

ADJOINT SYSTEMS IN THE PROBLEM OF MAYER UNDER GENERAL END-CONDITIONS*

BY S. B. MYERS

1. *Introduction.* Necessary conditions in the problems of Lagrange and Mayer with variable end points, analogous to the Euler equations and transversality condition in the ordinary problem of the calculus of variations in the plane, have recently been given by Morse and the author.† The methods used in the derivation of these conditions, as well as those used by Bliss‡ in his derivation of somewhat more complicated transversality conditions in the same problem, consist primarily of certain proofs which more properly belong to the theory of differential equations.

In the present paper, certain adjoint relationships in the problem hitherto unnoticed are pointed out which make it possible to derive the necessary conditions mentioned above merely by a consideration of the compatibility of the adjoint system of a certain set of differential equations and boundary conditions. In fact, in the abnormal case it is shown that these necessary conditions form an adjoint system of the differential and terminal equations of variation. The normal case is treated essentially by making it the abnormal case of a slightly different problem. Such considerations have the advantage of referring a major part of the proofs to the theory of differential equations, and thus greatly simplifying them.

2. *The Definition of the Adjoint System.* Consider the set of m linear homogeneous differential equations of the first order§ in the n variables η_1, \dots, η_n ,

$$(1) \quad L_\beta(\eta) \equiv p_{\beta i} \eta'_i + q_{\beta i} \eta_i = 0, \quad m \leq n, \\ (\beta = 1, \dots, m; i = 1, \dots, n),$$

* Presented to the Society, April 3, 1931.

† Morse and Myers, *Proceedings of the American Academy of Arts and Sciences*, vol. 66 (1931), pp. 235–253.

‡ Bliss, *Transactions of this Society*, vol. 19 (1918), pp. 305–314.

§ The usual convention of summation with respect to an index repeated in the same term is used throughout.

where the functions $p_{\beta i}(x)$, $q_{\beta i}(x)$ are continuous on the interval $x^1 \leq x \leq x^2$ and the matrix $\|p_{\beta i}(x)\|$ is of rank m throughout this interval.

By the *adjoint set of equations* we shall mean the following set of n equations in the m dependent variables $\lambda_1, \dots, \lambda_m$:

$$(2) \quad M_i(\lambda) \equiv -\frac{d}{dx}(p_{\beta i}\lambda_\beta) + q_{\beta i}\lambda_\beta = 0, \\ (\beta = 1, \dots, m; i = 1, \dots, n).$$

By a *continuous solution* of these quasi-differential equations* is meant a set (λ) of m continuous functions of x such that the functions $p_{\beta i}\lambda_\beta$ are of class C' and the equations are satisfied.

We note the following identity† which holds for arbitrary functions $(\eta) = (\eta_1, \dots, \eta_n)$ of class C' and arbitrary continuous functions $(\lambda) = (\lambda_1, \dots, \lambda_m)$ such that the functions $p_{\beta i}\lambda_\beta$ are of class C' :

$$(3) \quad \int_{x^1}^{x^2} [\lambda_\beta L_\beta(\eta) - \eta_i M_i(\lambda)] dx \equiv [p_{\beta i}\lambda_\beta \eta_i^s]_1^2.$$

Here η_i^s represents the value of $\eta_i(x)$ at $x = x^s$, ($s = 1, 2$), and λ_β^s the value of $\lambda_\beta(x)$ at $x = x^s$.

Consider any set of homogeneous linear boundary conditions on η_i^s . By a set of *adjoint boundary conditions* of these conditions shall be meant a set of linear homogeneous boundary conditions on λ_β^s with the following properties.

I. The bilinear form $[p_{\beta i}\lambda_\beta^s \eta_i^s]_1^2$ vanishes for all λ_β^s which are subject to these conditions while the η_i^s are subject to the original boundary conditions.

II. Of all sets of conditions with the above property (I), these conditions restrict λ_β^s as little as possible; that is, the points in $2m$ -space with coordinates λ_β^s fill out a subspace of maximum possible dimensionality.

