t,x) \mapsto F(t,x) = \langle \lambda(t), x \rangle \tag{1}$$

and $\lambda \in RCS^1[a,b]^n$ is a function that has to be determined in a suitable way.

As $F_x(t,x) = \lambda(t)$ and $F_t(t,x) = \langle \dot{\lambda}(t), x \rangle$, the equivalent problem is:

$$\text{Minimize } g(x) = \int_a^b [L(x(t), \dot{x}(t), t) - \langle \lambda(t), \dot{x}(t) \rangle - \langle \dot{\lambda}(t), x(t) \rangle] \, dt \text{ on } S. \tag{2}$$

We shall now attempt to solve this problem through pointwise minimization of the integrand. This simple approach already leads to a very efficient method for the treatment of variational problems.

We have to minimize

$$(p,q) \mapsto \ell_t(p,q) := L(p,q,t) - \langle \lambda(t), q \rangle - \langle \dot{\lambda}(t), p \rangle$$

on U_t. Let W be an open superset of U in the relative topology of $\mathbb{R}^n \times \mathbb{R}^n \times [a,b]$, and let $L : W \to \mathbb{R}$ be piecewise continuous and have L_p and L_q piecewise continuous. If for fixed $t \in [a,b]$ a point $(p_t, q_t) \in \operatorname{Int} U_t$ is a corresponding minimal solution, then the partial derivatives of ℓ_t have to be zero at this point. This leads to the equations

$$L_p(p_t, q_t, t) = \dot{\lambda}(t), \tag{3}$$

$$L_q(p_t, q_t, t) = \lambda(t). \tag{4}$$

The pointwise minimum (p_t, q_t) yields a function $t \mapsto (p_t, q_t)$. It is our aim to show that this pair provides a solution x^* of the variational problem, where $x^*(t) := p_t$ and $\dot{x}^*(t) = q_t$. In the spirit of the supplement method this means that the global minimum is an element of the restriction set S. The freedom of choosing a suitable function λ is exploited to achieve this goal.

Definition 1.8. A function $x^* \in RCS^1[a,b]^n$ is called an *extremaloid* (see Hestenes [10, p. 60]) if it satisfies the Euler–Lagrange equation in integral form, i.e. there is a constant c such that

$$\int_a^t L_x(x(\tau), \dot{x}(\tau), \tau) \, d\tau + c = L_{\dot{x}}(x(t), \dot{x}(t), t) \quad \forall t \in (a,b]. \tag{5}$$

If the extremaloid x^* is a $C^1[a,b]$-function then x^* is called an *extremal*. An extremaloid x^* is called *admissible* if $x^* \in S$.

Remark 1.9. An extremaloid x^* always satisfies the *Weierstrass–Erdmann condition*, i.e.

$$t \mapsto L_{\dot{x}}(x^*(t), \dot{x}^*(t), t) \text{ is continuous.}$$

For an extremaloid x^* the definition

$$\lambda(t) := \int_a^t L_x(x^*(\tau), \dot{x}^*(\tau), \tau)\, d\tau + c \tag{6}$$

(c a constant) leads to a $\lambda \in RCS^1$.

A fundamental question of variational calculus is under what conditions an extremaloid is a minimal solution of the variational problem.

If x^* is an admissible extremaloid then for every $t \in [a, b]$ the pair $(x^*(t), \dot{x}^*(t)) \in \operatorname{Int} U_t$ satisfies the first necessary condition for a pointwise minimum at t. From the perspective of pointwise minimization we can state the following: if setting the partial derivatives of the integrand to zero leads to a pointwise (global) minimum on U_t for every $t \in [a, b]$, then indeed x^* is a solution of the variational problem. The principle of pointwise minimization also provides a criterion to decide which of the extremaloids is the global solution.

For convex integrands, an extremaloid already leads to sufficient conditions for a pointwise minimum.

If for every $t \in [a, b]$ the set U_t is convex and the Lagrangian $L(\cdot, \cdot, t) : U_t \to \mathbb{R}$ is a convex function, then an admissible extremaloid is a solution of the variational problem. This remains true if the extremaloid lies partially on the boundary of U (i.e. if the pointwise minimum lies on the boundary of U_t). As we require continuous differentiability in an open superset of U_t and the vector of partial derivatives (i.e. the gradient) represents the (total) derivative, all directional derivatives are equal to zero at $(x^*(t), \dot{x}^*(t))$. The characterization theorem (see [13, p. 66]) of convex optimization guarantees that this point is indeed a pointwise minimum:

THEOREM 1.10. *Let $X \subset \mathbb{R}^n \times \mathbb{R}^n$ be open and $\phi : X \to \mathbb{R}$ be differentiable. Let K be a convex subset of X and suppose $\phi : K \to \mathbb{R}$ is convex. If for $x^* \in K$ we have $\phi'(x^*) = 0$ then x^* is a minimal solution of ϕ on K.* ∎

We summarize this situation in the following:

THEOREM 1.11. *For convex problems every admissible extremaloid is a solution of the variational problem.* ∎

In view of this theorem it turns out that the method of pointwise minimization can be extended to a much larger class of variational problems, where the integrand can be convexified by use of a suitable supplement (see Section 3).

The invariance property stating that equivalent problems have the same extremaloids, established in the subsequent theorem, leads to the following principle: an extremaloid of a problem that is convexifiable is a solution of the original variational problem. In particular, explicit convexification does not have to be carried out, instead the solvability of certain ordinary differential equations has to be verified (see below).

THEOREM 1.12. *Every extremaloid for the Lagrangian L is an extremaloid for the supplemented Lagrangian*

$$\widetilde{L} := L - \langle F_x, \dot{x} \rangle - F_t$$

and vice versa, where F is a supplement potential.

Proof. We have

$$\widetilde{L}_{\dot{x}} = L_{\dot{x}} - F_x, \qquad \widetilde{L}_x = L_x - \dot{x}^T F_{xx} - F_{tx}.$$

Moreover

$$\frac{d}{dt} F_x(t, x(t)) = F_{xt}(t, x(t)) + \dot{x}(t)^T F_{xx}(t, x(t)).$$

If x satisfies the Euler–Lagrange equation in integral form with respect to L, i.e.

$$L_{\dot{x}} = \int_a^t L_x \, d\tau + c,$$

then there is a constant $\widetilde{c}$ such that

$$\widetilde{L}_{\dot{x}} = \int_a^t \widetilde{L}_x \, d\tau + \widetilde{c},$$

which we shall now show, using the continuity of F_x:

$$\begin{aligned} \int_a^t \widetilde{L}_x \, d\tau &= \int_a^t (L_x - \dot{x}^T F_{xx} - F_{tx}) \, d\tau = L_{\dot{x}} - c - \int_a^t (\dot{x}^T F_{xx} + F_{xt}) \, d\tau \\ &= L_{\dot{x}} - F_x + F_x(x(a), a) - c = \widetilde{L}_{\dot{x}} - \widetilde{c}. \ \blacksquare \end{aligned}$$

As equivalent variational problems have the same extremaloids, Theorem 1.11 also holds for convexifiable problems. The following theorem can be viewed as a central guideline for our further considerations:

THEOREM 1.13. *If, for a given variational problem, there exists an equivalent convex problem, then every admissible extremaloid is a minimal solution.*

2. Smoothness of solutions

THEOREM 2.1. *Let $L : U \to \mathbb{R}$ be such that L_q continuous. Let $x \in S$ be such that*

(i) $V_t := \{q \in \mathbb{R}^n \mid (x(t), q, t) \in U\}$ *is convex for all* $t \in [a, b]$.
(ii) $L(x(t), \cdot, t)$ *is strictly convex on* V_t *for all* $t \in [a, b]$.
(iii) $\lambda : [a, b] \to \mathbb{R}^n$, *where* $t \mapsto \lambda(t) := L_q(x(t), \dot{x}(t), t)$, *is continuous.*

Then x is smooth.

