A DIRECT SUFFICIENCY PROOF FOR THE
PROBLEM OF BOLZA IN THE CALCULUS
OF VARIATIONS*

BY
MAGNUS R. HESTENES

1. Introduction. The problem to be considered is that of minimizing an
expression of the form
\[ J = g[x^1, y(x^1), x^2, y(x^2)] + \int_{a}^{b} f(x, y, y')dx \]
in a class of arcs
\[ y_i = y_i(x) \quad (x^1 \leq x \leq x^2; i = 1, \ldots, n) \]
satisfying a set of differential equations
\[ \phi_\beta(x, y, y') = 0 \quad (\beta = 1, \ldots, m < n) \]
and having its end values \([x^1, y(x^1), x^2, y(x^2)]\) lying on a surface \(S\) in \((x^1, y^1, x^2, y^2)\)-space. The surface \(S\) may be defined by equations of the form
\[ \psi_\mu [x^1, y(x^1), x^2, y(x^2)] = 0 \quad (\mu = 1, \ldots, \rho \leq 2n + 2) \]
or by parametric equations
\[ x^s = x^s(\alpha_1, \ldots, \alpha_\rho), \quad y_i^s = y_i^s(\alpha_1, \ldots, \alpha_\rho) \quad (s = 1, 2). \]
In the latter case the function \(g\) of the end values may be considered as a
function \(g(\alpha)\) of the parameters \(\alpha_1, \ldots, \alpha_\rho\). The functional \(J\) can then be
taken in the form
\[ J = g(\alpha) + \int_{a}^{b} f(x, y, y')dx. \]
It is immaterial which of the forms (1.2) and (1.3) is used. The theory for the
one can readily be obtained from that of the other.

A sufficiency proof for the problem of Bolza independent of assumptions
of normality has recently been given by the author (IV)† and in a somewhat
modified form by Bliss (VII). The proof given is indirect in the sense that

* Presented to the Society, November 28, 1936; received by the editors November 9, 1936.
The results here given were obtained at the University of Chicago during the Summer Quarter, 1936,
and came as a discussion of the problem with Professor Bliss.
† Roman numerals in parentheses refer to the references at the end of the paper.
it is applicable only after the problem has been transformed into an equivalent problem of somewhat different form. This proof, although adequate, is not entirely satisfactory. For a satisfactory theory one should have a sufficiency proof which can be applied directly to the problem at hand. The purpose of this paper is to give such a proof. The proof here given is essentially a generalization of the earlier one and makes use of the notion of families of Mayer fields which is the basis of Hahn’s theorem as presented by Bliss (II). The only other sufficiency proof for the problem of Bolza known to the author which is independent of assumptions of normality is the one recently announced by Reid (IX). This proof has the same limitations as the earlier proof of the author.*

2. Preliminary remarks. The hypotheses upon which the analysis of the present paper is based and the terminology used will be that of the author (IV) in case the end conditions are of the form (1.3) and that of Bliss (VII) in case the end conditions are of the form (1.2). The definition of admissible arcs here used will be that of Bliss; that is, an arc (1.1) will be said to be admissible if it is continuous and is composed of a finite number of subarcs having a continuously turning tangent, satisfying the equations \( \phi_\theta = 0 \), and having its elements \((x, y, y')\) all admissible. As is customary we center our attention on a particular admissible arc \( E_0 \) with end values \((x_1, y_1, x_2, y_2)\) on \( S \). We shall assume that the parameters \((\alpha)\) in equations (1.3) have been chosen so that the end values of \( E_0 \) are given by the values \((\alpha) = (0)\). We shall assume further that the matrix of the derivatives of the second members of equations (1.3) has rank \( r \) at \((\alpha) = (0)\).

The arc \( E_0 \) will be said to satisfy the condition I with a set of multipliers \( \lambda_0, \lambda_\theta(x) \) if the Euler-Lagrange equations

\[
F_{\psi'} = \int_a^b F_{\psi'} dx + c_i, \quad \phi_\theta = 0 \quad (i = 1, \ldots, n)
\]

hold along \( E_0 \), where the \( c_i \)'s are constants and

\[
F(x, y, y', \lambda) = \lambda_0 f + \lambda_\theta(x) \phi_\theta \quad (\beta = 1, \ldots, m)
\]

and if further the transversality condition

\[
(F - y[F_{\psi'}] dx + F_{\psi'} dy_1)^2 + \lambda_0 dg = 0
\]

holds on \( E_0 \) for all directions \( dx^1, dy_1, dx^2, dy_2 \) tangent to \( S \) at the point on \( S \) determined by the end values of \( E_0 \). Here and elsewhere it will be understood

* Added in proof. Reid has recently removed these limitations in a paper entitled *A direct expansion proof of sufficient conditions for the non-parametric problem of Bolza* to appear in an early issue of these Transactions.
that the multiplier \( \lambda_0 \) is a constant and that the multipliers \( \lambda_\beta(x) \) have continuity properties like those of the functions \( y'_i(x) \) belonging to \( E_0 \). An admissible arc (1.1) and a set of multipliers \( \lambda_0, \lambda_\beta(x) \) having continuous derivatives \( y'_i, y''_i, \lambda'_\beta \) will be called an extremal if it satisfies equations (2.1) with a set of constants \( c_i \).

