THE INDEX OF AN EXTREMAL ARC

WILLIAM KARUSH

1. Introduction. We are concerned with extremal arcs for the problem of minimizing a function

$$I(C) = g(a) + \int_{t_1}^{t_2} f(a, y, y')dt \tag{1.1}$$

over a class of parametric curves C in ay-space of the form

$$a_h, y_i(t) \qquad (h = 1, 2, \ldots, r;\ i = 1, 2, \ldots, n;\ t_1 \leqslant t \leqslant t_2) \tag{1.2}$$

and satisfying end conditions of the type

$$y_i(t_s) = y_{is}(a) \qquad (s = 1, 2). \tag{1.3}$$

The components a_h are constants and the functions g, y_{is}, and f are given, with the last function positively homogeneous of degree one.

We propose to consider several definitions for the index of an extremal of this problem, and to show that under appropriate hypotheses these indices are equivalent.

The first index defined is the so-called index of the second variation I_2 of I. A second index is then formulated in terms of families of curves (1.2); this index is of interest because its definition does not make use of the second variation or of topological considerations. A third index, the isoperimetric index, is treated next; it is similar to that of Birkhoff and Hestenes [**1**] (see also [**2**]). It is then shown that these three indices are equal in the non-degenerate case. In the next section it is shown that these indices are equal to the index of a critical point for a certain function of a finite number of real variables, which is defined in terms of broken extremals. In the final section we treat the topological index. The definition is not the most general one but it serves to indicate the relationship with the other indices. (For a more complete treatment of topological critical curves see [**3**], [**5**], and [**7**].) We show under the additional assumptions of positive definiteness and positive regularity for I that the topological index is equal to the earlier indices.

The precise analytic formulation of the extremum problem treated here is given in a preceding paper by the author [**4**]. To save space, frequent references will be made to the definitions, formulae, and theorems of that paper.

Throughout the paper we shall assume that E denotes a *non-singular extremal which does not intersect itself and which satisfies the end conditions* (1.3) *and the transversality condition* [**4**, (2.8)].

2. Index of the second variation. We fix our attention upon a particular extremal E; the results below hold relative to this arc.

Received August 29, 1950.

https://doi.org/10.4153/CJM-1952-025-0 Published online by Cambridge University Press

In [**4**] we dealt with (admissible) variations a_h, η_i (a_h constant); we shall use simply η to designate such variations (a, η). In particular $\eta = 0$ is the variation $(a, \eta) = (0, 0)$. Let $\mathfrak{H}$ denote the linear class of variations (real scalars) which satisfy, along E, the accessory end conditions

$$\eta_i(t_s) = y_{ish}a_h \qquad (s = 1, 2). \tag{2.1}$$

(Here the first subscript "h" denotes differentiation with respect to a_h; also, a repeated index indicates summation over that index.)

The second variation of I along E is given by [**4**, (2.10)]; it is now designated by $I_2(\eta)$ instead of $I_2\ (a, \eta)$. Let $\mathfrak{H}'$ be a maximal linear subset of $\mathfrak{H}$ of finite dimension (possibly zero) on which I_2 is negative definite, i.e., for which $I_2(\eta) < 0$ for $\eta \neq 0$ in $\mathfrak{H}'$. The dimension k of $\mathfrak{H}'$ will be called the *index of* I_2 on $\mathfrak{H}$. If no such $\mathfrak{H}'$ of finite dimension exists, then we shall say that the index is infinite.

Associated with the second variation is the bilinear form

$$I_2(\eta, \bar{\eta}) = b_{hk}a_h\bar{a}_k + \int_{t_1}^{t_2} (\omega_{a_h}\bar{a}_k + \omega_{\eta_i}\bar{\eta}_i + \omega_{\eta'_i}\bar{\eta}'_i)dt, \tag{2.2}$$

where the arguments in the derivatives of ω are those belonging [**4**, (2.11)] to η. The function (2.2) is the first variation of $I_2(\eta)$. The following properties are obvious: (2.2) is linear in each of its arguments, it is symmetric in its arguments, and it reduces to the second variation for $\eta = \bar{\eta}$. We use I_2 to designate the second variation of I or (2.2); the context in every case will make the equation unambiguous.

A variation η in $\mathfrak{H}$ is *orthogonal* to a subset of $\mathfrak{H}$ in case $I_2(\eta, \bar{\eta}) = 0$ whenever the second argument belongs to the subset. A tangential variation (i.e., one of the form 0, $w(t)y_i(t)$ where y_i belongs to E) which belongs to $\mathfrak{H}$ is orthogonal to every variation; this follows from the homogeneity of f. One subset is orthogonal to another in case each variation of the first subset is orthogonal to each variation of the second.

The proofs below are phrased for finite-dimensional, linear subspaces. Where the complete proof requires extension to the infinite case, such extension will be obvious.

LEMMA 2.1. *The index k of I_2 on $\mathfrak{H}$ is uniquely defined.*

Let $\mathfrak{H}', \mathfrak{H}''$ be maximal linear subsets of $\mathfrak{H}$ of finite dimension on which I_2 is negative definite with bases $\eta_1, \ldots, \eta_p$ and $\bar{\eta}_1, \ldots, \bar{\eta}_q$ respectively. (Subscripts here denote different vector functions, not separate components as in (2.2)). Suppose $p > q$. Then there exist constants c_i, not all zero, such that

$$I_2(\bar{\eta}_j, \eta_i)c_i = I_2(\bar{\eta}_j, \eta) = 0,$$

where $\eta = c_i\eta_i$. Hence η is orthogonal to $\mathfrak{H}''$. For any variation $\bar{\eta} + b\eta$ with $b =$ constant, $\bar{\eta}$ in $\mathfrak{H}''$, and not both b and $\bar{\eta}$ zero,

$$I_2(\bar{\eta} + b\eta) = I_2(\bar{\eta}) + 2bI_2(\bar{\eta}, \eta) + b^2I_2(\eta) < 0.$$

It follows that $\mathfrak{H}''$ is not maximal. Hence $p \leqslant q$. Similarly, $q \leqslant p$. Therefore $p = q$, and the proof is complete.