Proof. Let $q \mapsto \phi_t(q) := L(x(t), q, t) - \langle \lambda(t), q \rangle$. Then ϕ_t is strictly convex on V_t. We have

$$\phi_t'(\dot{x}(t)) = L_q(x(t), \dot{x}(t), t) - \lambda(t) = 0$$

for all $t \in [a, b]$, hence $\dot{x}(t)$ is the unique minimal solution of ϕ_t on V_t. Let (t_k) be a sequence in $[a, b]$ with $t_k \uparrow t$. Then there exists an interval $I_t := [\bar{t}, t)$ and $K \in \mathbb{N}$ such that $t_k \in I_t$ for $k > K$, and $L_q(x, \dot{x}, \cdot)$ and x are continuous on I_t. We obtain

$$0 = L_q(x(t_k), \dot{x}(t_k), t_k) - \lambda(t_k) \to L_q(x(t), \dot{x}(t-), t) - \lambda(t).$$

Hence $\dot{x}(t-)$ is the minimal solution of ϕ_t on V_t. As ϕ_t is strictly convex, we finally obtain $\dot{x}(t) = \dot{x}(t-)$, which proves the theorem. ∎

COROLLARY 2.2. *Let x^* be an extremaloid, and suppose that*:

(i) $V_t := \{q \in \mathbb{R}^n \mid (x^*(t), q, t) \in U\}$ *is convex for all* $t \in [a, b]$.
(ii) $L(x^*(t), \cdot, t)$ *is strictly convex on* V_t.

Then x^* is an extremal.

Proof. $\lambda : [a, b] \to \mathbb{R}^n$ with $t \mapsto \lambda(t) := L_q(x^*(t), \dot{x}^*(t), t)$ is continuous (the Weierstrass–Erdmann condition is satisfied). ∎

The following example shows that the above theorem does not hold if the strict convexity of $L(x^*(t), \cdot, t)$ is violated:

EXAMPLE 2.3. Let $L(p, q, t) := \left(\left(t - \frac{1}{2}\right)_+\right)^2 q + \frac{1}{2} p^2$ for $t \in [0, 1]$. Observe that L is convex but not strictly convex with respect to q. We want to consider the corresponding variational problem for the boundary conditions $x(0) = 0$, $x(1) = 1$. For the Euler–Lagrange equation we obtain

$$\frac{d}{dt}\left(\left(t - \frac{1}{2}\right)_+\right)^2 = x \;\Rightarrow\; x^*(t) = 2 \cdot \left(t - \frac{1}{2}\right)_+,$$

i.e. x^* is not smooth.

On the other hand, $\lambda(t) = L_q(t) = \left(\left(t - \frac{1}{2}\right)_+\right)^2$ is continuous and hence the Weierstrass–Erdmann condition is satisfied. Thus x^* is an extremaloid of the convex variational problem and hence a minimal solution (Theorem 1.12). The function ϕ_t is not strictly convex:

$$\phi_t(q) = \left(\left(t - \frac{1}{2}\right)_+\right)^2 q + \frac{1}{2} p^2 - \left(\left(t - \frac{1}{2}\right)_+\right)^2 = \frac{1}{2} p^2,$$

i.e. ϕ_t is constant with respect to q, which means that every q is a minimal solution of ϕ_t. In particular the set of minimal solutions of ϕ_t is unbounded (and not just non-unique). ∎

DEFINITION 2.4. Let $x^* \in RCS^1[a,b]^n$ be an extremaloid, and let $L^0_{\dot{x}\dot{x}}(t) := L_{\dot{x}\dot{x}}(x^*(t), \dot{x}^*(t), t)$ satisfy the *strong Legendre–Clebsch condition*, i.e. $L^0_{\dot{x}\dot{x}}$ is positive definite on $[a,b]$. Then x^* is called a *regular extremaloid*.

REMARK 2.5. The Legendre–Clebsch condition, i.e. $L^0_{\dot{x}\dot{x}}$ positive semi-definite on $[a,b]$, is a classical necessary condition for a minimal solution of the variational problem (see [10]).

The following example shows that a regular extremaloid is not always an extremal:

EXAMPLE 2.6. Consider the variational problem on the interval $[-2\pi, 2\pi]$ with the boundary conditions $x^*(2\pi) = x^*(-2\pi) = 0$ given by the Lagrangian $(p,q,t) \mapsto L(p,q,t) := \cos q + (t/\gamma)q$ for $\gamma > 2\pi$ for $(p,q,t) \in U = \mathbb{R} \times (-3\pi/2, 3\pi/2) \times [-2\pi, 2\pi]$. Then $L_q = -\sin q + t/\gamma$ and $L_{qq} = -\cos q$, the latter being positive for $\pi/2 < |q| < 3\pi/2$. The Euler–Lagrange equation is
$$\frac{d}{dt}(-\sin \dot{x} + t/\gamma) = 0,$$
i.e.
$$-\sin \dot{x} + t/\gamma = c.$$
In particular, any solution of the above equation satisfies the Weierstrass–Erdmann condition.

Choosing x^* to be even (i.e. $\dot{x}^*$ odd), we obtain, according to (5),
$$c = \frac{1}{b-a}\int_a^b L_{\dot{x}}(x^*(t), \dot{x}^*(t), t)\, dt = \frac{1}{4\pi}\int_{-2\pi}^{2\pi}\left(-\sin(\dot{x}^*(t)) + \frac{t}{\gamma}\right) dt = 0$$
and hence we have to solve
$$\sin q = t/\gamma$$
with $\pi/2 < |q| < 3\pi/2$ in order to satisfy the strong Legendre–Clebsch condition. We obtain two different solutions:
$$q(t) = \begin{cases} -\arcsin(t/\gamma) + \pi & \text{for } 2\pi \geq t > 0 \text{ and } \pi/2 < q < 3\pi/2, \\ -\arcsin(t/\gamma) - \pi & \text{for } -2\pi \leq t < 0 \text{ and } -\pi/2 > q > -3\pi/2, \end{cases}$$
and
$$q(t) = \begin{cases} -\arcsin(t/\gamma) - \pi & \text{for } 2\pi \geq t > 0 \text{ and } -\pi/2 > q > -3\pi/2, \\ -\arcsin(t/\gamma) + \pi & \text{for } -2\pi \leq t < 0 \text{ and } \pi/2 < q < 3\pi/2. \end{cases}$$
Obviously, both solutions are discontinuous at $t = 0$. The extremals are then obtained via integration:
$$x^*(t) = \begin{cases} t(-\arcsin(t/\gamma) + \pi) - \gamma\sqrt{1-(t/\gamma)^2} - D_1 & \text{for } t > 0, \\ t(-\arcsin(t/\gamma) - \pi) - \gamma\sqrt{1-(t/\gamma)^2} - D_1 & \text{for } t < 0, \end{cases}$$

where $D_1 = 2\pi(-\arcsin(2\pi/\gamma) + \pi) - \gamma\sqrt{1-(2\pi/\gamma)^2}$ according to the boundary conditions; and

$$x^*(t) = \begin{cases} t(-\arcsin(t/\gamma) - \pi) - \gamma\sqrt{1-(t/\gamma)^2} - D_2 & \text{for } t > 0, \\ t(-\arcsin(t/\gamma) + \pi) - \gamma\sqrt{1-(t/\gamma)^2} - D_2 & \text{for } t < 0, \end{cases}$$

where $D_2 = 2\pi(-\arcsin(2\pi/\gamma) - \pi) - \gamma\sqrt{1-(2\pi/\gamma)^2}$. Both extremaloids are admissible and both satisfy the strong Legendre condition.