The arc \( E_0 \) will be said to satisfy the Weierstrass condition II with a set of multipliers \( \lambda_0, \lambda_\beta(x) \) if at each element \((x, y, y', \lambda)\) in a neighborhood \( \mathcal{N} \) of those on \( E_0 \) and having \((x, y, y')\) admissible, the inequality

\[
E(x, y, y', \lambda, Y') > 0
\]

holds for every admissible set \((x, y, Y') \neq (x, y, y')\), where

\[
E = F(x, y, Y', \lambda) - F(x, y, y', \lambda) - (Y'_i - y'_i)F_{y'_i}(x, y, y', \lambda).
\]

The arc \( E_0 \) will be said to satisfy the Clebsch condition III' if at each element \((x, y, y', \lambda)\) on \( E_0 \) the inequality

\[
F_{y'_i}y'_k \pi_i \pi_k > 0 \quad (i, k = 1, \ldots, n)
\]

holds for every set of constants \((\pi) \neq (0)\) satisfying the conditions \( \phi_{y'_i}(0) = 0 \).

The arc \( E_0 \) will be said to satisfy the condition IV with a set of multipliers \( \lambda_0 = 1, \lambda_\beta(x) \) if the second variation \( J_2 \) of \( J \) formed with these multipliers is positive definite along \( E_0 \). In order to define this condition more precisely, recall that when the end conditions are taken in the form (1.3) the second variation \( J_2 \) of \( J \) along \( E_0 \) is expressible in the form (II, pp. 520–521).

\[
J_2(\eta, w) = b_{hi}w_i w_l + \int_{x_1}^{x_2} 2\omega(x, \eta, \eta') dx,
\]

where \( h, l = 1, \ldots, r; i, k = 1, \ldots, n; s \) not summed,

\[
2\omega = F_{y'_i}y'_k \pi_i \pi_k + F_{y'_i}y'_k \pi_i \pi_k + F_{y'_i}y'_k \pi_i \pi_k
\]

(2.3)

\[
b_{hi} = g_{hi} + [(F - y'_i F_{y'_i}) x'_i x'_l + F_{y'_i} y''_{i\alpha_i} x_{i\alpha_i}^{\sum_{i=1}}
+ [(F_{x} - y'_i F_{x}) x'_l x'_l + F_{x} x'_l y''_{i\alpha_i} + x'_l y''_{i\alpha_i})]^{\sum_{i=1}}.
\]

In the last expression the subscripts \( h, l \) denote the derivatives of the functions \( g(\alpha), x'(\alpha), y''(\alpha) \) with respect to \( \alpha_h, \alpha_l \) at \( (\alpha) = (0) \). The variations \( \eta_i(x) \) are assumed to have derivatives like those of the functions \( y_i(x) \) and the \( w_i \)'s are constants. The second variation \( J_2(\eta, w) \) is said to be positive definite along \( E_0 \) if the inequality \( J_2(\eta, w) > 0 \) holds for every set of variations \((\eta, w) \neq (0, 0)\) satisfying along \( E_0 \) the differential equations

\[
\Phi_\beta(x, \eta, \eta') = \phi_{y'_i} \eta_i + \phi_{y'_i} \eta'_i = 0 \quad (\beta = 1, \ldots, m)
\]

and end conditions (\( s \) not summed; \( s = 1, 2 \))
The functions \( y'_i(x) \) being the slope functions of \( E_0 \).

The purpose of the present paper is to give a direct proof of the following theorem:

**Theorem 2.1.** If the arc \( E_0 \) has no corners and satisfies the conditions I, II', III', IV', there exist neighborhoods \( \tilde{g} \) of \( E_0 \) in \( xy \)-space and \( N \) of the end values of \( E_0 \) in \( (x^1, y^1, x^2, y^2) \)-space such that the inequality \( J(C) > J(E_0) \) holds for every admissible arc \( C \) in \( \tilde{g} \) with end values on \( S \) in \( N \) and not identical with \( E_0 \).

The condition IV' can be replaced by other conditions, commonly called conditions of Mayer, which together with the condition III' imply the condition IV' here defined. A discussion of these results has been given by Bliss (VII, pp. 95–140) and the author (IV, pp. 810–813; VIII). One should also consult the papers of Morse (V) and Reid (VI).

The differential equations associated with the second variation are given by the equations

\[
\eta_i(x^0) = c^i_{xk}w_k, \quad c^i_{xk} = y^i_{xk} - y^i'_i(x^0)x^k_i,
\]

(2.4) the functions \( y^i_x(x) \) being the slope functions of \( E_0 \).

Theorem 2.1. If the arc \( E_0 \) has no corners and satisfies the conditions I, II', III', IV', there exist neighborhoods \( \tilde{g} \) of \( E_0 \) in \( xy \)-space and \( N \) of the end values of \( E_0 \) in \( (x^1, y^1, x^2, y^2) \)-space such that the inequality \( J(C) > J(E_0) \) holds for every admissible arc \( C \) in \( \tilde{g} \) with end values on \( S \) in \( N \) and not identical with \( E_0 \).

The condition IV' can be replaced by other conditions, commonly called conditions of Mayer, which together with the condition III' imply the condition IV' here defined. A discussion of these results has been given by Bliss (VII, pp. 95–140) and the author (IV, pp. 810–813; VIII). One should also consult the papers of Morse (V) and Reid (VI).

