

AN ALTERNATE SUFFICIENCY PROOF FOR THE NORMAL PROBLEM OF BOLZA

BY
MAGNUS R. HESTENES

1. **Introduction.** In recent years various sufficiency proofs have been given for the problem of Bolza, the most general of which are those given by McShane [6]⁽¹⁾, Myers [7], and the author [5]. The purpose of the present paper is to show that the sufficiency theorem for a normal problem⁽²⁾ of Bolza can be obtained from that for the corresponding theorem for the case in which there are no differential side conditions. A similar result holds for isoperimetric problems.

The problem with which we shall be concerned is that of minimizing a function

$$I(C) = g(a) + \int_C f(a, y, \dot{y}) dt$$

in a class of arcs

$$C: \quad a^h, y^i(t), \quad (t^1 \leqq t \leqq t^2, h = 1, \dots, r; i = 0, 1, \dots, n)$$

satisfying the conditions

$$(1.1) \quad \phi^\beta(a, y, \dot{y}) = 0 \quad (\beta = 1, \dots, m < n),$$

$$(1.2) \quad y^i(t^1) = T^{i1}(a), \quad y^i(t^2) = T^{i2}(a).$$

The components a^1, \dots, a^r of C are independent of t . The sufficiency theorem given below consists of showing that if an arc C_0 satisfies the usual sufficiency conditions for a minimum for the problem just described, there exists a function of the form

$$J(C) = g(a) + \int_C \{f(a, y, \dot{y}) + m^\beta(a, y, \dot{y})\phi^\beta(a, y, \dot{y})\} dt$$

such that C_0 satisfies the corresponding sufficiency conditions for minimizing $J(C)$ in the class of arcs satisfying the end conditions (1.2). From the theory for the case when there are no differential side conditions (cf. [1, chap. 6]), it is seen that C_0 affords a proper strong relative minimum to $J(C)$ subject

Presented to the Society, April 26, 1946; received by the editors April 17, 1946.

(¹) Numbers in brackets refer to the list of references at the end of the paper.

(²) As a matter of fact, the problem discussed is not assumed to be normal. However, our hypotheses are such that by suitably enlarging the class of arcs considered, a normal problem can be obtained for which our hypotheses still hold.

to the conditions (1.2). Since $J(C) = I(C)$ when equations (1.1) hold, it follows that C_0 minimizes $I(C)$ subject to the conditions (1.1) and (1.2).

The existence of a function $J(C)$ having the properties described above can be obtained readily from known results [2, 3]. However, these results give at the same time the desired sufficiency theorem for the problem of Bolza. The information is therefore of no value in reducing the sufficiency proof for the problem of Bolza to the simpler case involving no differential side conditions.

2. The sufficiency theorem. The notations, terminology, and hypotheses here used are those given in three preceding papers on the problem of Bolza. It will be convenient to assume that the functions $f(a, y, p)$, $\phi^\beta(a, y, p)$, $g(a)$, $T^{i1}(a)$, $T^{i2}(a)$ are of class C^4 on the open set \mathfrak{R} of points (a, y, p) forming the domain of definition of our problem. The functions f, ϕ^β are positively homogeneous of order one in p . The arcs under consideration are rectifiable arcs C , with absolutely continuous representations

$$C: \quad a, y(t), \quad t^1 \leq t \leq t^2,$$

whose elements $(a, y(t), \dot{y}(t))$ are in \mathfrak{R} for almost all values of t on $t^1 t^2$.

