FUNCTIONAL TOPOLOGY

MARSTON MORSE

Introduction. The present paper gives the fundamental existence proof for a homotopic critical point under hypotheses which are less stringent than those employed in the author's recent fascicule on functional topology and abstract variational theory [2]. This relaxation of hypotheses seems to be necessary if one is to extend the variational theory in the large to non-regular problems. Non-regular integrals include some of the most important and interesting integrals such as the Jacobi least action integral[1] in the three body problem of celestial mechanics. The author believes that a topological basis for the planetary orbits will be disclosed by studying the contour manifolds of this Jacobi integral.

The results of this paper will be used to extend the theory in the large to non-regular problems in two papers by Morse and Ewing.

1. **The metric spaces** M, L, J. We are concerned with a compact metric space M of elements $p, q, r, \cdots$ and distance $pq, pr, \cdots$. We shall deal with two functions $J(p)$ and $L(p)$, bounded, single-valued, and lower semi-continuous on M. In the applications p will be a curve joining two fixed points in some space, the distance pq will be the Fréchet distance between curves, while $J(p)$ and $L(p)$ will be integrals along p, with L the length of p.

Beside the metric M we shall use two other metrics, an L-metric with a distance

$$|pq| = pq + |L(p) - L(q)|,$$

and a J-metric with a distance,

$$pq + |J(p) - J(q)|.$$

We shall refer to the corresponding spaces as the spaces L and J. We shall make the following hypothesis.

HYPOTHESIS. *Convergence on L to a point p shall imply convergence on J to p.*

We shall not assume that convergence on J to p implies convergence on L to p. Convergence on L or J clearly implies convergence on M. Terms such as neighborhood, compact, and so on

---

Received by the editors May 5, 1942.

[1] For other examples see paper [1] by McShane.

will be preceded by the letters M, L, or J according to which metric is used to define them. Every M-neighborhood of a point p contains an L-neighborhood of p, but not conversely. Every J-neighborhood of p contains an L-neighborhood of p, but not conversely.

The subset of points on which $J \leqq c$ will be denoted by J^c. The set J^c is M-compact since J is lower semi-continuous on M. But J^c will not in general be L or J-compact. In fact an M-compact subset A would be L-compact if and only if $L(p)$ were M-continuous on A. Similarly with J-compactness. In seeking to cover sets such as J^c with a finite number of neighborhoods we accordingly use M-neighborhoods, although it would be simpler if we could use L-neighborhoods.

On the other hand the deformations which we shall use in later papers are adequate only if their continuity is L-continuity. We shall thus be using L-continuous deformations defined over M-neighborhoods. But J-neighborhoods enter also, since there are important properties which can be established for J-neighborhoods but not for L or M-neighborhoods.

Our chains and cycles shall be defined on L using L-continuity. They shall be finite singular chains and cycles, taken mod 2, see [**3**, p. 146]. The point set bearing a singular chain is L-compact and hence J-compact. We shall term the least upper bound of J on an arbitrary set E, the *J-height* of E. In general we shall say that a point p is *above* or *below* q according as $J(p)$ is greater or less than $J(q)$. We shall admit *relative* k-cycles u in which the modulus is always a set J^c in which c is less than the J-height of u.

2. **J-deformations.** Let E be a subset of L. Let I be an interval $0 \leqq \tau \leqq a$ $(a>0)$. By $E \times I$ we shall mean the product of E and I, assigning the usual metric to the space $E \times I$. We shall admit deformations D of E which replace a point p found on E at the time $\tau=0$ by a point $q(p, \tau)$ on L at the time τ $(0 \leqq \tau \leqq a)$. If D is to be admissible we require:

(a) *That $q(p, \tau)$ map $E \times I$ continuously into L.*

(b) *That for p fixed $q(p, \tau)$ map I continuously into M uniformly with respect to (p, τ) on $E \times I$.*

We shall say that D is a *weak J-deformation*, if for p on E and for $q=q(p, \tau)$, $J(p)-J(q) \geqq 0$ for each τ on I. A weak J-formation $q(p, \tau)$ will be said to be *proper* on E, if when τ is bounded from 0 on I,

$$J(p) - J(q) > e > 0, \qquad q = q(p, \tau),$$

where e is a constant independent of p on E.

A point p will be said to be *homotopically J-ordinary* if some J-neighborhood of p admits a proper J-deformation. A point which is not homotopically J-ordinary will be termed *homotopically J-critical.*

We shall say that J is *upper-reducible* at p if for each constant $a > J(p)$ there exists a weak J-deformation of some M-neighborhood of p which is proper for points initially above a.