The differential equations associated with the second variation are given by the equations

\[
(2.5) \quad \Omega \eta_i - (d/dx)\Omega \eta_i' = 0, \quad \Phi_i = 0,
\]

where \( \Omega = \omega + \mu(x)\Phi_i \). When the condition III' is satisfied, a solution \( \eta, \mu(x) \) of equations (2.5) is completely determined by the functions \( \eta_i, \zeta_i = \Omega_n' \). A system \( \eta_{ik}, \zeta_{ik} (i, k = 1, \ldots, n) \) of \( n \) linearly dependent solutions of equations (2.5) will be called a conjugate system if the equations

\[
(2.6) \quad \zeta_{ij} \eta_{ik} - \zeta_{ik} \eta_{ij} = 0 \quad (i, j, k = 1, \ldots, n)
\]

hold identically on \( x^1, x^2 \).

### 3. Fundamental sufficiency theorems

In this section we shall assume that the arc \( E_0 \) forms an extremal arc with a set of multipliers \( \lambda_0 = 1, \lambda_0(x) \). The sufficiency theorems here given depend upon the notion of families of Mayer fields. The definition of Mayer fields here used is that of Bliss (I, p. 730). Let \( \tilde{g} \) and \( N \) be respectively neighborhoods of \( E_0 \) in \( xy \)-space and of the end values of \( E_0 \) in \( (x^1, y^1, x^2, y^2) \)-space such that the points \( (x^1, y^1, x^2, y^2) \) determined by points in \( N \) are in \( \tilde{g} \). Suppose that for each fixed point \( (x^1, y^1, x^2, y^2) \) in \( N \) the functions

\[
(3.1) \quad \rho_i(x, y, x^1, y^1, x^2, y^2), \quad l_0 = 1, \quad l_0(x, y, x^1, y^1, x^2, y^2)
\]

are the slope functions and multipliers of a Mayer field over the neighborhood \( \tilde{g} \) of \( E_0 \) and that \( E_0 \) is an extremal of the field determined by its end values \( (x^0, y^0, x^0, y^0) \). Let \( W(x^1, y^1, x^2, y^2) \) be the function defined over \( N \) by the equation
\[ W(x^1, y^1, x^2, y^2) = g(x^1, y^1, x^2, y^2) + \int_{(x^1, y^1)}^{(x^2, y^2)} \left\{ Fdx + (dy_i - p_idx)Fy_i \right\}, \]

where the integral in this expression is the Hilbert integral formed with the functions (3.1) and evaluated along an arc in \( \mathcal{S} \) joining the points \((x^1, y^1), (x^2, y^2)\). Since \( E_0 \) is an extremal of the field determined by its end values, it follows that

\[ J(E_0) = W(x^1_0, y^1_0, x^2_0, y^2_0). \]

A first sufficiency theorem is the following (cf. VII, p. 107).

**Theorem 3.1.** If at each point \((x, y)\) in \( \mathcal{S} \) and each element \((x^1, y^1, x^2, y^2)\) on \( S \) in \( N \) the Weierstrass \( E \)-function formed with the functions (3.1) satisfies the condition

\[ E(x, y, p, l, y') > 0 \]

for every admissible set \((x, y, y') \neq (x, y, p)\) and if the end values of \( E_0 \) afford a proper minimum for the function \( W(x^1, y^1, x^2, y^2) \) on \( S \) in \( N \), then the inequality \( J(C) > J(E_0) \) holds for every admissible arc \( C \) in \( \mathcal{S} \) with end values on \( S \) in \( N \) and not identical with \( E_0 \).

For from the definitions of the function \( W \) and the Weierstrass \( E \)-function it follows readily that

\[ J(C) = W(x^1_0, y^1_0, x^2_0, y^2_0) + \int_{c} E(x, y, p, l, y')dx, \]

where \((x^1_0, y^1_0, x^2_0, y^2_0)\) are the end values of \( C \) and the functions \( p_i, l_0, l_0 \) appearing in the \( E \)-function are the functions (3.1) determined by the end values of \( C \). Since the end values of the arc \( E_0 \) afford a proper minimum to the function \( W \) on \( S \) in \( N \) it follows from the relations (3.2), (3.3), and the last equation that

\[ J(C) \leq W(x^1, y^1, x^2, y^2) \leq W(x^1_0, y^1_0, x^2_0, y^2_0) = J(E_0), \]

the equality holding in both cases only in case the end values of \( C \) and \( E_0 \) coincide and the equations \( y_i' = p_i \) hold along \( C \). But in this case the arc \( C \) would be an extremal of the field determined by the common end values of \( C \) and \( E_0 \) and hence would coincide with \( E_0 \) since there is but one extremal of the field through each point of \( \mathcal{S} \). This proves the theorem.