Consider now a particular arc

$$C_0: \quad a_0, y_0(t), \quad t^1 \leq t \leq t^2,$$

of class C' which satisfies equations (1.1) and (1.2) and does not intersect itself. We suppose that there exists a set of absolutely continuous multipliers $\lambda^\beta(t)$ such that if we set

$$F(a, y, p, \lambda) = f(a, y, p) + \lambda^\beta \phi^\beta(a, y, p),$$

then:

(1) C_0 satisfies *Euler-Lagrange equations*

$$(2.1) \quad \frac{d}{dt} F_{p^i} = F_{y^i}.$$

(2) C_0 satisfies the *transversality condition*

$$(2.2) \quad g_h + F_{p^i}(t^2) T_h^{i2} - F_{p^i}(t^1) T_h^{i1} + \int_{C_0} F_h dt = 0.$$

Here and elsewhere the subscript h denotes the value of the derivative of a function with respect to a_h . The quantities $F_{p^i}(t^1)$, $F_{p^i}(t^2)$ denote the values of F_{p^i} at the initial and final end points of C_0 .

(3) C_0 is *nonsingular*, that is, the matrix

$$\begin{pmatrix} F_{p^i p^i} & \phi_{p^i}^\gamma \\ \phi_{p^i}^\beta & 0 \end{pmatrix} \quad \begin{matrix} (i, j = 0, 1, \dots, n) \\ (\beta, \gamma = 1, \dots, m), \end{matrix}$$

has rank $n+m$ along C_0 . This assumption together with (1) implies that the functions $\lambda^\beta(t)$, $\dot{y}_0^i(t)$ have continuous third derivatives.

(4) C_0 satisfies the *strengthened condition of Weierstrass*, that is, there is a neighborhood N of the values (a, y, \dot{p}, λ) on C_0 such that the inequality

$$E(a, y, \dot{p}, \lambda, q) = F(a, y, q, \lambda) - q^i F_{p^i}(a, y, \dot{p}, \lambda) \geq 0$$

holds whenever (a, y, \dot{p}, λ) is in N and (a, y, \dot{p}) , (a, y, q) are in \mathfrak{R} and satisfy the conditions $\phi^\beta = 0$.

(5) C_0 satisfies the *strengthened Jacobi condition*, that is, the second variation

$$I_2(\gamma) = 2q(\alpha) + \int_{t^1}^{t^2} 2\omega(\alpha, \eta, \dot{\eta}) dt$$

is positive for every variation

$$\gamma: \quad \alpha^h, \eta^i(t) \quad (t^1 \leq t \leq t^2; h = 1, \dots, r; i = 0, 1, \dots, n)$$

satisfying the following conditions:

(a) The functions $\eta^i(t)$ are absolutely continuous on $t^1 t^2$ and have derivatives that are integrable square. The components α^h are independent of t .

(b) The equations

$$(2.3) \quad \Phi^\beta(\alpha, \eta, \dot{\eta}) = \phi_h^\beta \alpha^h + \phi_{y^i}^\beta \eta^i(t) + \phi_{p^i}^\beta \dot{\eta}^i(t) = 0$$

hold for almost all values of t , the arguments in the derivatives being $(a_0, y_0(t), \dot{y}_0(t))$.

(c) The variation γ satisfies the accessory end conditions

$$(2.4) \quad \eta^i(t^1) = T_h^{i1} \alpha^h, \quad \eta^i(t^2) = T_h^{i2} \alpha^h.$$

(d) The variation γ is not essentially null, that is, it is not of the form

$$(2.5) \quad \alpha^h = 0, \quad \eta^i(t) = \theta(t) \dot{y}_0^i(t),$$

where $\theta(t^1) = \theta(t^2) = 0$.

The quantity $2q(\alpha)$ appearing in $I_2(\gamma)$ is the quadratic form

$$2q(\alpha) = [g_{hk} + F_{p^i}(\dot{t}^2) T_{hk}^{i2} - F_{p^i}(\dot{t}^1) T_{hk}^{i1}] \alpha^h \alpha^k$$

evaluated on C_0 , where the subscripts h, k denote the derivatives with respect to α^h, α^k . The integrand $2\omega(\alpha, \eta, \pi)$ of $I_2(\gamma)$ is the second differential along C_0 of $F(a, y, \dot{p}, \lambda)$ with respect to the variables (a, y, \dot{p}) .