Suppose $\mathfrak{H}$ contains a maximal linear subset $\mathfrak{H}_0$ of finite dimension which is orthogonal to $\mathfrak{H}$ and contains no non-zero tangential variation. Then the dimension d of $\mathfrak{H}_0$ will be called the *order of degeneracy* of I_2 on $\mathfrak{H}$. If no such subset exists the order of degeneracy will be said to be infinite. In case $d = 0$ the second variation I_2 will be called *non-degenerate on* $\mathfrak{H}$.

LEMMA 2.2. *The order of degeneracy d is uniquely defined.*

Let $\mathfrak{H}_0$, $\mathfrak{H}_1$, be maximal linear subsets of finite dimension which contain no non-zero tangential variation and are orthogonal to $\mathfrak{H}$. Suppose the dimension p of the first subset exceeds the dimension q of the second. Then at least one variation $\eta \neq 0$ of $\mathfrak{H}_0$ is not expressible as an element of $\mathfrak{H}_1$ plus a tangential variation (possibly 0). Then the subset with dimension $q + 1$ spanned by $\mathfrak{H}_1$ and η contains no non-tangential variation and is orthogonal to $\mathfrak{H}$, contrary to $\mathfrak{H}_1$ being maximal. Hence $p \leqslant q$. Similarly $q \leqslant p$, and thus $p = q$.

LEMMA 2.3 [cf. **1**; **2**]. *If E satisfies the Clebsch condition the integers k and d are finite.*

LEMMA 2.4. *Let $\mathfrak{H}'$ be a maximal linear subset of $\mathfrak{H}$ on which I_2 is negative definite. Let $\mathfrak{D}$ be the set of elements in $\mathfrak{H}$ which are orthogonal to $\mathfrak{H}'$. Then every element of $\mathfrak{H}$ can be written uniquely as the sum of an element of $\mathfrak{H}'$ and an element of $\mathfrak{D}$. Furthermore, $I_2(\eta) \geqslant 0$ for every η in $\mathfrak{D}$, the equality holding only in case η is orthogonal to the whole space $\mathfrak{H}$.*

Let $\bar{\eta}_1, \ldots, \bar{\eta}_k$ be a basis for $\mathfrak{H}'$ and η be an arbitrary element of $\mathfrak{H}$. Since the determinant $|I_2(\bar{\eta}_i, \bar{\eta}_j)|$ does not vanish we can select (c) such that

$$I_2(\eta, \bar{\eta}_j) = c_i I_2(\bar{\eta}_i, \bar{\eta}_j).$$

Hence $\eta - c_i\bar{\eta}_i$ is in $\mathfrak{D}$, and we have the decomposition $\eta = c_i\bar{\eta}_i + (\eta - c_i\bar{\eta}_i)$ To prove the uniqueness assume for the moment that the second conclusion of the theorem holds. Let

$$\eta = \eta_1 + \eta_2 = \eta^*_1 + \eta^*_2$$

with η_1, η^*_1 in $\mathfrak{H}'$ and η_2, η^*_2 in $\mathfrak{D}$. Then $\eta^*_2 - \eta_2 = \eta_1 - \eta^*_1$. Since the left member is in $\mathfrak{D}$ and the right member is in $\mathfrak{H}'$, we have

$$0 \leqslant I_2(\eta^*_2 - \eta_2) = I_2(\eta_1 - \eta^*_1) \leqslant 0.$$

Whence $\eta_1 = \eta^*_1$, $\eta_2 = \eta^*_2$ as desired. To establish the second half of the theorem suppose η were an element of $\mathfrak{D}$ with $I_2(\eta) < 0$. For b = constant and $\bar{\eta}$ in $\mathfrak{H}'$ we would have

$$I_2(\bar{\eta} + b\eta) = I_2(\bar{\eta}) + b^2 I_2(\eta) < 0,$$

unless $b = 0$, $\bar{\eta} = 0$, contrary to $\mathfrak{H}'$ being maximal. Finally, let $I_2(\eta) = 0$. Suppose there were an η^* in $\mathfrak{D}$ not orthogonal to η. Then

$$I_2(\eta^* + b\eta) = I_2(\eta^*) + 2bI_2(\eta^*, \eta) < 0,$$

for a suitable b. As just seen, this is a contradiction. Thus η is orthogonal to $\mathfrak{D}$, and since it is also in this subset it is orthogonal to $\mathfrak{H}'$. From the first part of the proof it follows that η is orthogonal to $\mathfrak{H}$, as desired.

LEMMA 2.5. *Let $\mathfrak{H}_1$ be a linear subset, containing no non-zero tangential variation, on which I_2 is non-positive. Then* dim $\mathfrak{H}_1 \leqslant k + d$. *The equality holds in case $\mathfrak{H}_1$ is maximal with these properties.*

Suppose dim $\mathfrak{H}_1 > k + d$. Select a subset $\mathfrak{H}^*_1$ of $\mathfrak{H}_1$ with finite dimension q greater than $k + d$, and let $\mathfrak{D}_1$ be the elements of $\mathfrak{H}^*_1$ which are orthogonal to $\mathfrak{H}$. Then $\mathfrak{H}^*_1$ has a basis in which the first d_1 elements form a basis for $\mathfrak{D}_1$ and the remaining elements span a subset which we shall denote by $\mathfrak{H}_2$. Since dim $\mathfrak{D}_1 \leqslant d$, we have dim $\mathfrak{H}_2 > k$. Following the device used in the proof of Lemma 2.1 we can choose an element $\eta \neq 0$ in $\mathfrak{H}_2$ which is orthogonal to $\mathfrak{H}'$ of Lemma 2.4. From that lemma, $I_2(\eta) \geqslant 0$. But $I_2(\eta) \leqslant 0$ since η is in $\mathfrak{H}_1$. Hence $I_2(\eta) = 0$, η is orthogonal to $\mathfrak{H}$ by the same lemma, and η belongs to $\mathfrak{D}_1$. This implies that $\mathfrak{H}_2$ and $\mathfrak{D}_1$ have a non-zero element in common, which contradicts the definition of $\mathfrak{H}_2$. Inasmuch as we shall not make use of the second part of the lemma we shall not pause to prove it.