The original boundary conditions may be written in the form

$$(4) \quad \eta_i^s = c_{ih}^s u_h, \quad (h = 1, \dots, r; s = 1, 2; i = 1, \dots, n),$$

where $(u) = (u_1, \dots, u_r)$ are parameters which are allowed to

* See Bôcher, *Leçons sur les Méthodes de Sturm*, p. 10.

† The notation $[\quad]_1^2$ means the value of $[\quad]$ for $s=2$ minus its value for $s=1$.

take on any values whatever, and $\|c_{ih}^s\|$ is a $(2nr)$ -dimensional matrix of constants.

THEOREM 1. *A necessary and sufficient condition that a set of homogeneous linear boundary conditions on λ_β^s in the form*

$$(5) \quad b_{\beta\delta}^s \lambda_\beta^s = 0, \quad (\delta = 1, \dots, p; p \text{ arbitrary}),$$

be a set of adjoint boundary conditions of (4) is that they be equivalent to the following conditions:

$$(6) \quad V_h \equiv [p_{\beta_i}^s \lambda_\beta^s c_{ih}^s]_1^2 = 0, \quad (h = 1, \dots, r).$$

The proof is simple, and will be left to the reader.

Equations (2) and conditions (6) together form an *adjoint system* of (1) and (4).

3. *Compatibility of the Adjoint System.* In applying our definition of an adjoint system to the problem of Mayer, the chief interest will lie in the question as to when an adjoint system has a solution. To answer this question, we start by adjoining to the differential equations (1) $n - m$ more differential equations so that the extended set of equations becomes

$$(7) \quad p_{ki} \eta_i' + q_{ki} \eta_i = \zeta_k(x), \\ (\zeta_\beta \equiv 0), \quad (\beta = 1, \dots, m; i, k = 1, \dots, n).$$

Here the functions $q_{\tau i}(x)$, $p_{\tau i}(x)$, ($\tau = m + 1, \dots, n$), are any continuous functions subject only to the restriction that the determinant $|p_{ki}(x)|$ be different from zero along (x^1, x^2) .*

If we have any solution (η) of (1) of class C' , we can use (7) to define continuous functions $\zeta_\tau(x)$ which we will say *correspond* to (η) . On the other hand, we can choose the functions ζ_τ as arbitrary continuous functions of x and obtain a solution (η) of (7) of class C' taking on arbitrary values at one end point. This set (η) will automatically be a solution of (1).

We now write down an extended adjoint set of equations

$$(8) \quad \frac{d}{dx}(p_{ki} \lambda_k) - q_{ki} \lambda_k = 0, \quad (i, k = 1, \dots, n).$$

We will use later the fact that these equations have a continuous

* For a proof of the possibility of this adjunction, see Bliss, loc. cit., p. 312.

solution $(\lambda_1, \dots, \lambda_n)$ such that the functions $p_{ki}\lambda_k$ take on arbitrary values at one end point (since $|p_{ki}|$ is not zero).

LEMMA. *Between functions $\eta_i(x)$ of class C' satisfying (7) with the corresponding ζ 's and continuous solutions λ_k of (8) we have the following relation:*

$$(9) \quad [p_{ki}\lambda_k\eta_i]_1^2 = \int_{x^1}^{x^2} \lambda_\tau \zeta_\tau dx.$$

For it is easily verified that the relation

$$\frac{d}{dx}(p_{ki}\lambda_k\eta_i) - \lambda_\tau \zeta_\tau = 0, \quad (i, k = 1, 2, \dots, n; \tau = m + 1, \dots, n),$$

holds, and we need only integrate this to obtain (9).

THEOREM 2. *A necessary and sufficient condition that the adjoint system (2), (6) have a continuous solution $(\lambda_1, \dots, \lambda_m)$ not zero at any point of (x^1, x^2) is that all determinants of the form*

$$(A) \quad \begin{vmatrix} \eta_{ij}^s - \bar{\eta}_{ij}^s \end{vmatrix}, \quad \begin{matrix} (2n \text{ columns : } j = 1, \dots, 2n), \\ (2n \text{ rows : } i = 1, \dots, n; s = 1, 2), \end{matrix}$$

vanish, where $\eta_{i1}^s, \dots, \eta_{i2n}^s$ are the end values of $2n$ sets of solutions of class C' of the original differential equations (1) and $\bar{\eta}_{i1}^s, \dots, \bar{\eta}_{i2n}^s$ are $2n$ sets of values which satisfy the boundary conditions (4) with $2n$ sets of values of the parameters (u) .