As has been shown by McShane [6, §13] no generality is lost in the strengthened condition of Jacobi if the variations γ were restricted to be of class C' . However for our purposes it is convenient to use the larger class described above.

The main theorem to be established is the following:

THEOREM 1. *If the arc C_0 satisfies, with a set of multipliers $\lambda^\beta(t)$, the Euler-Lagrange equations, the transversality condition, the strengthened conditions of Weierstrass and Jacobi, and is nonsingular, there is a function of the form*

$$(2.6) \quad J(C) = g(a) + \int_C F(a, y, \dot{y}, m(a, y, \dot{y})) dt$$

and a neighborhood \mathfrak{F} of C_0 in ay -space such that the inequality $J(C) > J(C_0)$ holds for every rectifiable arc $C \neq C_0$ in \mathfrak{F} satisfying the end conditions (1.2). If C also satisfies the differential equations (1.1), then $I(C) > I(C_0)$. The multipliers $m^\beta(a, y, \dot{y})$ can be chosen to be of class C''' .

It is understood of course that the elements $(a, y(t), \dot{y}(t))$ on C are in \mathfrak{R} for almost all values of the parameter t along C .

The corresponding theorem for the weak relative minimum is obtained by replacing the strengthened condition of Weierstrass by the following condition:

(4') C_0 satisfies the *strengthened condition of Clebsch*, that is, the quadratic form

$$(2.7) \quad F_{p^i p^i}(a, y, \dot{y}, \lambda) \pi^i \pi^i > 0$$

holds along C for every solution $\pi \neq \rho \dot{y}$ of the equations

$$\phi_{p^i}^\beta(a, y, \dot{y}) \pi^i = 0.$$

This condition implies that C_0 is nonsingular.

The theorem obtained is the following:

THEOREM 2. *If the arc C_0 satisfies, with a set of multipliers $\lambda^\beta(t)$, the Euler-Lagrange equations, the transversality condition, and the strengthened conditions of Clebsch and Jacobi, there is a function of the form (2.6) and a neighborhood \mathfrak{R}_0 of C_0 in ayp -space such that the inequality $J(C) > J(C_0)$ holds for every rectifiable arc $C \neq C_0$ in \mathfrak{R}_0 that satisfies the end conditions (1.2). If $\phi^\beta = 0$ along C , then $I(C) > I(C_0)$.*

By a neighborhood of \mathfrak{R}_0 of C_0 in ayp -space is meant an open subset of R containing the elements $(a_0, y_0(t), \dot{y}_0(t))$ on C_0 and having the property that if (a, y, \dot{y}) is in \mathfrak{R}_0 so also is $(a, y, k\dot{y})$ if $k > 0$. An arc C is in \mathfrak{R}_0 if almost all of its elements $(a, y(t), \dot{y}(t))$ are in \mathfrak{R}_0 .

3. Proof of Theorem 2. It will be convenient to prove Theorem 2 first and to use the results obtained in order to prove Theorem 1. We accordingly suppose that C_0 satisfies the conditions imposed upon it in Theorem 2. Let $\lambda^\beta(a, y)$ be a set of functions of class C''' such that

$$\lambda^\beta(a_0, y_0(t)) = \lambda^\beta(t),$$

where $\lambda^\beta(t)$ are the multipliers associated with C_0 . We shall show that a set

of multipliers of the form

$$(3.1) \quad m^\beta(a, y, p) = \lambda^\beta(a, y) + b\phi^\beta(a, y, p)(p^i p^i)^{-1/2}$$

are effective for $J(C)$ in Theorem 2, provided the constant b is chosen sufficiently large. More precisely, we shall show that if b is suitably chosen, then C_0 satisfies the Euler equations, the transversality condition and the strengthened conditions of Legendre (Clebsch) and Jacobi for the problem of minimizing the function