Pointwise minimization. The method of pointwise minimization (with respect to p, q for fixed t) leads directly to a global solution of the variational problem. For the linear supplement choose the Dubois-Reymond form $\lambda(t) = \int_{-2\pi}^{t} L_x\, d\tau + c = c$. Let $\Phi_t : \mathbb{R} \times (-3\pi/2, 3\pi/2) \to \mathbb{R}$ with

$$\Phi_t(p,q) := L(p,q,t) + \lambda(t)q + \dot{\lambda}(t)p = \cos q + \frac{t}{\gamma}\, q + cq =: \phi_t(q),$$

where $\phi_t : (-3\pi/2, 3\pi/2) \to \mathbb{R}$. Choosing x^* to be even (i.e. $\dot{x}^*$ odd), we obtain as above (see (5))

$$c = \frac{1}{b-a}\int_a^b L_{\dot{x}}(x^*(t), \dot{x}^*(t), t)\, dt = \frac{1}{4\pi}\int_{-2\pi}^{2\pi} (-\sin(\dot{x}^*(t)) + t/\gamma)\, dt = 0.$$

For $t > 0$ and $q > 0$ we obtain $\phi_t(q) > -1$ and $\phi_t(-\pi) = \cos(-\pi) - (t/\gamma)\pi$, hence we must look for minimal solutions on $(-3\pi/2, -\pi/2)$. On this interval ϕ_t is convex.

The necessary condition yields

$$\sin q = t/\gamma,$$

for which we obtain the following solution: let $q = -\pi + r$, where $r \in [-\pi/2, \pi/2]$; then $\sin q = \sin(-\pi+r) = -\sin r = t/\gamma$, hence $r = -\arcsin(t/\gamma)$.

For $t < 0$ for analogous reasons minimal solutions are found on $(\pi/2, 3\pi/2)$. Hence,

$$q(t) = \begin{cases} -\arcsin(t/\gamma) - \pi & \text{for } 2\pi \ge t > 0 \text{ and } -\pi/2 > q > -3\pi/2, \\ -\arcsin(t/\gamma) + \pi & \text{for } -2\pi \le t < 0 \text{ and } \pi/2 < q < 3\pi/2. \end{cases}$$

As ϕ_t does not depend on p, every pair $(p, q(t))$ is a pointwise minimal solution of Φ_t on $\mathbb{R} \times (-3\pi/2, 3\pi/2)$. We choose

$$x^*(t) = p(t) = \int_{-2\pi}^{t} q(\tau)\, d\tau.$$

Then x^* is in $RCS^1[-2\pi, 2\pi]$, even, and satisfies the boundary conditions. According to the principle of pointwise minimization, x^* is the global minimal solution of the variational problem. ∎

We point out that Carathéodory in [4] always assumes that x^* is smooth.

3. Weak local minima. In what follows, let U be open (and as in Section 1). For our subsequent discussion that is based on convexification of the Lagrangian we need the following

LEMMA 3.1. *Let V be an open and A a compact subset of a metric space (X,d), and suppose $A \subset V$. Then there is a positive δ such that*

$$\bigcup_{x\in A} K(x,\delta) \subset V.$$

Proof. Suppose for every $n \in \mathbb{N}$ there exist $x_n \in X\backslash V$ and $a_n \in A$ such that $d(x_n, a_n) < 1/n$. As A is compact there is a convergent subsequence (a_k) such that $a_k \to \overline{a} \in A \subset V$. We obtain

$$d(x_k, \overline{a}) \le d(x_k, a_k) + d(a_k, \overline{a}) \to 0,$$

a contradiction to $X\backslash V$ being closed. ∎

LEMMA 3.2. *Let $M \in L(\mathbb{R}^{2n})$ be a matrix of the structure*

$$M = \begin{pmatrix} A & C^T \\ C & D \end{pmatrix},$$

where $A, C, D \in L(\mathbb{R}^n)$ and D is positive definite and symmetric. Then M is positive (*semi-*)*definite if and only if $A - C^T D^{-1} C$ is positive* (*semi-*) *definite.*

Proof. Let $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be defined by

$$f(p,q) := \begin{pmatrix} p \\ q \end{pmatrix}^T \begin{pmatrix} A & C^T \\ C & D \end{pmatrix} \begin{pmatrix} p \\ q \end{pmatrix} = p^T A p + 2q^T C p + q^T D q.$$

Minimization with respect to q for fixed p yields $2Dq = -2Cp$ and hence

$$q(p) = -D^{-1} C p.$$

By inserting this result into f we obtain

$$f(p, q(p)) = p^T A p - 2p^T C^T D^{-1} C p + p^T C^T D^{-1} C p = p^T (A - C^T D^{-1} C) p.$$

If $A - C^T D^{-1} C$ is positive (semi-)definite it follows that $f(p, q(p)) > 0$ for $p \neq 0$ ($f(p, q(p)) \ge 0$ in the semi-definite case). For $p = 0$ and $q \neq 0$ obviously $f(p,q) > 0$.

Conversely, suppose M is positive (semi-)definite. Then $(0,0)$ is the only minimal solution of f. Hence the function $p \mapsto f(p, q(p))$ has 0 as the minimal solution, i.e. $A - C^T D^{-1} C$ is positive (semi-)definite. ∎

DEFINITION 3.3. We say that the *Legendre–Riccati condition* is satisfied if there exists a continuously differentiable symmetric matrix function $W : [a,b] \to L(\mathbb{R}^n)$ such that for all $t \in [a,b]$ the expression

$$L^0_{xx} + \dot{W} - (L^0_{x\dot{x}} + W)(L^0_{\dot{x}\dot{x}})^{-1}(L^0_{\dot{x}x} + W) \tag{7}$$

is positive definite.

We point out that Zeidan (see [21]) treats variational problems by discussing the Hamiltonian. She presents a corresponding condition for W with respect to the Hamiltonian in order to obtain sufficient conditions.

If the Legendre–Riccati condition is satisfied, we shall introduce a *quadratic supplement potential* based on the corresponding matrix W such that the supplemented Lagrangian is strictly convex (compare also [8, Vol. I, p. 251]). Klötzler [5, p. 325] uses a modification of the Hamiltonian, also leading to Riccati's equation, such that the resulting function is concave, in the context of extensions of field theory.

THEOREM 3.4 (Fundamental theorem). *Let $L : U \to \mathbb{R}$ be continuous and $L(\cdot,\cdot,t)$ twice continuously differentiable. If the Legendre–Riccati condition is satisfied, then an admissible regular extremaloid x^* is a weak local minimal solution of the given variational problem.*

Proof. Let a differentiable $W : [a,b] \to L(\mathbb{R}^n)$ satisfy the Legendre–Riccati condition. Then we choose the quadratic supplement potential

$$F : [a,b] \times \mathbb{R}^n \to \mathbb{R}$$

with $F(t,p) = -\frac{1}{2}p^T W(t)p$, which leads to an equivalent variational problem with the modified Lagrange function

$$\begin{aligned}\widetilde{L}(p,q,t) &:= L(p,q,t) - \langle q, F_p(t,p)\rangle - F_t(t,p)\\ &= L(p,q,t) + \langle q, W(t)p\rangle + \frac{1}{2}\,\langle p, \dot{W}(t)p\rangle.\end{aligned}$$

We shall now show that there is a $\delta > 0$ such that for all $t \in [a,b]$ the function $(p,q) \mapsto \phi_t(p,q) := \widetilde{L}(p,q,t)$ is strictly convex on $K_t := K((x^*(t), \dot{x}^*(t)), \delta)$. Indeed,

$$M := \phi_t''(x^*(t), \dot{x}^*(t)) = \begin{pmatrix} L^0_{pp} + \dot{W} & L^0_{pq} + W \\ L^0_{qp} + W & L^0_{qq} \end{pmatrix}(t)$$

is positive definite by the Legendre–Riccati condition and Lemma 3.2 (note that $L^0_{qp} = (L^0_{pq})^T$). Then $\phi_t''(p,q)$ is positive definite on an open neighbourhood of $(x^*(t), \dot{x}^*(t))$. As the set $\{(x^*(t), \dot{x}^*(t)) \cup (x^*(t), \dot{x}^*(t-)) \mid t \in [a,b]\}$ is compact, according to Lemma 3.1 there is a (universal) δ such that on K_t the function ϕ_t'' is positive definite, and therefore ϕ_t is strictly convex on K_t for all $t \in [a,b]$. As the RCS^1-ball with center x^* and radius δ is contained in the set

$$S_\delta := \{x \in RCS^1[a,b]^n \mid (x(t), \dot{x}(t)) \in K_t \ \forall t \in [a,b]\},$$

we deduce that the extremal x^* is a (proper) weak local minimum. Thus have identified a (locally) convex variational problem with the Lagrangian $\widetilde{L}$ that is equivalent to the problem involving L (Theorems 1.11, 1.12). ∎