To prove the necessity of the condition, assume that there is a continuous solution $(\lambda_1, \dots, \lambda_m)$ of (2) and (6) not zero at any point of (x^1, x^2) . Then the set $(\lambda_1, \dots, \lambda_m, 0, \dots, 0)$ is a solution of (8) and the preceding lemma proves that

$$(10) \quad [p_{\beta r}\lambda_\beta\eta_i]_1^2 = 0.$$

If we multiply equations (6) respectively by arbitrary constants u_1, \dots, u_r and add, and subtract from (10), we obtain

$$(11) \quad [p_{\beta i}\lambda_\beta(\eta_i^s - c_{ih}^s u_h)]_1^2 = 0,$$

which may be written as

$$(12) \quad [p_{\beta i}\lambda_\beta(\eta_i^s - \bar{\eta}_i^s)]_1^2 = 0.$$

The $2n$ coefficients $p_{\beta_i^1} \lambda_{\beta_i^1}$, $p_{\beta_i^2} \lambda_{\beta_i^2}$ are not all zero, for otherwise we would have $\lambda_{\beta_i^s} = 0$, which is contrary to hypothesis. Hence all determinants of form (A) must vanish, and the condition of our theorem is necessary.

Now suppose that all determinants of form (A) vanish. Then there must exist constants m_i^s not all zero such that the equation

$$(13) \quad m_i^s (\eta_i^s - \bar{\eta}_i^s) = 0$$

holds for all solutions $\eta_i(x)$ of class C' of (1), and all solutions $\bar{\eta}_i^s$ of (4). This is proved as in Morse and Myers (loc. cit., p. 243). Using (4), we may write this in the form

$$(14) \quad m_i^s (\eta_i^s - c_{ih}^s u_h) = 0.$$

Let $(\lambda_1, \dots, \lambda_n)$ be any continuous solution of (8) not identically zero. Then $(\lambda_1, \dots, \lambda_n)$ cannot all be zero at any one point of (x^1, x^2) because of the homogeneity of (8). Subtract (9) from (14). We obtain

$$(15) \quad \eta_i^2 (m_i^2 - p_{ki}^2 \lambda_k^2) + \eta_i^1 (m_i^1 + p_{ki}^1 \lambda_k^1) - c_{ih}^s u_h m_i^s + \int_{x^1}^x \lambda_r \zeta_r dx = 0,$$

$$(i, k = 1, 2, \dots, n; h = 1, 2, \dots, r; \tau = m + 1, \dots, n).$$

According to the sentence following (8), we may consider $p_{ki}^1 \lambda_k^1$ as having arbitrary values. Furthermore, according to the remarks immediately preceding (8), we may consider ζ_r as arbitrary continuous functions of x , and η_i^2 as having arbitrary values. The values of the parameters u_h are also arbitrary.

Let us choose $p_{ki}^1 \lambda_k^1 = -m_i^1$. Then if we choose temporarily $\eta_i^2 = 0$ and $u_h = 0$, we find that $\lambda_r = 0$. Equation (15) may now be written in the form

$$(16) \quad \eta_i^2 (m_i^2 - p_{ki}^2 \lambda_k^2) - c_{ih}^s u_h m_i^s = 0.$$

From the arbitrariness of u_h and η_i^2 we deduce that

$$p_{ki}^2 \lambda_k^2 = m_i^2, \quad m_i^s c_{ih}^s = 0.$$

Thus the conditions

$$[p_{ki}^s \lambda_k^s c_{ih}^s]_1^2 = 0, \quad \frac{d}{dx} (p_{ki} \lambda_k) - q_{ki} \lambda_k = 0$$

are satisfied with $\lambda_\tau \equiv 0$, ($\tau = m+1, \dots, n$), or, in other words, conditions (6) and equations (2) have a continuous solution not zero. Hence the condition of our theorem is sufficient.