$$J(C) = g(a) + \int_c G(a, y, \dot{y}) dt$$

subject to the end conditions (1.2). Here

$$G(a, y, p) = F(a, y, p, m(a, y, p)).$$

Since $\phi^\beta = 0$ along C_0 we have

$$G_{p^i} = F_{p^i}, \quad G_{y^i} = F_{y^i}, \quad G_h = F_h$$

along C_0 . It follows that the Euler equations (2.1) and the transversality condition (2.2) hold along C_0 with F replaced by G , whatever value is assigned to b . If b is taken sufficiently large, say $b \geq b_0$, then, as has been shown by the author ([3, Theorem 3.1]; see also [8]), the inequality

$$(3.2) \quad G_{p^i p^i}(a, y, p) \pi^i \pi^i > 0 \quad (b \geq b_0)$$

will hold along C_0 for every set $\pi \neq \rho p$. It remains to show that b can be increased so that the second variation $J_2(\gamma)$ of $J(C)$ will be positive along C_0 for every variation γ , not essentially null, that satisfies the accessory end conditions (2.4). When this result has been established, Theorem 2 will follow from the sufficiency theorem for the case in which there are no differential side conditions.

A simple computation will reveal that the second variation $J_2(\gamma)$ of $J(C)$ along C_0 is given by the formula

$$J_2(\gamma) = P(\gamma) + bQ(\gamma),$$

where

$$P(\gamma) = I_2(\gamma) + 2 \int_{t_1}^{t_2} (\lambda_{\alpha\alpha}^\beta h + \lambda_{y^i y^i}^\beta) \Phi^\beta dt,$$

$$Q(\gamma) = 2 \int_{t_1}^{t_2} \Phi^\beta \Phi^\beta (\dot{y}_0^i \dot{y}_0^i)^{-1/2} dt,$$

and Φ^β is the variation of ϕ^β along C_0 . It should be recalled that by virtue of the homogeneity properties of f and ϕ^β , the values of $P(\gamma)$, $Q(\gamma)$ and $J_2(\gamma)$ are unchanged if we add to γ an essentially null variation (2.5). We can accord-

ingly restrict ourselves to variations that satisfy, with constants c and d , a condition of the form

$$\dot{y}_0(t)\dot{\eta}(t) = ct + d$$

on $t^{1/2}$. The class of variations of this type that satisfy the accessory end conditions (2.4) will be denoted by Γ . The only essentially null variation in Γ is the null variation.

The quadratic functions $P(\gamma)$ and $Q(\gamma)$ defined above have the following three properties:

- (1) $Q(\gamma) \geq 0$ on Γ ;
- (2) $P(\gamma) > 0$ for every non-null variation γ in Γ for which $Q(\gamma) = 0$;
- (3) There is a constant b_0 such that the strengthened condition of Legendre holds for $P(\gamma) + b_0Q(\gamma)$.

Theorem 2 will be established if we show that these properties imply the existence of a constant $b \geq b_0$ such that $P(\gamma) + bQ(\gamma)$ is positive for every non-null variation γ in Γ . In the proof we can and shall assume that $b_0 = 0$. This is permissible since the pair $P(\gamma) + b_0Q(\gamma)$, $Q(\gamma)$ have the properties (1), (2), (3) of $P(\gamma)$, $Q(\gamma)$ with $b_0 = 0$.

If a constant b of the type just described fails to exist, there is for every integer q a variation

$$\gamma_q: \quad \alpha_q, \quad \eta_q(t) \quad (t^1 \leq t \leq t^2)$$

in Γ satisfying the inequality

$$(3.3) \quad P(\gamma_q) + qQ(\gamma_q) \leq 0.$$

Since $P(\gamma)$ and $Q(\gamma)$ are homogeneous in γ we can suppose that γ_q satisfies the condition

$$(3.4) \quad |\alpha_q|^2 + \max_{t^1 \leq t \leq t^2} |\eta_q(t)|^2 + \int_{t^1}^{t^2} |\dot{\eta}_q(t)|^2 dt = 1,$$

where a pair of vertical bars denotes the usual norm of a vector, namely, the square root of the sum of the squares of its components. Since γ_q is in Γ there exist constants c_q and d_q such that one has