3. Index of an extremal. In this section we propose a definition of index without resort to the second variation or to topological methods.

Let $\mathfrak{C}$ be the class of admissible arcs a_h, $y_i(t)$ which satisfy the end conditions (1.3). Consider a q-parameter family of arcs

$$a_h = a_h(b_1, \ldots, b_q), \quad y_i = y_i(t, b_1, \ldots, b_q) \tag{3.1}$$

$$(t_1 \leqslant t \leqslant t_2, \quad h = 1, 2, \ldots, r; \quad i = 1, 2, \ldots, n)$$

in $\mathfrak{C}$ which contains E for $(b) = (0)$. For (t, b), near the values belonging to E, the functions in (3.1) are supposed to have continuous first and second derivatives relative to the components of (b), and these derivatives, in turn, are to have piecewise continuous derivatives relative to t. The family is said to be of dimension q on E if the variations

$$\eta_p\colon \quad a_{hb_p}(0), \quad y_{ib_p}(t, 0) \qquad (p = 1, 2, \ldots, q) \tag{3.2}$$

span a linear space of dimension q which contains no non-zero tangential variation. Along (3.1) the integral I defines a function $I(b)$. By the *index m of E on $\mathfrak{C}$* will be meant the least upper bound of the integers q for which there exists a family (3.1) of dimension q on E such that the corresponding function $I(b)$ has a proper relative maximum at $(b) = (0)$.

THEOREM 3.1. *The index m of E on $\mathfrak{C}$ satisfies the inequality $k \leqslant m \leqslant k + d$ where k is the index of I_2 on $\mathfrak{H}$ and d in the order of degeneracy of I_2 on $\mathfrak{H}$. If I_2 is non-degenerate on $\mathfrak{H}$ then $m = k$.*

Let $\eta_1, \ldots, \eta_k$ be the basis of a linear subset $\mathfrak{H}'$ as in Lemma 2.4. Let $a_{h0}, y_{i0}(t)$ define E and let

$$a_h(b) = a_{h0} + a_{hp}b_p, \quad Y_i(t, b) = y_{i0}(t) + \eta_{ip}(t)b_p \qquad (p = 1, 2, \ldots, k).$$

Then the k-parameter family

$$a_h(b), \quad y_i(t, b) = Y_i(t, b) + h_{i1}(b)(t_2 - t) + h_{i2}(b)(t - t_1)$$

where

$$h_{is}(b) = \frac{y_{is}(a(b)) - Y_i(t_s, b)}{t_2 - t_1} \qquad (s = 1, 2),$$

lies in $\mathfrak{C}$ and has η_p as its variations (3.2). Hence the family has dimension k. The function $I(b)$ has at $(b) = (0)$ the differentials

$$dI = I_1(\delta y), \quad d^2I = I_2(\delta y)$$

where δy is $\eta_p db_p$. Since E is an extremal satisfying the transversality condition, $dI = 0$. Furthermore $d^2I < 0$ for all $(db) \neq (0)$. Thus $I(b)$ has a proper relative maximum at $(b) = (0)$, and $k \leqslant m$. To prove $m \leqslant k + d$ let (3.1) define a family of dimension q for which $I(b)$ has a proper relative maximum at $(b) = (0)$. It follows that

$$dI(0) = I_1(\delta y) = 0, \quad d^2I(0) = I_2(\delta y) \leqslant 0,$$

with δy as above and variations (3.2). The linear subset spanned by (3.2) satisfies the hypotheses of Lemma 2.5. Hence $q \leqslant k + d$, $m \leqslant k + d$, as desired.

4. Isoperimetric index. Consider a set of q isoperimetric conditions

$$I_p(C) = g_p(a) + \int_{t_1}^{t_2} f_p(a, y, y')dt = 0 \qquad (p = 1, 2, \ldots, q) \tag{4.1}$$

for which the arc E is normal, i.e., there is no linear combination $c_pI_p(c) \neq (0)$ for which E is an extremal satisfying the transversality condition. By the *isoperimetric index* k' of E will be meant the least of the integers q for which there exist conditions (4.1) such that E affords I a weak relative minimum in the class of admissible arcs C satisfying (1.3) and (4.1). If no such finite (possibly 0) integer q exists the isoperimetric index will be said to be infinite. The idea of defining index by the adjunction of isoperimetric conditions was used by Birkhoff and Hestenes in [**1**] where they employed "natural" isoperimetric conditions.

LEMMA 4.1. *For any variation $\bar\eta$: $\bar a_h$, $\bar\eta_i(t)$ in $\mathfrak{H}$ of class C''' there exists a set of functions $\xi_i(a, y)$ defined and of class C''' in a neighbourhood of the values (a, y) belonging to E such that along E we have $\xi_i(a_0, y_0(t)) = \bar\eta_i(t)$ and such that the first variation of the function*

$$I_1(\xi; C) = g_h\bar a_h + \int_C (f_{a_h}\bar a_h + f_{y_i}\xi_i + f_{y'_i}\xi_{iy_j}y'_j)dt \tag{4.2}$$

has the form $I_2(\bar\eta, \eta)$ along E for η in $\mathfrak{H}$.

https://doi.org/10.4153/CJM-1952-025-0 Published online by Cambridge University Press