When the surface \( S \) is defined by the parametric equations (1.3), the parameter values \((\alpha) = (0)\) giving the end values of \( E_0 \), the sufficiency theorem described above can be stated in somewhat different form. Let \( \mathcal{S} \) be a neighborhood of \( E_0 \) in \( xy \)-space and \( A \) be a neighborhood of \((\alpha) = (0)\) in \((\alpha_1, \cdots, \alpha_r)\).
space such that the points \((x^1, y^1), (x^2, y^2)\) determined by points on \(S\) with parameter values \((\alpha)\) in \(A\) are in \(\mathcal{H}\). Suppose that for each element \((\alpha)\) in \(A\) the functions
\[
(3.5) \quad p_i(x, y, \alpha), \quad l_0 = 1, \quad l_\beta(x, y, \alpha)
\]
are the slope functions and multipliers of a Mayer field defined over \(\mathcal{H}\) and that \(E_0\) is an extremal of the field determined by \((\alpha) = (0)\). Let \(W(\alpha_1, \cdots, \alpha_r)\) be the function defined by the equation
\[
(3.6) \quad W(\alpha) = g(x^1, y^1, x^2, y^2) + \int_{(x^1, y^1)}^{(x^2, y^2)} \left\{ Fdx + (dy_i - p_i dx)F_{\nu_i} \right\},
\]
where \((x^1, y^1, x^2, y^2)\) is the point \((1.3)\) on \(S\) determined by \((\alpha_1, \cdots, \alpha_r)\) and the integral is the Hilbert integral formed with the functions \((3.5)\) and evaluated along an arc in \(\mathcal{H}\) joining the points \((x^1, y^1), (x^2, y^2)\). One readily verifies that \(J(E_0) = W(0)\). The following sufficiency theorem can now be established by an argument like that given above.

**Theorem 3.2.** If at each point \((x, y)\) in \(\mathcal{H}\) and each element \((\alpha)\) in \(A\) the inequality
\[
(3.7) \quad E[x, y, p(x, y, \alpha), l(x, y, \alpha), y'] > 0
\]
holds for every admissible element \((x, y, y') \neq (x, y, p)\) and the element \((\alpha) = (0)\) affords a proper relative minimum to the function \(W(\alpha)\), then the inequality \(J(C) > J(E_0)\) is true for every admissible arc \(C\) in \(\mathcal{H}\) not identical with \(E_0\) and having its end values on \(S\) in a sufficiently small neighborhood \(N\) of the end values of \(E_0\).

One readily verifies that Theorems 3.1 and 3.2 are equivalent whenever, as we have assumed, the surface \(S\) is non-singular. The earlier sufficiency theorem given by Hestenes (IV, p. 805) and in somewhat different form by Bliss (VII, p. 107) is the special case in which the functions \((3.1)\) and \((3.5)\) depend only on the values \((x, y)\). When this earlier theorem is applied to the problem of Bolza with separated end conditions obtained from the original problem of Bolza here described by a suitably chosen transformation, and the result is reinterpreted in terms of the original problem, one obtains a theorem of the type here described.

4. **Two lemmas.** The proof of Theorem 2.1 to be given in the next section will be based on Theorem 3.2 and two lemmas, the first of which is the following:

**Lemma 4.1.** Let \(E_0\) be a non-singular extremal arc with multipliers \(\lambda_0 = 1, \lambda_\beta(x)\) for which there exists a conjugate system \(\eta_{ik}, \zeta_{ik}\) of solutions of the accessory
equations (2.5) whose determinant \( |\eta_{ik}(x)\| \) is different from zero along \( E_0 \). Let \( u_{ih}, v_{ih} (h = 1, \ldots, r) \) be a set of \( r \) further solutions of equations (2.5). There exists an \((n+r)\)-parameter family of extremals
\[
y_i = y_i(x, a_1, \ldots, a_n, \alpha_1, \ldots, \alpha_r), \\
\lambda_0 = 1, \quad \lambda_\beta = \lambda_\beta(x, a_1, \ldots, a_n, \alpha_1, \ldots, \alpha_r)
\]
containing \( E_0 \) for values \( x_0^0 \leq x \leq x_0^\delta \), \( a_i = 0, \alpha_i = 0 \) and such that the functions \( y_i, y_{ia}, z_i = F_{ui'}, z_{ia} \) have continuous first and second partial derivatives in a neighborhood of the values \((x, a, \alpha)\) belonging to \( E_0 \) and have as its variations along \( E_0 \)
\[
y_{ia} = \eta_{ik}, \quad z_{ia} = \xi_{ik}, \quad y_{ia} = u_{ik}, \quad z_{ia} = v_{ik}.
\]
Moreover there is a neighborhood \( \mathfrak{F} \) of \( E_0 \) in \( xy\)-space and \( A \) of \((\alpha) = (0)\) such that for every fixed set \((\alpha)\) in \( A \) the family (4.1) defines a Mayer field over \( \mathfrak{F} \). The slope functions and multipliers of the field are given by the equations
\[
p_i(x, y, \alpha) = y_{ia}[x, a(x, y, \alpha), \alpha], \\
l_0 = 1, \quad l_\beta(x, y, \alpha) = \lambda_\beta[x, a(x, y, \alpha), \alpha],
\]
where the values \( a_i(x, y, \alpha) \) are determined by the first \( n \) of the equations (4.1). The arc \( E_0 \) is an extremal of the field determined by the set \((\alpha) = (0)\).