REMARK 3.5. The Legendre–Riccati condition is guaranteed if the *Legendre–Riccati matrix differential equation*

$$L^0_{xx} + \dot{W} - (L^0_{x\dot{x}} + W)(L^0_{\dot{x}\dot{x}})^{-1}(L^0_{\dot{x}x} + W) = cI$$

for a positive $c \in \mathbb{R}$ has a symmetric solution on $[a, b]$. Set $R := L^0_{\dot{x}\dot{x}}$, $Q := L^0_{\dot{x}x}$, $P := L^0_{xx}$ and $A := -R^{-1}Q$, $B := R^{-1}$, $C := P - Q^T R^{-1} Q - cI$. Then for $V := -W$ the above equation assumes the equivalent form (Legendre–Riccati equation)

$$\dot{V} + VA + A^T V + VBV - C = 0.$$

DEFINITION 3.6. The first order system

$$\dot{Z} = CY - A^T Z, \qquad \dot{Y} = AY + BZ$$

is called the *Jacobi equation in canonical form*. A pair of solutions (Z, Y) is called *self-conjugate* if $Z^T Y = Y^T Z$.

For the proof of the subsequent theorem we need the following

LEMMA 3.7 (Quotient Rule).

$$\frac{d}{dt}(A^{-1}(t)) = -A^{-1}(t)\dot{A}(t)A^{-1}(t),$$
$$\frac{d}{dt}(B(t)A^{-1}(t)) = \dot{B}(t)A^{-1}(t) - B(t)A^{-1}(t)\dot{A}(t)A^{-1}(t),$$
$$\frac{d}{dt}(A^{-1}(t)B(t)) = A^{-1}(t)\dot{B}(t) - A^{-1}(t)\dot{A}(t)A^{-1}(t)B(t).$$

Proof. We have

$$0 = \frac{d}{dt}I = \frac{d}{dt}(A(t)A^{-1}(t)) = \dot{A}(t)A^{-1}(t) + A(t)\,\frac{d}{dt}(A^{-1}(t)). \ \blacksquare$$

THEOREM 3.8. *If the Jacobi equation has a solution (Z, Y) such that $Y^T Z$ is symmetric and Y is invertible then $V := ZY^{-1}$ a symmetric solution of the Legendre–Riccati equation*

$$\dot{V} + VA + A^T V + VBV - C = 0.$$

Proof. If $Y^T Z$ is symmetric, then $V := ZY^{-1}$ is also symmetric, as $(Y^T)^{-1}(Y^T Z)Y^{-1} = ZY^{-1} = V$.

Now V is a solution of the Legendre–Riccati equation, because according to the quotient rule (Lemma 3.7) we have

$$\dot{V} = \dot{Z}Y^{-1} - ZY^{-1}\dot{Y}Y^{-1} = (CY - A^T Z)Y^{-1} - ZY^{-1}(AY + BZ)Y^{-1}$$
$$= C - A^T V - VA - VBV. \ \blacksquare$$

DEFINITION 3.9. Let x^* be an extremaloid and let A, B, C be as in Remark 3.5. Let (x, y) be a solution of the Jacobi equation in vector form

$$\dot{z} = Cy - A^T z, \qquad \dot{y} = Ay + Bz$$

such that $y(a) = 0$ and $z(a) \neq 0$. If there is a point $t_0 \in (a, b]$ such that $y(t_0) = 0$ then we say t_0 is a *conjugate point* of x^* with respect to a.

For the next theorem compare Hartman [9, Theorem 10.2, p. 388].

THEOREM 3.10. *If the regular extremaloid x^* does not have a conjugate point in $(a, b]$ then there exists a self-conjugate solution (Z_1, Y_1) of the matrix Jacobi equation such that Y_1 is invertible on $[a, b]$. Furthermore, the Legendre–Riccati condition is satisfied.*

Proof. For any solution (Z, Y) of the matrix Jacobi equation

$$\dot{Z} = CY - A^T Z, \qquad \dot{Y} = AY + BZ$$

one can prove that $\frac{d}{dt}(Z^T Y - Y^T Z) = 0$ using the product rule and the fact that the matrices B and C are symmetric. Hence $Z^T Y - Y^T Z$ is a constant matrix K on $[a, b]$. If we consider the following initial value problem:

$$\begin{aligned} Y(a) &= I, & Y_0(a) &= 0, \\ Z(a) &= 0, & Z_0(a) &= I, \end{aligned}$$

then obviously for the solution (Z_0, Y_0) the matrix K is zero, i.e. (Z_0, Y_0) is self-conjugate. Clearly,

$$\begin{pmatrix} Y & Y_0 \\ Z & Z_0 \end{pmatrix}$$

is a fundamental system. Hence any solution (y, z) can be represented in the following way: $y = Yc_1 + Y_0c_2$ and $z = Zc_1 + Z_0c_2$. If $y(a) = 0$ then $0 = y(a) = Ic_1 = c_1$ and $z(a) = Ic_2$, thus $y = Y_0c_2$ and $z = Z_0c_2$.

It turns out that $Y_0(t)$ is nonsingular on $(a, b]$. For suppose there is $t_0 \in (a, b]$ such that $Y_0(t_0)$ is singular; then the linear equation $Y_0(t_0)c = 0$ has a nontrivial solution c_0. Let $y_0(t) := Y_0(t)c_0$ on $[a, b]$. Then $y_0(a) = 0$ and $y_0(t_0) = 0$. Moreover for $z_0(t) := Z_0(t)c_0$ we have $z_0(a) = Ic_0 = c_0 \neq 0$, hence t_0 is a conjugate point of a, a contradiction.

We now use a construction that can be found in Hestenes [10]: there is an $\varepsilon > 0$ such that x^* can be extended to a function $\overline{x^*}$ on $[a - \varepsilon, b]$ and $L_{\dot{x}\dot{x}}(\overline{x^*}, \dot{\overline{x^*}}, \cdot)$ remains positive definite on $[a - \varepsilon, b]$. If we insert $\overline{x^*}$ into $L_{x\dot{x}}$ and L_{xx} then the matrices A, B, C (in the notation of Remark 3.5) retain their properties and we can consider the corresponding Jacobi equation extended to $[a - \varepsilon, b]$. Then Lemma 5.1 in [10, p. 129] yields an $a_0 < a$ such that $\overline{x^*}$ has no conjugate point with respect to a_0 on $(a_0, b]$. But then, using the same initial conditions at a_0 and the same argument as in the first part of the proof, we obtain a self-conjugate solution (Z_1, Y_1) on $[a_0, b]$ such that Y_1 is nonsingular on $(a_0, b]$. The restriction of Y_1 is, of course, a solution of the Jacobi equation on $[a, b]$ that is nonsingular there.