We show first the existence of functions ξ_i for which the first equation in the conclusion holds and such that

$$\xi_i(a, y_s(a)) = y_{ish}(a)\bar{a}_h. \tag{4.3}$$

Assume $(\bar{a}) \neq 0$; say the first component does not vanish. Select $r - 1$ variations η_k: $\alpha_{hk}, \eta_{ik}(t)$ $(k = 2, 3, \ldots, r)$ in $\mathfrak{H}$ of class C''' for which

$$\alpha_{hk} = \delta_{hk} \tag{4.4}$$

where the symbol on the right is the Kronecker delta. Since the variations satisfy (2.1),

$$\eta_{ik}(t_s) = y_{isk}(a_0). \tag{4.5}$$

Now if we define η_{i1} by

$$\bar{a}_1\eta_{i1} + \bar{a}_2\eta_{i2} + \ldots + \bar{a}_r\eta_{ir} = \bar{\eta}(t) \tag{4.6}$$

and select α_{h1} to satisfy (4.4), then equations (4.4) and (4.5) hold for $k = 1, 2, \ldots, r$.

Next, define the family $a_h, y_i(t, a)$ as in [4, (3.4)]. Then

$$y_i(t_s, a) = y_{is}(a), \quad y_{ia_h}(t, a_0) = \eta_{ih}(t). \tag{4.7}$$

Choose functions $u_{ij}(t)$ $(j = 1, 2, \ldots, n - 1)$ of class C''' such that the determinant $|y'_{i0}\, u_{ij}|$ does not vanish on t_1t_2. Then the equations $y_i = y_i(t, a) + u_{ij}(t)e_j$ have initial solutions (a, y, t, e) for values (a, y, t) belonging to E and $(e) = (0)$, the functional determinant with respect to (t, e) does not vanish on t_1t_2, and no two distinct points (a, y, t, e) have the same projection (a, y) since E does not intersect itself. Hence there exist unique solutions $t = t(a, y), e_j = e_j(a, y)$ of these equations defined and of class C''' in a neighbourhood of the values (a, y) belonging to E. Furthermore, the solutions satisfy

$$t(a_0, y_0(t)) = t, \quad t(a, y_s(a)) = t_s \qquad (s = 1, 2). \tag{4.8}$$

Define

$$\eta_i(a, t) = y_{ia_h}(t, a)\bar{a}_h.$$

Then the desired functions are $\xi_i(a, y) = \eta_i(a, t(a, y))$. For, from equations (4.6), (4.7), and (4.8),

$$\xi_i(a_0, y_0(t)) = \eta_i(a_0, t) = y_{ia_h}(t, a_0)\bar{a}_h = \eta_{ih}(t)\bar{a}_h = \bar{\eta}_i(t),$$

$$\xi_i(a, y_s(a)) = \eta_i(a, t_s) = y_{ia_h}(t_s, a)\bar{a}_h = y_{ish}(a)\bar{a}_h.$$

Now suppose $(\bar{a}) = (0)$. In this case we select r arbitrary variations satisfying (4.4) and (4.5) and determine $t(a, y)$ as above. Then the desired functions are $\xi_i(a, y) = \bar{\eta}_i(t(a, y))$.

It remains to calculate the first variation of (4.2) along E. We find it to be

$$g_{hk}\bar{a}_h\alpha_k + \int_{t_1}^{t_2}\left(f_{y_i}\delta\xi_i + f_{y'_i}\frac{d}{dt}\delta\xi_i\right)dt + \int_{t_1}^{t_2}(\omega_{a_h}\bar{a}_h + \omega_{\eta_i}\bar{\eta}_i + \omega_{\eta'_i}\bar{\eta}'_i)dt,$$

where

https://doi.org/10.4153/CJM-1952-025-0 Published online by Cambridge University Press

$$\delta\xi_i = \xi_{ia_h}a_h + \xi_{iy_j}\eta_j,$$

evaluated at E. Using (2.1) and the result of differentiating (4.3) we find the second term above to equal

$$\int_{t_1}^{t_2} \frac{d}{dt}(f_{y'_i}\delta\xi_i)dt = [f_{y'_i}\delta\xi_i]_{t_1}^{t_2} = [f_{y'_i}(\xi_{ia_h}a_h + \xi_{iy_j}y_{isk}a_k)]_{t_1}^{t_2} = [f_{y'_i}y_{ishk}\bar{a}_h a_k]_{t_1}^{t_2}.$$

The theorem follows from (2.2) and [**4**, (2.11)].

THEOREM 4.1. *Let E be a non-singular extremal for the function I which does not intersect itself and satisfies the end conditions* (1.3) *and the transversality condition. Suppose that E satisfies the Clebsch condition and that the second variation I_2 of I along E is non-degenerate. Then the isoperimetric index k', the index m of E, and the index k of the second variation of I along E are all finite and equal.*

Lemma 2.3 establishes the finiteness of k, and Theorem 3.1 the equality $m = k$. We shall show $k = k'$. Select a maximal linear subspace $\mathfrak{H}'$ as in Lemma 2.4. By appropriate modification of a basis of $\mathfrak{H}'$ we obtain another subspace $\mathfrak{H}'$ for which each element of its basis is of class C'''. Consider such a subspace. For each η_q construct the function ξ_q of Lemma 4.1. Consider the isoperimetric conditions