In order to establish this result let
\[
y_i = Y_i(x, b_1, \ldots, b_n, c_1, \ldots, c_n) \\
\lambda_0 = 1, \quad \lambda_\beta = \Lambda_\beta(x, b_1, \ldots, b_n, c_1, \ldots, c_n)
\]
be a \(2n\)-parameter family of extremals containing \( E_0 \) for values \( b_i = b_{i0}, c_i = c_{i0}, x_0^0 \leq x \leq x_0^\delta \) and such that the parameters \( b_i, c_i \) are the values of the functions \( Y_i, Z_i = F_{ui'}, Y_{ia}, Z_{ia}, \Lambda_\beta \) at \( x = x_0^\delta \), the functions \( Y_i, Y_{ia}, Z_i, Z_{ia}, \Lambda_\beta \) having continuous first and second partial derivatives in a neighborhood of the values \((x, b, c)\) belonging to \( E_0 \). The \((n+r)\)-parameter family (4.1) obtained from the family (4.4) by setting
\[
b_i = b_{i0} + \eta_{ik}(x_0^\delta) a_k + u_{ik}(x_0^\delta) \alpha_k, \\
c_i = c_{i0} + \xi_{ik}(x_0^\delta) a_k + v_{ik}(x_0^\delta) \alpha_k
\]
can now be shown to have the properties described in the lemma. For, by differentiating the identities
\[
b_i = y_i(x_0^\delta, a, \alpha), \quad c_i = z_i(x_0^\delta, a, \alpha),
\]
where \( z_i = F_{ui'} \), with respect to \( a_k \) and \( \alpha_k \) it is found that the relations (4.2) hold at \( x = x_0^\delta \) and hence also along \( E_0 \), since along \( E_0 \) the right and left mem-
bers of equations (4.2) are solutions of the accessory equations (2.5) with the same initial values at \( x = x_0^1 \). The determinant \( \frac{\partial y}{\partial \alpha} \) is accordingly different from zero along \( E_0 \). It follows that the first \( n \) equations (4.1) have unique solutions \( \hat{u}_i(x, y, \alpha) \) which are continuous and have continuous first and second partial derivatives for all \( (x, y, \alpha) \) with \( (x, y) \) in a neighborhood \( \mathcal{G} \) of \( E_0 \) and \( (\alpha) \) in a neighborhood \( A \) of \( (\alpha) = (0) \). For each fixed set \( (\alpha) \) in \( A \) the family (4.1) not only simply covers \( \mathcal{G} \) but also defines a Mayer field over \( \mathcal{G} \) with the slope functions and multipliers (4.3). This follows because in view of equations (4.5), \( \alpha_h = \text{const.}, (2.6) \) the Hilbert integral takes the form

\[
\int F_{\nu,i} d\nu_i = \int c_i d\beta_i = \frac{1}{2} \int d\left[(2c_i - \xi_{ij}a_j)\eta_{ik}a_k\right]
\]

on the hyperplane \( x = x_0^1 \) and is accordingly independent of the path on \( x = x_1^1 \) and hence also in \( \mathcal{G} \) (I, p. 733). This completes the proof of Lemma 4.1.

A second lemma is the following. (Cf. VII, p. 110.)

**Lemma 4.2.** Let \( E_0 \) be an extremal arc having the properties described in Lemma 4.1. The function \( W(\alpha) \) defined by equations (3.6) and (4.3) will have a proper relative minimum at \( (\alpha) = (0) \) if \( E_0 \) satisfies the transversality condition (2.2) and the condition \( Q(w) > 0 \) for every set \( (w) \neq (0) \), where

\[
Q(w) = b_{hi}w_hw_i + \left[c_i(\xi_{ij}d_k + v_i) + (c_i - u_i)(v_i - \xi_{ij}e_i)\right]^2
\]

\[
(4.7) \quad u_i = u_{ih}w_h, \quad v_i = v_{ih}w_h, \quad c_i = c_{ih}w_h, \quad d_i = d_{ih}w_h, \quad e_i = e_{ih}w_h,
\]

\[
(b_{hi} \text{ are constants defined by equations (2.3), and } c_{ih}(x), d_{ih}(x), e_{ih}(x) \text{ are functions satisfying the conditions } (s=1, 2))
\]

It is well known that the function \( W(\alpha) \) will have a proper minimum at \( (\alpha) = (0) \) whenever the conditions \( dW = 0, d^2W > 0 \) hold at \( (\alpha) = (0) \) for all differentials \( d\alpha_1, \ldots, d\alpha_r \) not all zero. The lemma will accordingly be established if we show that the equations

\[
(4.9) \quad dW = 0, \quad d^2W = Q(d\alpha)
\]

are identities in \( d\alpha_1, \ldots, d\alpha_r \). To this end let \( \delta \) denote the differential with respect to the variables \( \alpha_1, \ldots, \alpha_r \) of a function of \( (x, y, \alpha) \). Thus

\[
\delta p_i = p_i \delta \alpha_h, \quad \delta F_{\nu,i} = \delta \alpha_h \left(\frac{\partial}{\partial \alpha_h}\right) F_{\nu,i}(x, y, p_i, l).
\]

We then have
\[ dW = dg + \left[ (F - p_1 F y_i)dx + F y_i dy_i \right]^2 + \int_{\alpha_0}^{\alpha_1} (dy_i - p_1 dx) \delta F y_i, \]

where here and elsewhere the differentials \( dx^1, dy^1, dx^2, dy^2 \) are to be evaluated in terms of \( da_\alpha \) by means of equations (1.3). This expression vanishes identically at \( \alpha = 0 \), since \( E_0 \) satisfies the transversality condition (2.2) and the differential equations \( y'_i = p_i(x, y, 0) \) of the field determined by \( \alpha = 0 \). This proves the first relation (4.9). In order to prove the second relation we note first that at \( \alpha = 0 \)
\[
\begin{align*}
\frac{d^2 W}{d\alpha^2} & = d^2 g + \left[ (F - p_1 F y_i) d^2 x + F y_i d^2 y_i + d(F - p_1 F y_i) dx + dF y_i dy_i \right]^2 \\
& + \left[ (dy_i - p_1 dx) \delta F y_i \right]^2 + \int_{E_0} \left\{ - \delta p_1 \delta F y_i dx + (dy_i - p_1 dx) \delta^2 F y_i \right\}.
\end{align*}
\]