From Theorem 3.8 it follows that $Z_1 Y_1^{-1}$ is a symmetric solution of the Legendre–Riccati equation. ∎

THEOREM 3.11 (Fundamental Theorem of Jacobi–Weierstrass). *If the regular extremal x^* does not have a conjugate point on $(a,b]$ then x^* is a weak local minimal solution of the given variational problem.* ∎

3.1. *Carathéodory minimale*

DEFINITION 3.12. A function $x^* \in RCS^1[a,b]^n$ is called a *Carathéodory minimale* if for every $t_0 \in (a,b)$ there are $s,t \in (a,b)$ with $s < t_0 < t$ such that $x^*|_{[s,t]}$ is a weak local solution of the variational problem

$$\min \int_s^t L(x(\tau), \dot{x}(\tau), \tau)\, d\tau$$

on

$$S_{s,t} := \{x \in RCS^1[s,t]^n \mid (x(\tau), \dot{x}(\tau), \tau) \in U \ \forall \tau \in [s,t], \\ x(s) = x^*(s),\ x(t) = x^*(t)\}.$$

As a consequence of Theorem 3.4 we obtain (compare [4, p. 210])

THEOREM 3.13. *Every regular extremal is a Carathéodory minimale.*

Proof. The matrix $W := r(t-t_0)\cdot I$ satisfies the Legendre–Riccati condition for large r on the interval $[t_0 - 1/r, t_0 + 1/r]$: in fact, for $p \in \mathbb{R}^n$ with $\|p\| = 1$ we obtain

$$p^T(P(t) + rI - (Q^T(t) + r(t-t_0)I)R^{-1}(t)(Q(t) + r(t-t_0)I))p \\ \geq r - (\|Q^T\| + 1)\|R^{-1}\|(\|Q(t)\| + 1) - \|P(t)\| > 0$$

for r large enough. ∎

4. Strong convexity and strong local minima

LEMMA 4.1. *Let X be a normed space, $U \subset X$ open, and $g : U \to \mathbb{R}$ twice continuously differentiable. If for an $x^* \in U$ and $c > 0$ we have*

$$\langle g''(x^*)h, h\rangle \geq c\|h\|^2 \quad \text{for all } h \in X,$$

then there exists a $\delta > 0$ such that g is strongly convex on $K := K(x^*, \delta)$. In particular,

$$g\left(\frac{x+y}{2}\right) \leq \frac{1}{2}\, g(x) + \frac{1}{2}\, g(y) - \frac{c}{8}\, \|x-y\|^2 \quad \text{for all } x, y \in K.$$

Proof. As g'' is continuous, there is a $\delta > 0$ such that $\|g''(x) - g''(x^*)\| \leq c/2$ for all $x \in K(x^*, \delta)$. Hence for all $x \in K$ and all $h \in X$ we obtain

$$\langle g''(x)h, h\rangle = \langle (g''(x) - g''(x^*))h, h\rangle + \langle g''(x^*)h, h\rangle \\ \geq -\|g''(x) - g''(x^*)\|\|h\|^2 + c\|h\|^2 \geq \frac{c}{2}\, \|h\|^2.$$

From [13, Theorem 5, p. 42] it follows that g is strongly convex on K. ∎

Theorem 4.2 (Uniform strong convexity of the Lagrangian). *Let x^* be an extremal and suppose the Legendre–Riccati condition is satisfied. Then there is a $\delta > 0$ and a $c > 0$ such that for all $(p,q),(u,v) \in K_t := K((x^*(t),\dot{x}^*(t)),\delta)$ and for all $t \in [a,b]$ we have*

$$\widetilde{L}\left(\frac{p+u}{2},\frac{q+v}{2},t\right) \le \frac{1}{2}\,\widetilde{L}(p,q,t)+\frac{1}{2}\,\widetilde{L}(u,v,t)-\frac{c}{8}\,(\|p-u\|^2+\|q-v\|^2).$$

Proof. Let ϕ_t be as in Theorem 3.4. As the set $K_1 := \{(p,q) \in \mathbb{R}^{2n} \mid \|p\|^2+\|q\|^2 = 1\}$ is compact and as $t \mapsto \phi_t''(x^*(t),\dot{x}^*(t))$ is continuous on $[a,b]$ there is a positive $c \in \mathbb{R}$ such that for all $t \in [a,b]$,

$$\begin{pmatrix} p \\ q \end{pmatrix}^T \phi_t''(x^*(t),\dot{x}^*(t)) \begin{pmatrix} p \\ q \end{pmatrix} \ge c \tag{8}$$

on K_1, i.e. $t \mapsto \phi_t''(x^*(t),\dot{x}^*(t))$ is uniformly positive definite on $[a,b]$.

According to Lemma 3.1 there is a $\varrho > 0$ such that on the compact set

$$\overline{\bigcup_{t\in[a,b]} K((x^*(t),\dot{x}^*(t),t),\varrho)} \subset U$$

in $\mathbb{R}^{2n+1}$ we have uniform continuity of $(p,q,t) \mapsto \phi_t''(p,q)$. Hence there is a $\delta > 0$ such that for all $(u,v) \in K_t := K((x^*(t),\dot{x}^*(t),t),\delta)$ we have

$$\|\phi_t''(u,v) - \phi_t''(x^*(t),\dot{x}^*(t))\| \le c/2$$

and hence on that set

$$\begin{pmatrix} p \\ q \end{pmatrix}^T \phi_t''(u,v) \begin{pmatrix} p \\ q \end{pmatrix} \ge \frac{c}{2}.$$

Thus $(p,q) \mapsto \widetilde{L}(p,q,t)$ is uniformly strongly convex on K_t for all $t \in [a,b]$, i.e.

$$\widetilde{L}\left(\frac{p+u}{2},\frac{q+v}{2},t\right) \le \frac{1}{2}\,\widetilde{L}(p,q,t)+\frac{1}{2}\,\widetilde{L}(u,v,t)-\frac{c}{8}\,(\|p-u\|^2+\|q-v\|^2)$$

for all $(p,q),(u,v) \in K_t$ and all $t \in [a,b]$. ■

For the corresponding variational functional the above theorem leads to

Corollary 4.3 (Strong convexity of the variational functional). *Let $B(x^*,\delta) := \{x \in RCS^1[a,b]^n \mid \|x(t)-x^*(t)\|^2+\|\dot{x}(t)-\dot{x}^*(t)\|^2 < \delta\ \forall t \in [a,b]\}$. Then the variational functional $\widetilde{f}$ belonging to the Lagrangian $\widetilde{L}$ is uniformly strongly convex with respect to the Sobolev norm*:

$$\widetilde{f}\left(\frac{x+y}{2}\right) \le \frac{1}{2}\,\widetilde{f}(x)+\frac{1}{2}\,\widetilde{f}(y)-\frac{c}{8}\,\|x-y\|_W^2.$$

Let $V := \{x \in B(x^*,\delta) \mid x(a)=\alpha,\ x(b)=\beta\}$ be the subset of those functions satisfying the boundary conditions. Then on V the original functional f is *also uniformly strongly convex with respect to the Sobolev norm*:

$$f\left(\frac{x+y}{2}\right) \le \frac{1}{2}\, f(x) + \frac{1}{2}\, f(y) - \frac{c}{8}\, \|x-y\|_W^2.$$

Furthermore, *every minimizing sequence converges to the minimal solution with respect to the Sobolev norm* (*strong solvability*).