$$(4.9) \qquad I_1(\xi_q; C) = 0 \qquad (q = 1, 2, \ldots, k)$$

determined by (4.2). We now make use of the sufficiency theorem [**4**, Theorem 10.1]. We note first that E satisfies (4.9), since along E the left side of (4.9) is the first variation of I evaluated at η_q. The arc E with multipliers $l_q = 0$ will satisfy the conditions of the above-mentioned sufficiency theorem for the problem of minimizing I relative to the end conditions (1.3) and the isoperimetric conditions (4.9) if we show the following: that the second variation I_2 of I along E is positive for non-tangential variations in $\mathfrak{H}$ which make the first variation of each function on the left in (4.9) vanish. By Lemma 4.1, the condition that the first variation vanish is $I_2(\eta_q, \eta) = 0$, that is, orthogonality to $\mathfrak{H}'$. By Lemma 2.4, for such an orthogonal variation $I_2(\eta) \geqslant 0$, the equality holding just in case η is orthogonal to $\mathfrak{H}$. From the non-degeneracy of I_2 the only non-zero variations orthogonal to $\mathfrak{H}$ are the tangential variations. From this the desired condition on I_2 is verified and we may deduce that $I(E)$ is a weak relative minimum relative to (1.3) and (4.9). Also, E is normal relative to (4.9); for otherwise E would be an extremal satisfying the transversality condition for a function which is a linear combination of the left sides of (4.9) with coefficients $(c) \neq (0)$. Thus the first variation $c_q I_2(\eta_q, \eta)$ of this function would vanish for all η in $\mathfrak{H}$, contrary to the assumption of non-degeneracy. We may conclude that $k' \leqslant k$.

To show the reverse inequality, consider a set of isoperimetric conditions $I_p(C) = 0$ $(p = 1, 2, \ldots, k')$ for which E is normal and such that $I(E)$ is a weak relative minimum relative to these conditions and (1.3). Then a necessary condition on E is $I_2(\eta) \geqslant 0$ for η in $\mathfrak{H}$ satisfying the first variation conditions

$I_{p1}(\eta) = 0$. Suppose $k' < k$. Let $\eta_1, \ldots, \eta_k$ be a basis for $\mathfrak{H}'$ as in Lemma 2.4. There exist coefficients $(c) \neq (0)$ such that $c_q I_{p1}(\eta_q) = I_{p1}(\eta) = 0$, where $\eta = c_q \eta_q \neq 0$. Thus $I_2(\eta) \geqslant 0$, contrary to η being in $\mathfrak{H}'$. Thus $k \leqslant k'$, $k = k'$, and the proof is complete.

5. Additional results. Let $f(z)$ be a function of a finite number of variables $(z_1, \ldots, z_m)$ which is of class C'' in the neighbourhood of a point $(z) = (z_0)$. We shall say that $(z) = (z_0)$ is an *analytical critical point of index k* in case at (z_0) we have (i),

$$f_{z_j} = 0 \qquad (j = 1, \ldots, m),$$

and (ii), the quadratic form

$$f_{z_i z_j} dz_i dz_j$$

has negative index k. An analytical critical point will be called *non-degenerate* in case the Hessian

$$|f_{z_i z_j}|$$

does not vanish there. It is well known that the index of a non-degenerate critical point (z_0) is equal to the dimension of a maximal linear subspace of (dz)-space on which the quadratic form

$$f_{z_i z_j} dz_i dz_j$$

is negative definite.

We wish to make use of certain results in the proof of [**4**, Theorem 8.1]. These results depend upon the assumptions on E made in the last paragraph of §1 and, in addition, the hypothesis that E satisfies the Weierstrass condition II_N. (Not all the hypotheses of [**4**, Theorem 8.1] are required for the results we shall use.) Instead of this Weierstrass condition we shall assume the Clebsch condition on E. This will only change the extremum properties in [**4**] from those of strong relative minima to those of weak relative minima. We may see this as follows. Let $\mathfrak{R}$ be the set of admissible points (a, y, y') for which the original problem of §1 is defined. From non-singularity and the Clebsch condition we deduce the existence of a neighbourhood $\mathfrak{R}_0$ of the values of (a, y, y') belonging to E for which E satisfies II_N (see first statement in the proof of [**4**, Theorem 10.1]). Then the results we shall carry over from [**4**] are valid when $\mathfrak{R}$ is replaced by $\mathfrak{R}_0$, i.e., when in the original problem we replace strong minima by weak minima.

We now make use of the family of broken extremals

$$a_h, \; y_i = y_i(t, a, e) \tag{5.1}$$

given by [**4**, (8.9)]. (The variables (l) do not appear since there are no isoperimetric conditions.) This family contains the given arc E for values (a_0, e_0), and the extremal segments of these curves have the (weak) minimizing property described in [**4**, Theorem 7.2]. The variations

$$\bar{\eta}: \quad a_h = da_h, \quad \eta_i = y_{ia_k} da_h + y_{ie_r} de_r \tag{5.2}$$

https://doi.org/10.4153/CJM-1952-025-0 Published online by Cambridge University Press

of (5.1) form a linear space of broken special accessory extremals satisfying (2.1) and containing no tangential variation for $(da, de) \neq (0, 0)$. This space, call it $\mathfrak{H}_1$, is a subspace of $\mathfrak{H}$. (For the properties of (5.1) and (5.2) see [**4**, Theorem 7.2] and the proof of [**4**, Theorem 8.1].)

LEMMA 5.1. *Let k be the index of I_2. Then there exists a linear subset $\mathfrak{H}'_1$ of $\mathfrak{H}_1$ which is of dimension k and on which I_2 is negative definite.*

We shall show that for an arbitrary admissible variation η: a_h, η_i in $\mathfrak{H}$ there exists a variation $\bar{\eta}$ in $\mathfrak{H}_1$ with the same components (a) such that $I_2(\bar{\eta}) \leqslant I_2(\eta)$. It is clear from [**4**, (8.13)] that we can determine constants (da, de) and a function $w(t)$ of class C'' which vanishes at the end points such that

$$\eta_i(s_j) = \bar{\eta}_i(s_j) + w(s_j)y'_{i0}(s_j) \qquad (j = 1, 2, \ldots, q),$$

where $\bar{\eta}$ is given by (5.2) with $(da) = (a)$. Let η^* represent the tangential variation $(0, wy'_0)$ and $\eta_0 = \eta - (\bar{\eta} + \eta^*)$. Then η_0 vanishes at the corner points s_j. Furthermore, since each segment of E between corner points affords I a weak relative minimum relative to admissible arcs joining its end points we must have $I_2(\eta_0) \geqslant 0$. Now

$$I_2(\eta_0) = I_2(\eta) - 2I_2(\eta, \bar{\eta} + \eta^*) + I_2(\bar{\eta} + \eta^*).$$