4. *The Problem of Mayer.* Let us consider arcs $y_i = y_i(x)$ of class C' in $(n+1)$ -space, let us denote the end points of these arcs by $(x^s, y_1^s, \dots, y_n^s)$ (x^s, y^s), where $s=1$ at the initial end point and $s=2$ at the final end point.

Such arcs which satisfy the differential equations*

$$(17) \quad \phi_\beta(x, y, y') = 0$$

are called *differentially admissible* arcs. The functions ϕ_β are of class C'' . Arcs of the same type which satisfy, for some value of the parameters $(\alpha) = (\alpha_1, \dots, \alpha_r)$, the conditions

$$(18) \quad x^s = x^s(\alpha), \quad y_i^s = y_i^s(\alpha),$$

will be called *terminally admissible* arcs. The end point functions $x^s(\alpha)$, $y_i^s(\alpha)$ are to be of class C' . An arc which is both differentially and terminally admissible will be called *admissible*. A set (α) will be called *admissible* if it determines through (18) the end points of some admissible arc.

Let E be an admissible arc $y_i = \bar{y}_i(x)$, $a^1 \leq x \leq a^2$, which satisfies (18) for $(\alpha) = (0)$ and along which the functional matrix $\|\phi_{\beta y_i'}\|$ is of rank m . We seek conditions under which E and the set $(\alpha) = (0)$ afford a minimum for α_1 among admissible sets (α) and the corresponding admissible arcs.

This simple form of the problem is new. It produces symmetrical results and proofs, and makes clear certain adjoint relationships in the problem. It is as general as any other form previously used in the sense that any of the other forms can be put into this form by simple transformations. The general form used by Morse and Myers (loc. cit.), in which the expression to be minimized is

$$\int_{x^1}^{x^2} f(x, y, y') dx + \theta(\alpha),$$

can be put into the form used in this paper by introducing a new dependent variable y_0 and a parameter α_0 , and adjoining the new differential equation $y_0' - f(x, y, y') = 0$, and the new end

* We use the notation (x, y, y') for $(x, y_1, \dots, y_n, y_1', \dots, y_n')$.

conditions $y_0^1 = 0$, $y_0^2 = \alpha_0 - \theta(\alpha)$. The expression to be minimized becomes simply α_0 .

5. *The Euler-Lagrange Equations and the Transversality Condition.* A set of functions $\eta_i(x)$ of class C' which satisfy the differential equations of variation

$$(19) \quad \phi_{\beta y_i} \eta_i' + \phi_{\beta u_i} \eta_i = 0, \quad (x, y, y') \text{ on } E,$$

will be called a set of *differentially admissible variations*.

We shall attach the subscript h to $x^s(\alpha)$ or $y_i^s(\alpha)$ to denote differentiation with respect to α_h . Let us set

$$c_{ih}^s = y_{ih}^s - \dot{y}_i^s x_h^s, \quad (h = 1, \dots, r).$$

Then a set of $2n$ numbers η_i^s which satisfy the *terminal equations of variation*

$$(20) \quad \eta_i^s = c_{ih}^s u_h$$

for some set of numbers $(u) = (u_1, \dots, u_r)$ will be called a set of *terminally admissible variations* and will be said to be determined by the set (u) .

Let $\|\eta_{ij}(x)\|$, $(i = 1, \dots, n; j = 1, \dots, 2n+2)$, be any matrix consisting of $2n+2$ columns of differentially admissible variations. Let γ_j^1, γ_j^2 be the j th of $2n+2$ sets of 2 arbitrary constants. Then the following theorem can be proved.*

THEOREM 3. *There exists a $(2n+2)$ -parameter family of differentially admissible arcs*

$$(21) \quad \begin{aligned} y_i &= y_i(x, e_1, \dots, e_{2n+2}), \\ x_0^1(e_1, \dots, e_{2n+2}) &\leq x \leq x_0^2(e_1, \dots, e_{2n+2}), \end{aligned}$$

containing the arc E for $(e_1, \dots, e_{2n+2}) = (0, \dots, 0)$, and such that

$$\begin{aligned} x_0^{e_j}(0, \dots, 0) &= \gamma_j^s, \quad y_{ie_j}(x, 0, \dots, 0) = \eta_{ij}(x), \\ &(j = 1, \dots, 2n+2). \end{aligned}$$

The functions y_i , y_{ix} , and x_0^s are of class C' in e_1, \dots, e_{2n+2} .