Proof. For the variational functional $\widetilde{f}(x) := \int_a^b \widetilde{L}(x(t), \dot{x}(t), t)\, dt$ we obtain

$$\begin{aligned}
\widetilde{f}\left(\frac{x+y}{2}\right) &= \int_a^b \widetilde{L}\left(\frac{x(t)+y(t)}{2}, \frac{\dot{x}(t)+\dot{y}(t)}{2}, t\right) dt \\
&\le \int_a^b \left(\frac{1}{2}\, \widetilde{L}(x(t), \dot{x}(t), t) + \frac{1}{2}\, \widetilde{L}(y(t), \dot{y}(t), t)\right) dt \\
&\quad - \frac{c}{8} \int_a^b (\|x(t)-y(t)\|^2 + \|\dot{x}(t) - \dot{y}(t)\|^2)\, dt \\
&= \frac{1}{2}\, \widetilde{f}(x) + \frac{1}{2}\, \widetilde{f}(y) - \frac{c}{8}\, \|x-y\|_W^2
\end{aligned}$$

for all $x, y \in B(x^*, \delta)$. Thus $\widetilde{f}$ is strictly convex on $B(x^*, \delta)$. On V the functional $f(x) := \int_a^b L(x(t), \dot{x}(t), t)\, dt$ differs from $\widetilde{f}$ only by a constant, hence has the same minimal solutions. On V the inequality

$$f\left(\frac{x+y}{2}\right) \le \frac{1}{2}\, f(x) + \frac{1}{2}\, f(y) - \frac{c}{8}\, \|x-y\|_W^2$$

is satisfied, i.e. f is strongly convex on V with respect to the Sobolev norm. ∎

Using Theorem 3.10 we obtain the following corollary (retaining the notation of the previous corollary):

Corollary 4.4 (Strong convexity). *If the regular extremal x^* does not have a conjugate point then there is a $\delta > 0$ such that on $B(x^*, \delta)$ the modified variational functional $\widetilde{f}$ is uniformly strongly convex with respect to the Sobolev norm*: *and on V the original functional f is also uniformly strongly convex with respect to the Sobolev norm.* ∎

4.1. *Strong local minima.* We now investigate the question under what conditions we can guarantee that weak local minimal solutions are in fact strong local minimal solutions. It turns out that such a strong property can be proved without the use of embedding theory. We will show that strong local minima require a supplement for the Lagrangian that is generated by the supplement potential

$$(t, p) \mapsto F(t, p) := \frac{1}{2}\, \langle p, W(t)p \rangle - \langle \lambda(t), p \rangle,$$

where $W(t)$ is symmetric and the linear term is chosen in such a way that—along the extremal—the necessary conditions for optimality for $\widehat{L}$ coincide with the fulfillment of the Euler–Lagrange equation, and the quadratic term, with its convexification property, provides sufficient optimality conditions.

The supplemented Lagrangian then has the following structure:

$$\begin{aligned}\widehat{L}(p,q,t) &= L(p,q,t) + F_t + \langle F_p, q\rangle \\ &= L(p,q,t) + \frac{1}{2}\langle p, \dot{W}(t)p\rangle - \langle \dot{\lambda}(t), p\rangle + \langle W(t)p, q\rangle - \langle \lambda(t), q\rangle.\end{aligned}$$

If we define

$$\lambda(t) := L_q(x^*(t), \dot{x}^*(t), t) + W(t)x^*(t)$$

then using the Euler–Lagrange equation for the original Lagrangian L, i.e. $\frac{d}{dt}L_q = L_p$, we obtain in fact the necessary conditions:

$$\widehat{L}_p(x^*(t), \dot{x}^*(t), t) = \widehat{L}_q(x^*(t), \dot{x}^*(t), t) = 0.$$

In the subsequent theorem we make use of the following lemma which is a well known immediate consequence of Brouwer's fixed point theorem:

LEMMA 4.5. *Let $r > 0$ and $x_0 \in \mathbb{R}^n$. Let $A : \overline{K}(x_0, r) \to \mathbb{R}^n$ be continuous, and let $\langle Ax, x - x_0\rangle \geq 0$ for all $x \in S(x_0, r)$. Then the nonlinear equation $Ax = 0$ has a solution in $\overline{K}(x_0, r)$.*

Proof. Otherwise Brouwer's fixed point theorem applied to the mapping

$$x \mapsto g(x) := -r\left(\frac{Ax}{\|Ax\|}\right) + x_0$$

would lead to a contradiction. ∎

THEOREM 4.6 (Strong local minimum). *Let x^* be an admissible, regular extremal and suppose the Legendre–Riccati condition is satisfied. Suppose there exists a $\kappa > 0$ such that for all $t \in [a,b]$ and all p with $\|p - x^*(t)\| < \kappa$ the set $V_{t,p} := \{q \in \mathbb{R}^n \mid (p,q,t) \in U\}$ is convex and the function $L(p,\cdot,t)$ is convex on $V_{t,p}$. Then x^* is a locally strong minimal solution of the variational problem, i.e. there is a positive d such that for all $x \in K := \{x \in S \mid \|x - x^*\|_\infty < d\}$ we have*

$$\int_a^b L(x^*(t), \dot{x}^*(t), t)\, dt \leq \int_a^b L(x(t), \dot{x}(t), t)\, dt.$$

Proof. In Theorem 4.2 we have constructed positive constants c and δ such that for all $(p,v),(u,v) \in K_t := K((x^*(t), \dot{x}^*(t)), \delta)$ we have

$$\widetilde{L}\left(\frac{p+u}{2}, \frac{q+v}{2}, t\right) \leq \frac{1}{2}\widetilde{L}(p,q,t) + \frac{1}{2}\widetilde{L}(u,v,t) - \frac{c}{8}(\|p-u\|^2 + \|q-v\|^2).$$

But $\widehat{L}$ differs from $\widetilde{L}$ only by the linear term

$$-\langle\dot{\lambda}(t), p\rangle - \langle\lambda(t), q\rangle.$$

In particular $\widetilde{L}'' = \widehat{L}''$, i.e. the convexity properties remain unchanged, and hence $\widehat{L}(\cdot,\cdot,t)$ is strongly convex on K_t with a uniform constant c for all $t \in [a,b]$.

From [14, p. 39 ff (Satz 4 and 5)] we deduce that for this c,

$$\begin{aligned}\langle\widehat{L}_q(x^*(t),q,t) - \widehat{L}_q(x^*(t),\dot{x}^*(t),t), q - \dot{x}^*(t)\rangle &= \langle\widehat{L}_q(x^*(t),q,t), q - \dot{x}^*(t)\rangle \\ &\geq 4c\|q - \dot{x}^*(t)\|^2\end{aligned}$$

(strong monotonicity of $\widehat{L}_q(x^*(t),\cdot,t)$). The uniform continuity of $\widehat{L}_q$ guarantees that for $0 < \varepsilon < c\cdot\delta/2$ there is a $0 < d \leq \min\{\delta/2, \kappa\}$ such that for all $t \in [a,b]$,

$$\|\widehat{L}_q(x^*(t),q,t) - \widehat{L}_q(p,q,t)\| < \varepsilon$$

for all $p \in x^*(t) + K(0,d)$. We obtain

$$\begin{aligned}&\langle\widehat{L}_q(x^*(t),q,t) - \widehat{L}_q(p,q,t), q - x^*(t)\rangle \\ &\qquad\leq \|\widehat{L}_q(x^*(t),q,t) - \widehat{L}_q(p,q,t)\|\,\|q - x^*(t)\| < \varepsilon\|q - x^*(t)\|.\end{aligned}$$

Let $\varrho = \delta/2$. Then on the sphere $\{q \mid \|q - \dot{x}^*(t)\| = \varrho\}$ for all $p \in x^*(t) + K(0,d)$ we have

$$\langle\widehat{L}_q(p,q,t), q - \dot{x}^*(t)\rangle \geq c\|q - \dot{x}^*(t)\|^2 - \varepsilon\|q - x^*(t)\| = c\varrho^2 - \varepsilon\varrho > 0.$$

Hence from Lemma 4.5, for every $p \in x^*(t) + K(0,d)$ we obtain a $q^*(p) \in K(\dot{x}^*(t),\varrho)$ such that

$$\widehat{L}_q(p, q^*(p), t) = 0.$$

From the convexity of $\widehat{L}(p,\cdot,t)$ we conclude that $q^*(p)$ is a minimal solution of $\widehat{L}(p,\cdot,t)$ on $V_{t,p}$.

We shall now show that for all $t \in [a,b]$, $(x^*(t), \dot{x}^*(t))$ is a minimal solution of $L(\cdot,\cdot,t)$ on

$$W_t := \{(p,q) \in \mathbb{R}^n \times \mathbb{R}^n \mid (p,q,t) \in U \text{ and } \|p - x^*(t)\| < d\}.$$

For, suppose there exists $(p,q) \in W_t$ such that

$$L(x^*(t), \dot{x}^*(t), t) > L(p,q,t).$$

As x^* is an extremal, and as $\widehat{L}(\cdot,\cdot,t)$ is convex on K_t, $(x^*(t), \dot{x}^*(t))$ is a minimal solution of $\widehat{L}(\cdot,\cdot,t)$ on K_t by construction of $\widehat{L}$. As $(p, q^*(p)) \in K_t$ we obtain

$$\widehat{L}(x^*(t), \dot{x}^*(t), t) \leq \widehat{L}(p, q^*(p), t) \leq \widehat{L}(p,q,t) < \widehat{L}(x^*(t), \dot{x}^*(t), t),$$

a contradiction.