Using, in particular, the property that η^* and $\bar{\eta}$ satisfy the accessory Euler equations, we find that

$$I_2(\eta, \bar{\eta} + \eta^*) = I_2(\bar{\eta} + \eta^*) + I_2(\eta_0, \bar{\eta} + \eta^*) = I_2(\bar{\eta} + \eta^*),$$

$$I_2(\bar{\eta} + \eta^*) = I_2(\bar{\eta}) + 2I_2(\bar{\eta}, \eta^*) + I_2(\eta^*) = I_2(\bar{\eta}).$$

Hence $I_2(\eta_0) = I_2(\eta) - I_2(\bar{\eta})$, $I_2(\bar{\eta}) \leqslant I_2(\eta)$. Now let $\eta_1, \ldots, \eta_k$ be a basis of a linear subset $\mathfrak{H}'$ as in Lemma 2.4. For each η_p deternine the variation $\bar{\eta}_p$ as above. If $\eta = c_p\eta_p$ is an arbitrary element of $\mathfrak{H}'$ the associated variation $\bar{\eta}$ determined above is $c_p\bar{\eta}_p$. Hence the linear subset spanned by $\bar{\eta}_p$ will satisfy the conclusion of the lemma.

THEOREM 5.1. *Let E satisfy the hypotheses of Theorem* 4.1. *Then the function $I(a, e)$ obtained by evaluating I along* (5.1) *has a non-degenerate analytical critical point at $(a, e) = (a_0, e_0)$ of index k if and only if the extremal E has index k.*

By Theorem 4.1 the index of E equals the index of I_2. For the function $I(a, e)$ we have, at (a_0, e_0), $dI = I_1(\bar{\eta})$, $d^2I = I_2(\bar{\eta})$ where $\bar{\eta}$ is given by (5.2). Suppose that (a_0, e_0) is a non-degenerate analytical critical point of index k_0. To a maximal linear subspace of (da, de)-space on which d^2I is negative definite there corresponds, through (5.2), a linear subset of $\mathfrak{H}$ of dimension k_0 on which I_2 is negative definite. Hence $k_0 \leqslant k$. Also, to the subset $\mathfrak{H}'_1$ of Lemma 5.1 there corresponds a linear subset in (da, de)-space of dimension k on which d^2I is negative definite. Hence $k \leqslant k_0$, $k = k_0$. Conversely, let E be an extremal of index k. Since the first variation of I vanishes at $\bar{\eta}$, the function $I(a, e)$ has an analytical critical point at (a_0, e_0). To prove non-degeneracy suppose that the

https://doi.org/10.4153/CJM-1952-025-0 Published online by Cambridge University Press

Hessian of $I(a, e)$ vanished there. There would exist constants $(da, de) \neq (0, 0)$ such that the corresponding non-tangential variation $\bar{\eta}$ of $\mathfrak{H}_1$ would be orthogonal to $\mathfrak{H}_1$. Since the latter set contains $\mathfrak{H}'_1$ of Lemma 5.1, by Lemma 2.4, $I_2(\bar{\eta}) \geqslant 0$. From the non-degeneracy of I_2 the equality sign could not hold. This is contrary to the fact that $\bar{\eta}$ is orthogonal to itself. Finally, the proof of the equality $k_0 = k$ is a repetition of the first part of the proof.

THEOREM 5.2. *Let the domain $\mathfrak{R}$ of elements (a, y, y') for which the extremum problem of* §1 *is defined consist of elements with (a, y) in an open set in ay-space and $y'_i y'_i \neq 0$. In Theorem* 4.1 *replace the Clebsch condition by the Weierstrass condition* II. *The modified theorem is valid even if in the definition of isoperimetric index it is required that $I(E)$ be a strong, rather than weak, relative minimum.*

(See [**4**, §10] for the definition of condition II.) The proof is identical to that of Theorem 4.1 except for the use of [**4**, Theorem 10.2] instead of [**4**, Theorem 10.1].

We conclude the section with a method, based on Lemma 4.1, for constructing abnormal isoperimetric problems.

THEOREM 5.3. *Let η_p $(p = 1, 2, \ldots, q)$ be admissible variations of class C''' which satisfy* (2.1). *Suppose there exist constants c_p not all zero such that $\eta = c_p \eta_p$ is an accessory extremal satisfying the transversality condition for the second variation. Then the extremal E is abnormal relative to the isoperimetric conditions*

$$I_1(\xi_p; C) = 0 \qquad (p = 1, 2, \ldots, q), \tag{5.3}$$

where the functions appearing in (5.3) *are defined in Lemma* 4.1.

The condition of abnormality in this case is the existence of $(c) \neq 0$ such that

$$c_p\left(\omega_{\eta_{ip}} - \frac{d}{dt}\omega_{\eta'_{ip}}\right) = 0,$$

$$c_p\left(b_{hk}a_{hp} + [\omega_{\eta'_{ip}} y_{ish}]_1^2 + \int_{t_1}^{t_2} \omega_{a_{kp}} dt\right) = 0.$$

(See [**4**, (8.1)]. Here $\omega_{\eta_{ip}}$ means ω_{η_i} evaluated at the variation η_p, etc.) The first condition is that η be an accessory extremal, and the second that this variation satisfy the transversality condition for I_2. The conclusion now follows [**4**, §8].

COROLLARY. *If in Theorem* 5.3 *we do not require η to satisfy the accessory transversality condition then the conclusion holds with "abnormal" replaced by "not strongly normal."*

6. Topological critical curves. In this section we shall assume a knowledge of modular Vietoris cycles on a metric space.