With the help of equations (2.3),
\[
\frac{d}{dx} (F - p_1 F y_i) = F x dx + F y_i dy_i - p_1 F y_i,
\]
and the relations \( y'_i = p_i \) along \( E_0 \) it is found that
\[
\frac{d^2 W}{d\alpha^2} = b_{hi} d\alpha_0 d\alpha_1 + \left[ (dy_i - p_1 dx) (dF y_i - F y_i dx + \delta F y_i) \right]^2 - \int_{E_0} \delta p_1 \delta F y_i dx.
\]

It is clear that this expression for \( d^2 W \) will reduce to the form \( Q(\alpha) \) if we establish the relations
\[
\begin{align*}
\left[ (dy_i - p_1 dx) (dF y_i - F y_i dx) \right]^2_1 &= \left[ c_i (\xi_i d_x + v_i) \right]^2_1, \\
\left[ (dy_i - p_1 dx) \delta F y_i \right]^2_1 &= \left[ c_i (v_i - \xi_i e_\alpha) \right]^2_1, \\
\int_{E_0} \delta p_1 \delta F y_i dx &= \left[ u_i (v_i - \xi_i e_\alpha) \right]^2_1,
\end{align*}
\]
where the functions \( c_i, d_i, e_i, u_i, v_i \) are defined by equations (4.7) with \( \omega_h = da_\alpha \). As a first step in the proof of these equations we note that in view of equations (2.4) and \( y'_i = p_i \) along \( E_0 \) we have at \( \alpha = 0 \)
\[
\begin{align*}
\left. \frac{dy_i}{dx} \right|_{\alpha = 0} &= p_i(x^s, y^s, 0) dx^s = c_i d\alpha_h = c_i(x^s) \quad (s = 1, 2),
\end{align*}
\]
where \( x^s, y^s \) and their differentials are to be evaluated by means of equations (1.3). Moreover if we set \( \delta a_i = a_{can} da_\alpha \) and let \( da_i^s \) be the differential of the function \( a_i[x^s(\alpha), y^s(\alpha), \alpha] \), then along \( E_0 \) with \( \alpha = 0 \)
\[
\begin{align*}
\delta a_i &= - e_i, \\
da_i^s &= d_i(x^s) \quad (s = 1, 2),
\end{align*}
\]
where \( d_i(x), e_i(x) \) are defined as above. For, by differentiating the identity
for $\alpha_h$ at $(\alpha) = (0)$, it is found with the help of equations (4.2), (4.7) that along $E_0$

$$0 = y_{i\alpha h} \delta a_h + y_{i\alpha h} \delta a_h = \eta_{ih} \delta a_h + u_i,$$

and hence that $\delta a_i = -e_i$ by virtue of the definition of $e_i$. Similarly from equations (4.13) with $x, y_i$ replaced by $x'(\alpha), y_i'(\alpha)$ it follows that at $(\alpha) = (0)$

$$d y_i' - y_{i\alpha} dx' = y_{i\alpha} d a_i' + y_{i\alpha} d a_h = \eta_{ih} d a_i' + u_i,$$

and hence by equations (4.11) and $y_{i\alpha} = p_i$ along $E_0$ that

$$\eta_{ih}(x') d a_i' = c_i(x') - u_i(x'), \quad (s = 1, 2).$$

This is clearly only possible in case $d_i(x') = d a_i'$ by virtue of equations (4.7) and (4.8) with $w_h = da_h$. Finally by the use of the identity

$$F_{\nu'}(x, y, p, l) = z_i[x, a(x, y, \alpha), \alpha]$$

and the Euler-Lagrange equations $z_{i\alpha} = F_{\nu'}$ along the extremal $E_0$ one finds with the help of equations (4.2), (4.7), (4.12), and $w_h = da_h$ that

$$d F_{\nu'} - F_{\nu'} dx = z_{i\alpha} d a_h = \eta_{ih} d a_h + u_i, \quad (4.14)$$

$$\delta F_{\nu'} = z_{i\alpha} d a_h = -\xi_{ih} e_h + v_i.$$

The first two equations (4.10) are now an easy consequence of equations (4.11), (4.14), and (4.12). In order to establish the last of these equations we need the further relation along $E_0$

$$\delta p_i = y_{i\alpha h} \delta a_h + y_{i\alpha h} \delta a_h = -\eta_{ih} e_h + u_i = \eta_{ih} e_i'$$

which follows readily from equations (4.3), (4.2), and (4.12), the last equation being obtained by differentiating the equation $u_i = \eta_{ih} e_h$. Moreover along $E_0$ the expression $\eta_{ih} \delta F_{\nu'}$ is a constant. For, by means of equations (4.14) and $u_i = \eta_{ih} e_i$ it is found that along $E_0$

$$\eta_{ih} \delta F_{\nu'} = (\eta_{ih} v_i - u_{i\alpha} v_h) + (\eta_{ih} \xi_{ih} - \eta_{ih} u_i) e_i.$$

In this expression the second term on the right is zero by virtue of equations (2.6). The first term is a constant since (I, p. 738) the sets $\eta_{ih}, \xi_{ih}$ and $u_i, v_i$ are solutions of the accessory equations (2.5). The expression $\eta_{ih} \delta F_{\nu'}$ is accordingly constant along $E_0$ and hence by equations (4.15) we have

$$(d/dx)(u_i \delta F_{\nu'}) = (d/dx)(\delta F_{\nu'} \eta_{ih} e_h) = \delta F_{\nu'} \eta_{ih} e_i' = \delta p_i \delta F_{\nu'}.$$

From this relation and equation (4.14) one readily obtains the last equation (4.10). This completes the proof of Lemma 4.2.
5. Proof of Theorem 2.1. In this section we shall consider as admissible arc $E_0$ satisfying the conditions described in Theorem 2.1 with a set of multipliers $\lambda_0 = 1, \lambda_0(x)$. It is well known (I, p. 735) that $E_0$ forms with these multipliers a non-singular extremal arc. The following lemma will be useful in the proof of Theorem 2.1.