$$(3.5) \quad \dot{y}_0(t)\dot{\eta}_q(t) = c_q t + d_q$$

on $t^{1/2}$. As was shown by McShane [6, §5] the relations (3.4) and (3.5) imply the existence of a subsequence (which for our purposes we can and shall suppose to be the original sequence) that converges uniformly on $t^{1/2}$ to a variation

$$\gamma_0: \quad \alpha_0, \quad \eta_0(t) \quad (t^1 \leq t \leq t^2)$$

satisfying with constants c_0 , d_0 the relation

$$\dot{\gamma}_0^i(t)\eta_0^i(t) = c_0t + d_0$$

on t^1t^2 . The functions $\eta_0^i(t)$ are absolutely continuous with integrable square derivatives. Since the variations γ_q satisfy the accessory end conditions (2.4), the variation γ_0 also has this property. It follows that γ_0 is in Γ . In view of the fact that $P(\gamma)$ satisfies the strengthened condition of Legendre, b_0 being zero, it is lower semi-continuous [6, §7]. Hence

$$(3.6) \quad \liminf_{q \rightarrow \infty} P(\gamma_q) \geq P(\gamma_0)$$

and, by (3.3),

$$P(\gamma_0) + \liminf_{q \rightarrow \infty} qQ(\gamma_q) \leq 0.$$

Since $Q(\gamma_q) \geq 0$ this relation can be satisfied only if

$$(3.7) \quad P(\gamma_0) \leq 0, \quad \liminf_{q \rightarrow \infty} Q(\gamma_q) = 0.$$

But the relation $Q(\gamma) \geq 0$ implies that $Q(\gamma)$ satisfies the Legendre condition and is accordingly lower semicontinuous. The last equation therefore tells us that $Q(\gamma_0) = 0$ and hence that $P(\gamma_0) \geq 0$, by property (2) for P and Q . It follows from (3.7) that $P(\gamma_0) = 0$. Consequently γ_0 is null and

$$(3.8) \quad \lim_{q \rightarrow \infty} \alpha_q^h = 0, \quad \lim_{q \rightarrow \infty} \eta_q^i(t) \equiv 0 \text{ (uniformly)}, \quad \lim_{q \rightarrow \infty} c_q = 0.$$

Using the relation (3.5) we find by differentiation that the relation

$$\dot{\gamma}_0^i \eta_q^i + \dot{\gamma}_0^i \dot{\eta}_q^i = c_q$$

holds for almost all values of t on t^1t^2 . This with (3.8) tells us that

$$(3.9) \quad \lim_{q \rightarrow \infty} \dot{\gamma}_0^i(t) \dot{\eta}_q^i(t) = 0$$

uniformly on t^1t^2 , if we set $\dot{\eta}_q^i = 0$ where it is not defined.

Turning now to the relations (3.3), (3.6), $P(\gamma_0) = 0$, and $Q(\gamma) \geq 0$ we find that

$$\lim_{q \rightarrow \infty} P(\gamma_q) = 0,$$

which by virtue of (3.8) takes the form

$$(3.10) \quad \lim_{q \rightarrow \infty} \int_{t^1}^{t^2} A_{ij} \dot{\eta}_q^i \dot{\eta}_q^j dt = 0,$$

the integrand being the Legendre form for $P(\gamma)$. Since the strengthened condition of Legendre holds for $P(\gamma)$ there is a constant $\epsilon > 0$ such that

$$A_{ij}(t)\pi^i\pi^j + (\dot{y}_0\pi^i)^2 \geq \epsilon\pi^i\pi^i$$

on t^2 . It follows from (3.9) and (3.10) that

$$0 = \lim_{q=\infty} \int_{\mu^1}^{t^2} \{A_{ij}\dot{\eta}_q^i\dot{\eta}_q^j + (\dot{y}_0\dot{\eta}_q^i)^2\} dt \geq \epsilon \limsup_{q=\infty} \int_{\mu^1}^{t^2} |\dot{\eta}_q|^2 dt.$$