Let S be a metric space whose elements we denote by C. Let I be a real single-valued function on S and $\mathfrak{K}$ be the class of subsets K of S determined by $I \leqslant c$ for c real or $+\infty$. In the sequel K, K', etc., will denote members of $\mathfrak{K}$. Suppose that all the sets K for $c \neq \infty$ are compact. Consider a particular point C_0 of S

https://doi.org/10.4153/CJM-1952-025-0 Published online by Cambridge University Press

and let K denote the fixed set determined by $c = I(C_0)$. Then we shall say that C_0 is an *isolated topological critical point of index k and count* 1 in case:

(1) There exists a Vietoris k-cycle u on K mod K' such that $u \not\sim 0$ on K mod $(K - C_0)$, and if N is a neighbourhood of C_0 then there is a set $K'' \subset K$ such that $u \sim 0$ on K mod $(K'' + KN)$.

(2) If V is a k-cycle on K mod K_0 with property (1), then there exists a set $K'_0 \subset K$ such that $u \sim v$ on K mod K'_0. (Here $\subset$ denotes strict inclusion.)

Clearly C_0 is a critical point of index k and count 1 relative to the space S if and only if it is a similar critical point relative to the space K. The above definition of a critical point is not the most general one but it will serve our purpose of relating the topological index to the indices defined earlier. The definition here given was suggested by [**3**].

A homotopy Δ of a subset P of S into a subset Q of S is called an *I-deformation* in case I never increases under Δ, that is, if $h(C, t)$ (C on P, t on $0 \leqslant t \leqslant 1$) is the function defining Δ and $C' = h(C, t_1)$, then $I(C'') \leqslant I(C')$ for every $C'' = h(C, t)$ with $t_1 \leqslant t \leqslant 1$. For C in P, let $\Delta C = h(C, 1)$. Let ΔP be the set of points ΔC with C in P.

THEOREM 6.1. *Let C_0 be an isolated topological critical point of index k and count* 1. *Let N be a neighbourhood of C_0 and Δ be an I-deformation of N. Then C is a fixed point under Δ.*

Suppose that C_0 is not a fixed point. Then there exists a neighbourhood $N_0 \subseteq N$ of C_0 and an I-deformation Δ_1 such that $\Delta_1(KN_0) \subseteq (K - C_0)$. Let u be a modular k-cycle related to C_0 as in the definition of a critical point. From the second part of (1) of this definition we can assume that u is on KN_0. It follows that under Δ_1 the k-cycle u is deformed into a k-cycle on $(K - C_0)$. Hence $u \sim 0$ on K mod $(K - C_0)$, contrary to our choice of u. This proves the theorem.

We return now to the function $I(C)$ given by (1.1). In this section we shall make the following additional assumptions on I: (1) The region of admissible points (a, y, y') has the form (a, y) in a region $\mathfrak{R}$ of ay-space and $(y') \neq (0)$ arbitrary. (2) *Positive definiteness*, that is, the integrand function f is positive everywhere. (3) *Positive regularity*, that is, for each admissible element (a, y, y'),

$$f_{y'_i y'_j} \sigma_i \sigma_j > 0, \qquad (\sigma) \neq (ky').$$

We fix attention on a particular admissible curve E in $\mathfrak{R}$ satisfying the conditions stated at the end of §1. (Notice that positive regularity of I automatically ensures non-singularity and the Clebsch condition for E.) Let $\mathfrak{R}'$ be a neighbourhood of E such that the closure $\bar{\mathfrak{R}}'$ is contained in $\mathfrak{R}$. Let R be the class of all rectifiable curves in $\bar{\mathfrak{R}}'$ which satisfy (1.3). We introduce the usual Frechet metric in R. By standard theory, our hypotheses imply that the subsets of R determined by inequalities ($I(C) \leqslant c$, $c \neq \infty$), are compact. We now make the following definition. The arc E is an *isolated topological critical curve of index k* *and count* 1 in case there exists a closed Fréchet neighbourhood $S \subset R$ of E such that E is a critical point according to the earlier definition in the space S. Such a curve will be called simply a "critical curve." Let K be the compact subset of R determined by $I(C) \leqslant I(E)$. From the remark following the definition of critical point, we see that we can limit ourselves to the space K. Henceforth we shall do so and consider only subsets and neighbourhoods relative to K.

It is well known that under our hypotheses on the function I there exists a number $c > 0$ such that any two points in $\bar{\mathfrak{R}}'$ within e-distance of each other can be joined in either direction by a unique extremal in $\mathfrak{R}$ which affords I a proper minimum relative to arcs in $\bar{\mathfrak{R}}'$ joining its end points. These short minimizing extremals vary continuously with their end points, and I is continuous on this class of arcs. Pick a neighbourhood $\mathfrak{R}_0$ of E such that the short minimizing arcs with end points in $\mathfrak{R}_0$ all lie in $\bar{\mathfrak{R}}'$.

Let

$$a_h(C), \quad y_i(t, C) \qquad (0 \leqslant t \leqslant 1, \quad C \text{ in } K) \tag{6.1}$$

be the special parameterization of curves given by Morse [**6**]. For each C in K the functions (6.1) define a parameterization of C, and the functions (6.1) are continuous in t and C simultaneously. From the compactness of K, there exists a number $d > 0$ such that if $|t_1 - t_2| \leqslant d$ and C is in K then the points given by (6.1) for $t = t_1$, and $t = t_2$ are within e-distance of each other. Let S be a closed Frechet neighbourhood of E such that all the curves of S lie in $\mathfrak{R}_0$. Let t_1t_2 be a sub-interval of $0 \leqslant t \leqslant 1$ of length less than d. We shall describe an I-deformation $\Delta(t_1, t_2)$ of S. Let C be any curve of S. When $t = 0$, leave C unaltered, when $0 < \tau \leqslant 1$ replace the sub-arc of C between t_1 and $t_1 + \tau(t_2 - t_1)$ by the short minimizing extremal joining its end points and directed in the same sense as C. By performing this deformation simultaneously on all the curves of S we obtain $\Delta(t_1, t_2)$.