* See Morse and Myers, loc. cit., p. 239.

variations, and where $u_{p\delta}$ is arbitrary, $\tilde{\eta}_{i\delta}^s$ being given by the following:

$$(26) \quad \tilde{\eta}_{i\delta}^s = c_{ip}^s u_{p\delta}, \quad (p = 2, \dots, r).$$

But, according to Theorem 2 of §3, equation (25) is a sufficient condition that the following adjoint system of (20) and (26) have a continuous solution $(\lambda_1, \dots, \lambda_m)$ not zero at any point of (a^1, a^2) :

$$(27) \quad \frac{d}{dx}(\phi_{\beta y_i'} \lambda_\beta) - \phi_{\beta y_i} \lambda_\beta = 0, \quad (\beta = 1, \dots, m; i = 1, \dots, n),$$

$$(28) \quad [\phi_{\beta y_i'}^s \lambda_\beta^s c_{ip}^s]_1^2 = 0, \quad (p = 2, \dots, r).$$

But from (28) follows an identity in the differentials $d\alpha_p$;

$$[\phi_{\beta y_i'}^s \lambda_\beta^s c_{ip}^s]_1^2 d\alpha_p = 0, \quad (p = 2, \dots, r).$$

Now $c_{ih}^s d\alpha_h = dy_i^s - \bar{y}_i^s dx^s$, ($h = 1, 2, \dots, r$), and so we obtain

$$(29) \quad [\phi_{\beta y_i'}^s \lambda_\beta^s (dy_i^s - \bar{y}_i^s dx^s)]_1^2 + K d\alpha_1 = 0, \quad (K = \text{constant}).$$

These results may be summarized in the following theorem.

THEOREM 4. *If E and the set $(\alpha) = (0)$ afford a minimum in the problem, there exist a constant K and a set of continuous functions $\lambda_1(x), \dots, \lambda_m(x)$, not all zero at any point of (a^1, a^2) , such that if we set $F = \lambda_\beta \phi_\beta$, the equations*

$$(30) \quad \frac{d}{dx} F_{y_i'} - F_{y_i} = 0$$

are satisfied at every point of E , and the following identity in $(d\alpha)$ holds:

$$(31) \quad [F_{y_i'}^s (dy_i^s - \bar{y}_i^s dx^s)]_1^2 + K d\alpha_1 = 0,$$

when dx^s and dy_i^s are expressed in terms of $(d\alpha)$ and evaluated for $(\alpha) = (0)$.

6. *Normalcy.* An admissible arc E is said to be *normal relative to conditions* (30) and (31) if there exists no set of continuous functions $\lambda_\beta \not\equiv 0$ with which it satisfies these conditions with $K=0$; that is, if it does not satisfy the differential equations and boundary conditions

$$(32) \quad \frac{d}{dx}(\phi_{\beta y_i} \lambda_\beta) - \phi_{\beta y_i} \lambda_\beta = 0,$$

$$(33) \quad [\phi_{\beta y_i}^s \lambda_{\beta c_{ih}^s}]_1^2 = 0, \quad (h = 1, \dots, r).$$

Otherwise E is said to be *anormal* relative to these conditions.

From Theorems 1 and 2 the following theorems are evident.

THEOREM 5. *The system (32) and (33) is an adjoint system of the differential and terminal equations of variations (19) and (20).*

THEOREM 6. *A necessary and sufficient condition for normalcy relative to (30) and (31) is that there exist a non-vanishing determinant of the form*

$$|\eta_{i\delta}^s - \bar{\eta}_{i\delta}^s|, \quad (\delta = 1, \dots, 2n; i = 1, \dots, n; s = 1, 2),$$

where $\eta_{i\delta}(x)$, ($i = 1, \dots, n$), is the δ th of any $2n$ sets of differentially admissible variations, and $\bar{\eta}_{i\delta}^s$ is the δ th of any $2n$ sets of terminally admissible variations.

HARVARD UNIVERSITY