**Lemma 5.1.** There exists for the extremal $E_0$ a conjugate system $\eta_{ik}, \xi_{ik}$ and a set of $r$ solutions $u_{ih}, v_{ih} (h = 1, \cdots, r)$ of the accessory equations (2.5) such that the determinant $|\eta_{ik}(x)|$ is different from zero on $E_0$ and the quadratic form $Q(w)$ described in Lemma 4.2 is positive definite on $E_0$.

If we accept the truth of Lemma 5.1 the proof of Theorem 2.1 can be made as follows: By virtue of Lemmas 5.1, 4.1, and 4.2 there exist neighborhoods $\mathcal{N}$ of $E_0$ and $A$ of $(\alpha) = (0)$ and a set of slope functions and multipliers (4.3) defining a Mayer field over $\mathcal{N}$ for each set $(\alpha)$ in $A$ and such that $W(\alpha) > W(0)$ for every set $(\alpha) \neq (0)$ in $A$. The arc $E_0$ is an extremal of the field determined by $(\alpha) = (0)$. Moreover by virtue of the condition H of $E_0$ the condition (3.7) of Theorem 3.2 will be satisfied if we take the neighborhoods $\mathcal{N}$ and $A$ so small that the elements $(x, y, \rho, I)$ will lie in the neighborhood $\mathcal{N}$ of the elements $(x, y, y', \lambda)$ on $E_0$ prescribed by the condition II of $E_0$. Theorem 2.1 now follows from Theorem 3.2.

In order to establish Lemma 5.1 we set, for convenience, $x^1 = x^3, x^2 = x^5$ and prove first the further lemma.

**Lemma 5.2.** If $c_i(x)$ ($x^1 \leq x \leq x^5$) is an arc satisfying the condition

\[ \xi_i(x^2)c_i(x^2) - \xi_i(x^1)c_i(x^1) = 0 \quad (i = 1, \cdots, n) \]

for every solution $\eta_i, \xi_i$ of the accessory equations (2.5) having $\eta_i \equiv 0$ on $x^1x^5$, then the end values of $c_i(x)$ can be joined by a solution $\eta_i = u_i(x), \xi_i = v_i(x)$ of the accessory equations.

This result follows at once from the proof of "Lemma 7.2" of Hestenes (IV, p. 809), since the end points of $E_0$ are not conjugate; that is, there is no solution $\eta_i, \xi_i$ of the accessory equations having $\eta_i(x^1) = \eta_i(x^5) = 0$ and $\eta_i \neq 0$ on $x^1x^5$. This follows because along such an arc one would have $J_2(\eta, 0) = 0$ contrary to the condition IV'. This lemma differs from "Lemma 7.2" of Hestenes in that we do not require the arc $c_i(x)$ to be admissible.

Returning to the proof of Lemma 5.1 we first note that by virtue of the condition IV' there exists a conjugate system $\eta_{ik}, \xi_{ik}$ of solutions of the accessory equations having its determinant $|\eta_{ik}(x)|$ different from zero on $E_0$ (IV, p. 811). Secondly, let $U_{i\gamma}, V_{i\gamma} (\gamma = 1, \cdots, q)$ be a maximal set of solutions of equations (2.5) such that $U_{i\gamma} \equiv 0$ on $x^1x^5$ and the matrix
\[ \| V_{i\gamma}(x^t) c_{ih}^2 - V_{i\gamma}(x^t) c_{ih} \| (\gamma = 1, \ldots, q; h = 1, \ldots, r) \]

has rank \( q \). In fact we can suppose that

\[ V_{i\gamma}(x^t) c_{ih}^2 - V_{i\gamma}(x^t) c_{ih} = \delta_{\gamma h} \quad (\gamma = 1, \ldots, q; h = 1, \ldots, r), \]

where \( \delta_{\gamma h} = 1, \delta_{\gamma h} = 0 \) \((\gamma \neq h)\), since this result can be brought about by a non-singular linear transformation on the parameters \( \alpha_1, \ldots, \alpha_r \) and by taking suitable linear combinations of the solutions \( U_{i\gamma}, V_{i\gamma} \). Let \( c_{ih}(x) \) be a set of arcs having \( c_{ih}(x^t) = c_{ih}^t \) \((s=1, 2; h=1, \ldots, r)\). For each value \( \sigma (q < \sigma \leq r) \) the arc \( c_{i\sigma} \) satisfies the conditions described in Lemma 5.2 by virtue of equations (5.1). Its end values can therefore be joined by a solution \( u_{i\sigma}, v_{i\sigma} \) of the accessory equations. We may accordingly suppose that \( c_{i\sigma} = u_{i\sigma} \). Let

\[ u_{i\gamma} = \epsilon U_{i\gamma} = 0, \quad v_{i\gamma} = \epsilon V_{i\gamma} \quad (\gamma = 1, \ldots, q), \]

where \( \epsilon \) is a constant chosen in the manner described below, and let \( d_{ih}(x), e_{ih}(x) \) \((h=1, \ldots, r)\) be solutions of equations (4.8). Since \( U_{i\gamma} = 0 \)