For all $x \in K = \{x \in S \mid \|x^* - x\|_\infty < d\}$ we then have

$$\int_a^b \widehat{L}(x^*(t), \dot{x}^*(t), t)\, dt \le \int_a^b \widehat{L}(x(t), \dot{x}(t), t)\, dt,$$

and as the integrals differ on S only by a constant, we obtain the corresponding inequality also for the original Lagrangian:

$$\int_a^b L(x^*(t), \dot{x}^*(t), t)\, dt \le \int_a^b L(x(t), \dot{x}(t), t)\, dt,$$

which completes the proof. ■

The previous theorem together with Theorem 3.10 leads to the following

COROLLARY 4.7 (Strong local minimum). *Let x^* be an admissible and regular extremal without conjugate points. Suppose there exists a $\kappa > 0$ such that for all $t \in [a, b]$ and all p with $\|p - x^*(t)\| < \kappa$ the set $V_{t,p} := \{q \in \mathbb{R}^n \mid (p, q, t) \in U\}$ is convex and the function $L(p, \cdot, t)$ is convex on $V_{t,p}$. Then x^* is a locally strong minimal solution of the variational problem, i.e. there is a positive d such that for all $x \in K := \{x \in S \mid \|x - x^*\|_\infty < d\}$ we have*

$$\int_a^b L(x^*(t), \dot{x}^*(t), t)\, dt \le \int_a^b L(x(t), \dot{x}(t), t)\, dt.$$

REMARK 4.8. If in particular $U = U_1 \times U_2 \times [a, b]$, where $U_1 \subset \mathbb{R}^n$ is open, $U_2 \subset \mathbb{R}^n$ is open and convex, and $L(p, \cdot, t) : U_2 \to \mathbb{R}$ is convex for all $(p, t) \in U_1 \times [a, b]$, then the requirements of the previous theorem are satisfied.

5. Necessary conditions. We briefly restate the Euler–Lagrange equation as a necessary condition in the piecewise continuous case. The standard proof carries over to this situation (see Hestenes [10, Lemma 5.1, p. 70]).

THEOREM 5.1. *Suppose $L, L_x, L_{\dot{x}}$ are piecewise continuous. Let x^* be a solution of the variational problem with graph contained in the interior of U. Then x^* is an extremaloid, i.e. there is a $c \in \mathbb{R}^n$ such that*

$$L_{\dot{x}}(x^*(t), \dot{x}^*(t), t) = \int_a^t L_x(x^*(\tau), \dot{x}^*(\tau), \tau)\, d\tau + c. \ ■$$

5.1. *The Jacobi equation as a necessary condition.* A different approach to obtaining the Jacobi equation is to consider the variational problem that corresponds to the second directional derivative of the original variational problem:

Let x^* be a solution of the original variational problem, let

$$V := \{h \in RCS^1[a,b]^n \mid h(a) = h(b) = 0\}$$

and let

$$\phi(\alpha) := f(x^* + \alpha h) = \int_a^b L(x^*(t) + \alpha h(t), \dot{x}^*(t) + \alpha \dot{h}(t), t)\, dt.$$

Then the necessary condition yields

$$\begin{aligned} 0 \le \phi''(0) = f''(x^*, h) &= \int_a^b (\langle L^0_{\dot{x}\dot{x}} \dot{h}, \dot{h}\rangle + 2\langle L^0_{\dot{x}x} h, \dot{h}\rangle + \langle L^0_{xx} h, h\rangle)\, dt \\ &= \int_a^b (\langle R\dot{h}, \dot{h}\rangle + 2\langle Qh, \dot{h}\rangle + \langle Ph, h\rangle)\, dt, \end{aligned}$$

using our notation of Remark 3.5. Then the (quadratic) variational problem

$$\text{minimize } f''(x^*, \cdot) \text{ on } V$$

is called the *accessory* (*secondary*) *variational problem*. It turns out that the corresponding Euler–Lagrange equation

$$\frac{d}{dt}(R\dot{h} + Qh) = Q^T \dot{h} + Ph$$

or in matrix form

$$\frac{d}{dt}(R\dot{Y} + QY) = Q^T \dot{Y} + PY$$

yields the Jacobi equation in canonical form

$$\dot{Z} = CY - A^T Z, \qquad \dot{Y} = AY + BZ$$

by setting $Z := R\dot{Y} + QY$ and using the notation of Remark 3.5.

LEMMA 5.2. *If $h^* \in V$ is a solution of the Jacobi equation in the piecewise sense then it is a minimal solution of the accessory problem.*

Proof. Let

$$\Omega(h, \dot{h}) := \langle R\dot{h}, \dot{h}\rangle + 2\langle Qh, \dot{h}\rangle + \langle Ph, h\rangle.$$

Since h^* is an extremal of the accessory problem, i.e. (in the piecewise sense)

$$\frac{d}{dt}(2R\dot{h}^* + 2Qh^*) = 2Ph^* + 2Q^T \dot{h}^*,$$

it follows that

$$\Omega(h^*, \dot{h}^*) = \langle R\dot{h}^* + Qh^*, \dot{h}^*\rangle + \left\langle \frac{d}{dt}(R\dot{h}^* + Qh^*), h^* \right\rangle = \frac{d}{dt}(\langle R\dot{h}^* + Qh^*, h^*\rangle).$$

Using the Weierstrass–Erdmann condition for the accessory problem, i.e.

$R\dot{h} + Qh$ continuous, we can apply the main theorem of differential and integral calculus:

$$\int_a^b 2\Omega(h^*, \dot{h}^*)\,dt = \langle R\dot{h}^* + Qh^*, h^*\rangle|_a^b = 0$$

as $h^*(a) = h^*(b) = 0$. Hence h^* is a minimal solution as $f''(x^*, \cdot)$ is nonnegative. ∎

THEOREM 5.3 (Jacobi's necessary condition for optimality). *If a regular extremal x^* is a minimal solution of the variational problem on S, then a has no conjugate point in (a, b).*

Proof. For otherwise, let (h^*, k^*) be a nontrivial solution of the Jacobi equation with $h(a) = 0$, and let $c \in (a, b)$ be such a conjugate point. Then $h^*(c) = 0$ and $k^*(c) \neq 0$. Because $k^*(c) = R\dot{h}^*(c) + Qh^*(c) = R\dot{h}^*(c)$ it follows that $\dot{h}^*(c) \neq 0$. We define $y(t) = h^*(t)$ for $t \in [a, c]$, and $y(t) = 0$ for $t \in [c, b]$. In particular $\dot{y}(c) \neq 0$. Obviously, y is a solution of the Jacobi equation (in the piecewise sense) and hence, according to the previous lemma, a minimal solution of the accessory problem, a contradiction to Corollary 2.2 on smoothness of solutions, as Ω is strictly convex with respect to $\dot{h}$. ∎

6. C^1-variational problems. Let

$$S_1 := \{x \in C^1[a, b]^n \mid (x(t), \dot{x}(t), t) \in U,\ x(a) = \alpha,\ x(b) = \beta\}$$

and

$$V_1 := \{h \in C^1[a, b]^n \mid h(a) = 0,\ h(b) = 0\}$$

As the variational problem is now considered on a smaller set ($S_1 \subset S$), sufficient conditions carry over to this situation.

Again, just as in the RCS^1-theory, the Jacobi condition is a necessary condition.