THEOREM 6.2. *A critical curve is an extremal.*

For, by the deformation $\Delta(t_1, t_2)$ and Theorem 6.1 every sufficiently short sub-arc of E is an extremal. Therefore E itself is an extremal.

We wish now to show the equivalence of the topological index with the earlier indices. For this purpose we make the following construction. Subdivide the interval $0 \leqslant t \leqslant 1$ by points $t_0 = 0, t_1, \ldots, t_q, t_{q+1} = 1$ so that (1) the length of each sub-interval is less than d, (2) the arc E has on it no pairs of conjugate points between t_j and t_{j+1}, and (3) the length of arc on E between t_j and t_{j+1} is less than $e/6$. Let

$$(a_0, y_1(a_0)), \quad (a_0, b_{10}), \ldots, (a_0, b_{q0}), \quad (a_0, y_2(a_0))$$

be the points of subdivision on E. For nearby values (a, b) we can obtain a broken extremal with end and corner points given by the preceding sequence with the subscript zero deleted. In this way we obtain a family

$$a_h, \quad y_i(t, a, b) \tag{6.2}$$

https://doi.org/10.4153/CJM-1952-025-0 Published online by Cambridge University Press

of broken extremals with the continuity properties of the family [**4**, (8.7)]. Let $\mathfrak{B}$ be a neighbourhood of (a_0, b_0) such that for (a, b) in $\mathfrak{B}$ the arc (6.2) lies in $\mathfrak{R}$, has the length of each sub-arc between corners less than $e/3$, and intersects π_j once and only once. (Here π_0 and π_{q+1} are the end manifolds of (1.3) and $\pi_1, \ldots, \pi_q$ are $(n-1)$-dimensional hyperplanes through $(a_0, b_{10}), \ldots, (a_0, b_{q0})$ orthogonal to E.)

Apply the above subdivision of the interval $0 \leqslant t \leqslant 1$ to each of the curves (6.1), and let S be a closed Fréchet neighbourhood of E such that each curve of S is in $\mathfrak{R}_0$ and has its values (a, b) corresponding to the points of subdivision of the interval in the neighbourhood $\mathfrak{B}$. Let Δ be the I-deformation obtained by applying

$$\Delta\ (t_j, t_{j+1}) \qquad (j = 0, 1, \ldots, q)$$

simultaneously for all j. The homotopy Δ deforms S into the family (6.2). Our next step is to I-deform (6.2) into its sub-family having its corners on the hyperplanes π_j. To this end let

$$a_h = a_h, \;\; y_i = b_i(e_{1,j}, e_{2,j}, \ldots, e_{n-1,j}) \qquad (j = 1, 2, \ldots, q)$$

be the equations of the hyperplanes π_j, where $b_i(e_{j0}) = b_{ij0}$. Then the sub-family is

$$(6.3) \qquad a_h, \;\; y_i(t, a, e) \equiv y_i[t, a, b(e)].$$

Let C be an arc of the family (6.2) and denote its corner points by $P_1, P_2, \ldots, P_q$. Let $Q_1, Q_2, \ldots, Q_q$ be the points of intersection of C with $\pi_1, \pi_2, \ldots, \pi_q$.For any time τ on the interval $0 \leqslant \tau \leqslant 1$ let R_j $(j = 1, 2, \ldots, q)$ be the point of the sub-arc of C between P_j and Q_j such that the ratio of the distance P_jR_j (j not summed) along C to the distance P_jQ_j along C equals τ. Let R_0, R_{q+1} be the fixed end points of C. For all $j = 0, 1, \ldots, q$, replace the sub-arc of C between R_j and R_{j+1} by the short minimizing arc joining these points and directed in the same sense as C. As τ varies this construction when applied to all the curves (6.2) yields an I-deformation Δ' into (6.3). The product $\Delta'\Delta$ is an I-deformation of S into (6.3). From the invariance of homology relations under homotopy and from the fact that $\Delta'\Delta$ is an I-deformation we infer the following result:

THEOREM 6.3. *Let E be an extremal which does not intersect itself and satisfies the end conditions* (1.3). *Then E is an isolated topological critical curve of index k and count* 1 *if and only if* (a_0, e_0) *is an isolated topological critical point of index k and count* 1 *of the function $I(a, e)$, where $I(a, e)$ is the value of the integral I along the family* (6.3).

Now it is known that (a_0, e_0) is an isolated topological critical point of index k and count 1 for $I(a, e)$ if and only if the point is a non-degenerate analytical critical point of $I(a, e)$ of index k. The results of §5 for the family (5.1) are valid for the family (6.3). Thus, from Theorems 6.3 and 5.1 we obtain the following result:

https://doi.org/10.4153/CJM-1952-025-0 Published online by Cambridge University Press

THEOREM 6.4. *Under the assumptions of this section and Theorem* 5.1 *an extremal E is an isolated topological critical curve of index k and count* 1 *if and only if it is an extremal of index k.*

REFERENCES

1. G. D. Birkhoff and M. R. Hestenes, *Natural isoperimetric conditions in the calculus of variations*, Duke Math. J., vol. 1 (1935), 198-286.
2. K. E. Hazard, *Index theorems for the problem of Bolza*, Contributions to the Calculus of Variations, 1938-41 (Chicago).
3. M. R. Hestenes, *A theory of critical points*, Amer. J. Math., vol. 67 (1945), 521-562.
4. W. Karush, *Isoperimetric problems in the calculus of variations*, Can. J. Math., vol. 4 (1952), 257-280.
5. M. Morse, *The calculus of variations in the large* (American Mathematical Society, Colloquium Publications, vol. 18, 1934).
6. ———, *A special parameterization of curves*, Bull. Amer. Math. Soc., vol. 42 (1936), 915-922.
7. ———, *Functional topology and abstract variational theory*, Ann. Math., vol. 38 (1937), 386-449.

The University of Chicago