We have accordingly

$$\lim_{q=\infty} \int_{\mu^1}^{t^2} |\dot{\eta}_q|^2 dt = 0.$$

This is inconsistent with the relations (3.4) which tell us that

$$\lim_{q=\infty} \int_{\mu^1}^{t^2} |\dot{\eta}_q|^2 dt = 1.$$

It follows that there is a constant $b > 0$ such that $P(\gamma) + bQ(\gamma) > 0$ for every non-null variation γ in Γ . This completes the proof of Theorem 2.

It should be observed that apart from the derivation of properties (1), (2), (3), for $P(\gamma)$ and $Q(\gamma)$ we have not used the special forms of the function, except that they be of the type of the second variation. For example, the result would still be valid if we take $Q(\gamma)$ of the form

$$Q(\gamma) = \alpha^h\alpha^h + \int_{\mu^1}^t \eta^i\eta^i dt,$$

giving a well known result in the calculus of variations.

4. Proof of Theorem 1. In order to prove Theorem 1 we observe first that the hypotheses of Theorem 1 imply those of Theorem 2. We accordingly select multipliers $m^\beta(a, y, p)$ in the manner described in the last section. This yields, as the integrand, the function

$$G(a, y, p) = f(a, y, p) + m^\beta(a, y, p)\phi^\beta(a, y, p).$$

Because of the strengthened condition of Weierstrass and the nonsingularity of C_0 the function $G(a, y, p)$ E -dominates (see [3, §9]) the integrand $L(p)$ of the length integral near C_0 on the set of values (a, y, p) for which $\phi^\beta = 0$. The author has shown [3, §8] that there is therefore a function $\theta(a, y, p)$ of class C^∞ such that the E -function $E_H(a, y, p, q)$ for the integrand

$$H(a, y, p) = G(a, y, p) + \theta\phi^\beta\phi^\beta$$

is positive for all (a, y, p) in a neighborhood \mathfrak{R}_0 of those on C_0 and all $(a, y, q) \neq (a, y, kp)$ ($k > 0$) in \mathfrak{R} . It is easily seen that the new function

$$J(C) = g(a) + \int_C H(a, y, \dot{y}) dt$$

retains the properties of the function $J(C)$ described in the last section. The arc C_0 therefore satisfies the Euler equation, the transversality condition, the strengthened conditions of Weierstrass, Legendre, and Jacobi for the problem of minimizing $J(C)$ subject to the end conditions (1.2). The function $J(C)$ therefore has the properties described in Theorem 1, as was to be proved.

REFERENCES

1. G. A. Bliss, *Lectures on the calculus of variations*, University of Chicago Press, 1946.
2. M. R. Hestenes, *Generalized problem of Bolza in the calculus of variations*, Duke Math. J. vol. 5 (1939) pp. 309–324.
3. ———, *The Weierstrass E-function in the calculus of variations*, Trans. Amer. Math. Soc. vol. 60 (1946) pp. 51–71.
4. ———, *Theorem of Lindeberg in the calculus of variations*, Trans. Amer. Math. Soc. vol. 60 (1946) pp. 72–92.
5. ———, *Sufficient conditions for the isoperimetric problem of Bolza in the calculus of variations*, Trans. Amer. Math. Soc. vol. 60 (1946) pp. 93–118.
6. E. J. McShane, *Sufficient conditions for a weak relative minimum in the problem of Bolza*, Trans. Amer. Math. Soc. vol. 52 (1942) pp. 344–379.
7. F. G. Myers, *Sufficient conditions for the problem of Lagrange*, Duke Math. J. vol. 10 (1943) pp. 93–97.
8. W. T. Reid, *Isoperimetric problems of Bolza in nonparametric form*, Duke Math. J. vol. 5 (1939) pp. 675–691.

UNIVERSITY OF CHICAGO,
CHICAGO, ILL.