THEOREM 6.1. *If a regular extremal $x^* \in C^1[a, b]^n$ is a minimal solution of the variational problem on S_1, then a has no conjugate point in (a, b).*

Proof. Let $c \in (a, b)$ be a conjugate point of a. We consider the quadratic functional

$$h \mapsto g(h) := \int_a^b \Omega(h, \dot{h})\,dt$$

for $h \in V_1$, where

$$\Omega(h, \dot{h}) := \langle R\dot{h}, \dot{h}\rangle + 2\langle Qh, \dot{h}\rangle + \langle Ph, h\rangle.$$

We show that there is $\widehat{h} \in V$ with $g(\widehat{h}) < 0$: for suppose $g(h) \geq 0$ for all $h \in V$; then, as in Lemma 5.2, every extremal h^* is a minimal solution of g, and we use the construction for y as in the proof of Theorem 5.3 to obtain a solution of the Jacobi equation in the piecewise sense (which is, again according to Lemma 5.2, a minimal solution of the accessory problem) which is not smooth, in contradiction to Corollary 2.2, as Ω is strictly convex with respect to $\dot{h}$. Hence there is $\widehat{h} \in V$ with $g(\widehat{h}) < 0$. Now we use the process of smoothing of corners as described in Carathéodory [4]. Thus we obtain an $\overline{h} \in C^1[a,b]^n$ with $g(\overline{h}) < 0$. Let $\alpha \mapsto \phi(\alpha) = f(x^* + \alpha\overline{h})$. Then $g(\overline{h}) = \phi''(0) \geq 0$ as x^* is a minimal solution of f on S_1, a contradiction. ■

7. Conclusion. For a comprehensive view of our results we introduce the notion of an optimal path. Our main objective in this context is to characterize necessary and sufficient conditions.

Definition 7.1. A function $x^* \in RCS^1[a,b]^n$ is called an *optimal path* if it is a weak local solution of the variational problem

$$\min \int_a^t L(x(\tau), \dot{x}(\tau), \tau)\, d\tau$$

on

$$S_t := \{x \in RCS^1[a,t]^n \mid (x(\tau), \dot{x}(\tau), \tau) \in U,\ \forall \tau \in [a,t],\\ x(a) = \alpha,\ x(t) = x^*(t)\}$$

for all $t \in [a,b)$.

Theorem 7.2. *Consider the variational problem*

$$\min \int_a^b L(x(\tau), \dot{x}(\tau), \tau)\, d\tau$$

on

$$S := \{x \in RCS^1[a,t]^n \mid (x(\tau), \dot{x}(\tau), \tau) \in U,\ \forall \tau \in [a,b],\ x(a) = \alpha,\ x(t) = \beta\}.$$

Let $x^* \in RCS^1[a,b]^n$ be a regular extremal. Then the following statements are equivalent:

(i) *x^* is an optimal path.*

(ii) *The variational problem has an equivalent convex problem in the following sense*: *for the original Lagrangian there is a locally convexified Lagrangian $\widetilde{L}$ such that for every subinterval $[a,\tau] \subset [a,b)$ there is a $\delta > 0$ with the property that $\widetilde{L}(\cdot,\cdot,t)$ is strictly convex on the ball $K((x^*(t), \dot{x}^*(t)), \delta)$ for all $t \in [a,\tau]$.*

(iii) *The variational problem has an equivalent convex problem* (*in the sense of* (ii)) *employing a quadratic supplement.*

(iv) *For every subinterval $[a,\tau] \subset [a,b)$ there is a $\delta > 0$ with the property that $(x^*(t), \dot{x}^*(t))$ is a pointwise minimum of $\widetilde{L}(\cdot,\cdot,t)$ on the ball $K((x^*(t), \dot{x}^*(t)), \delta)$ for all $t \in [a,\tau]$.*

(v) *The Legendre–Riccati condition is satisfied on* $[a, b)$.

(vi) *The Jacobi matrix equation has a nonsingular and self-conjugate solution on* $[a, b)$.

(vii) a *has no conjugate point in* (a, b).

REFERENCES

[1] R. E. Bellman, *Dynamic Programming*, Princeton Univ. Press, Princeton, NJ, 1957.

[2] G. A. Bliss, *The problem of Lagrange in the calculus of variations*, Amer. J. Math. 52 (1930), 673–744.

[3] O. Bolza, *Vorlesungen über Variationsrechnung*, Teubner, Leipzig, 1909.

[4] C. Carathéodory, *Variationsrechnung und partielle Differentialgleichungen erster Ordnung*, Teubner, Leipzig, 1935.

[5] —, *Variationsrechnung und partielle Differentialgleichungen erster Ordnung: Variationsrechnung*, herausgegeben, kommentiert und mit Erweiterungen zur Steuerungs- u. Dualitätstheorie versehen von R. Klötzler, B.G. Teubner, Stuttgart, 1994.

[6] L. Cesari, *Optimization Theory and Applications*, Springer, New York, 1983.

[7] W. H. Fleming and R. W. Rishel, *Deterministic and Stochastic Optimal Control*, Springer, Berlin, 1975.

[8] M. Giaquinta and S. Hildebrand, *Calculus of Variations I and II*, Springer, Berlin, 1996.

[9] Ph. Hartman, *Ordinary Differential Equations*, Wiley, New York, 1964.

[10] M. R. Hestenes, *Calculus of Variations and Optimal Control Theory*, Wiley, New York, 1966.

[11] A. D. Ioffe and V. M. Tichomirov, *Theorie der Extremalaufgaben*, Deutscher Verlag der Wiss., Berlin, 1979.

[12] R. Klötzler and S. Pickenhain, *Pontrjagin's Maximum Principle for multidimensional control problems*, in: Optimal Control (Freiburg, 1991), Internat. Ser. Numer. Math. 111, Birkhäuser, Basel, 1993, 21–30.

[13] P. Kosmol, *Optimierung und Approximation*, de Gruyter, Berlin, New York, 1991.

[14] —, *Methoden zur numerischen Behandlung nichtlinearer Gleichungen und Optimierungsaufgaben*, B.G. Teubner, Stuttgart, 1993.

[15] —, *An elementary approach to variational problems*, in: Proc. 12th Baikal Internat. Conf. on Optimization Methods and their Applications, Section 2, Optimal Control (Irkutsk, 2001), 202–208.

[16] P. Kosmol and M. Pavon, *Solving optimal control problems by means of general Lagrange functionals*, Automatica 37 (2001), 907–913.

[17] —, *Bemerkungen zur Brachistochrone*, Abh. Math. Univ. Sem. Hamburg 54 (1984), 91–94.

[18] P. Kosmol and D. Müller-Wichards, *On stability of families of nonlinear equations*, to appear.

[19] V. F. Krotov and V. I. Gurman, *Methods and Problems of Optimal Control*, Nauka, Moscow, 1973 (in Russian).

[20] W. T. Reid, *A matrix differential equation of Riccati type*, Amer. J. Math. 68 (1946), 237–246; addendum, ibid. 70 (1948), 460.

[21] V. Zeidan, *Sufficient conditions for the generalized problem of Bolza*, Trans. Amer. Math. Soc. 275 (1983), 561–586.

[22] V. Zeidan, *Extended Jacobi sufficient conditions for optimal control*, SIAM J. Control Optim. 22 (1984), 294–301.

[23] —, *First and second order sufficient conditions for optimal control and calculus of variations*, Appl. Math. Optim. 11 (1984), 209–226.

[24] E. Zeidler, *Nonlinear Functional Analysis and its Applications III*, Springer, Berlin, 1985.

Mathematisches Seminar
Universität Kiel
Ludewig-Meyn-Str. 4
24098 Kiel, Germany
E-mail: kosmol@math.uni-kiel.de

Fachbereich Elektrotechnik/Informatik
Hochschule f. Angewandte Wissenschaften
Berliner Tor 7
20099 Hamburg, Germany
E-mail: muewi@etech.haw-hamburg.de

Received 5 January 2004;
revised 30 April 2004 (4408)