\((\gamma=1, \ldots, q)\), \( c_{i\sigma} = u_{i\sigma} \) \((\sigma=q+1, \ldots, r)\), we have also \( e_{i\gamma} = 0, d_{i\sigma} = 0 \). If now we set

\[ A_{\beta\gamma} = b_{\beta\gamma} + (1/2) \left[ c_{i\sigma}^2 + c_{i\sigma} d_{i\gamma} \right]_{1}^2 \quad (\beta, \gamma = 1, \ldots, q), \]

\[ 2B_{\gamma r} = 2b_{\gamma r} + \left[ c_{i\sigma}^2 + c_{i\sigma} d_{i\gamma} - c_{i\gamma} (c_{i\sigma} d_{i\gamma} - 2v_{i\gamma}) \right], \]

\[ C_{\sigma r} = b_{\sigma r} + (1/2) \left[ u_{i\sigma} v_{i\sigma} + u_{i\sigma} v_{i\sigma} \right]_{1}^2 \quad (\sigma, \tau = q + 1, \ldots, r), \]

then by the use of equations (5.1), (5.2), \( e_{i\sigma} = 0, d_{i\sigma} = 0 \) it is found that the quadratic form \( Q(w) \) of Lemma 4.2 can be written in the form

\[ Q(w) = (A_{\beta\gamma} + 2\epsilon b_{\beta\gamma}) w_{\beta} w_{\gamma} + 2B_{\gamma r} w_{\gamma} w_{\tau} + C_{\sigma r} w_{\sigma} w_{\tau}, \]

where \( \delta_{\beta\gamma} = 1, \delta_{\beta\gamma} = 0 \) \((\beta = \gamma)\). The values \( A_{\beta\gamma} = A_{\gamma\beta}, B_{\gamma r}, C_{\sigma r} = C_{r\sigma} \) are independent of \( \epsilon \). If we set

\[ \eta_{i} = u_{i\sigma} w_{\sigma}, \quad z_{i} = v_{i\sigma} w_{\sigma}, \quad \bar{w}_{\gamma} = 0, \quad \bar{w}_{\sigma} = w_{\sigma} \quad (\gamma = 1, \ldots, q; \sigma = q + 1, \ldots, r), \]

we have \( \eta_{i}(x^t) = c_{ih}(x^t), \)

\((s=1, 2; h=1, \ldots, r)\) and

\[ C_{\sigma \tau} w_{\sigma} w_{\tau} = b_{\alpha i} \bar{w}_{\tau} \bar{w}_{\alpha} + [\eta_{i} z_{i}]_{1}^2 = J_{2}(\eta, \bar{w}), \]

the last equation being obtained by integration by parts with the help of equations (2.5). By the condition IV' we have \( J_{2}(\eta, \bar{w}) > 0 \) unless \( \eta_{i} = 0, \bar{w}_{\alpha} = 0 \), that is, unless \( w_{\sigma} = 0 \). The quadratic form (5.3) is accordingly positive definite. From the theory of quadratic forms it now follows that \( Q(w) \) will also be positive definite if the constant \( \epsilon \) is sufficiently large. For example, this can be seen by showing that the matrix
has a sequence of positive principal minors beginning in the lower right-hand corner. This proves Lemma 5.1. The proof of Theorem 2.1 is now complete.

6. Relation to the methods of Mayer, Hahn, and Bliss. The methods used in the last two sections are closely related to the methods of Mayer and Hahn as presented by Bliss (II) when properly interpreted. The earlier method seems to be applicable only in case $E_0$ is normal on the interval $x_1^0 x_2^0$. In making his sufficiency proof Bliss essentially establishes a result which when phrased in the terminology of the present paper can be stated as follows:

**Theorem 6.1.** If the arc $E_0$ is normal on $x_1^0 x_2^0$ and satisfies the conditions described in Theorem 2.1, there exist neighborhoods $\mathcal{E}$ of $E_0$ and $N$ of the end values of $E_0$ and a set of slope functions and multipliers (3.1) having the properties described in Theorem 3.1 and such that every point $(x^1, y^1, x^2, y^2)$ in $N$ is a set of end values of an extremal of the field determined by it. Moreover the function $W(x^1, y^1, x^2, y^2)$ is the extremal integral; that is, it represents the value of $J$ along the extremal in $\mathcal{E}$ whose end values are given by $(x^1, y^1, x^2, y^2)$.

Thus we see that in this case the inequality (3.4) can be written in the form

\[(6.1) \quad J(C) \geq J(E) \geq J(E_0),\]

where $E$ is the extremal in $\mathcal{E}$ joining the ends of $C$. The relation (6.1) and the last theorem indicate clearly the connection between our work and that of Mayer, Hahn, and Bliss.

**References**


VII. Bliss, *The problem of Bolza in the calculus of variations*, Lectures delivered at the University of Chicago during the Winter Quarter, 1935.


THE UNIVERSITY OF CALIFORNIA AT LOS ANGELES,
LOS ANGELES, CALIF.