The reader may wonder why a J-neighborhood is used in the above definition of a J-critical point while an M-neighborhood is used in the definition of upper-reducibility. A J-neighborhood is used in this connection because otherwise we are unable to prove that a rectifiable curve which is not an extremal is homotopically J-ordinary.[2] This reflects the fact that being an extremal is a consequence of properties possessed by weak neighborhoods such as J-neighborhoods. On the other hand an M-neighborhood is used in the definition of upper-reducibility since otherwise we could not prove the fundamental Theorem 3.1 of this paper. For, as one sees from its definition, upper-reducibility is not concerned merely with points q at which $J(q)$ is near $J(p)$, so that a J-neighborhood would be too restrictive.

We come to products of deformations. Let D be a deformation of a set A. The set of final images of points of A under D will be denoted by $D(A)$. Let $B_1, \cdots, B_n$ be a set of weak J-deformations such that B_1 is applicable to A, B_2 to $B_1(A)$, and more generally B_{i+1} is applicable to $B_i \cdots B_1(A)$. In such circumstances the deformations $B_n \cdots B_1$ will be said to define a product deformation Δ. The deformation Δ is applicable to A. One sees that Δ is a weak J-deformation of A.

The following lemma shows how a weak J-deformation can be extended as a weak J-deformation beyond a local domain of definition.

LEMMA 2.1. *Let B be a subset of L and s a point of B. Let S_r be the intersection with B of a spherical M-neighborhood of s of radius r. Suppose that S_{3e} admits a weak J-deformation D. Then D can be replaced by a weak J-deformation θ of B such that $\theta \equiv D$ for points initially on S_e and θ is the null deformation for points initially exterior to S_{2e}.*

Suppose the time τ in D varies on the interval $(0, a)$. Under θ, τ shall likewise vary on $(0, a)$. Points of B initially on S_e shall be deformed under θ as under D while points of B initially exterior to S_{2e} shall be held fast. Points p of B at a distance ps from s such that

$$e \leqq ps \leqq 2e, \tag{2.1}$$

[2] This theorem is the basic generalization of the Euler theorem that a curve (of class C') which is not an extremal is not minimizing.

shall be deformed under θ as follows. Let $t(p)$ divide the interval $(0, a)$ in the ratio inverse to that in which ps divides the interval (2.1). That is, let

$$\frac{t(p)}{a} = \frac{2e - ps}{e}, \qquad e \leqq ps \leqq 2e.$$

Under θ points p of B which satisfy (2.1) initially shall be deformed as under D until τ reaches $t(p)$ and shall be fast thereafter.

Let $q(p, \tau)$ be the image of p under θ. Recall that ps, and hence $t(p)$, vary continuously as p varies L-continuously on $\overline{S_{2e} - S_e}$. It will be convenient to set $t(p) \equiv a$ for p on S_e and $t(p) \equiv 0$ for p on $B - S_{2e}$. Then $t(p)$ is defined and continuous as p varies on B. To establish the continuity of $q(p, \tau)$ one breaks the domain of the pairs (p, τ) into the two domains

$$(2.3) \qquad [0 \leqq \tau \leqq t(p)] \qquad [p \text{ on } B],$$

$$(2.4) \qquad [t(p) \leqq \tau \leqq a] \qquad [p \text{ on } B],$$

with the set on which $\tau = t(p)$ in common. On the second domain $q(p, \tau)$ is constant. On the first domain $q(p, \tau)$ equals the point function defining D. The functions $q(p, \tau)$ defined over these separate domains obviously combine to define a function $q(p, \tau)$ with the properties of a weak J-deformation of B.

3. **The fundamental theorem.** Two different sets of neighborhoods U and V, respectively, covering a given space, will be said to be *equivalent* if each neighborhood U of each point p contains a neighborhood V of p, and conversely. In this sense the set of J-neighborhoods is equivalent to the set of neighborhoods $V(p)$ defined by conditions of the form

$$(3.1) \qquad pq < \delta, \qquad J(q) < J(p) + e,$$

where p is fixed and δ and e are arbitrary positive constants. This is a consequence of the lower semi-continuity of $J(p)$ for p on M. For the relation $pq < \delta$ implies the relation

$$(3.2) \qquad J(p) - e < J(q)$$

if δ is sufficiently small. Conditions (3.1) and (3.2) taken jointly define a set of neighborhoods equivalent to the set of J-neighborhoods, so that the set of neighborhoods (3.1) is equivalent to the set of J-neighborhoods. We shall therefore feel free to replace J-neighborhoods by neighborhoods of the type (3.1).

We continue with a fundamental theorem.

THEOREM 3.1. *Let C be an M-compact set with J-height c, containing no homotopic J-critical points at the level c. If J is upper-reducible at each point of C there exists a weak J-deformation of a J-neighborhood of C into a set with J-height less than c.*

I. If p is a point of C at the level c, p is homotopically J-ordinary. There accordingly exists a spherical M-neighborhood $V(p)$ of p and a positive constant $c(p)>c$ such that the intersection $J^{c(p)}\cdot V(p)$ admits a proper J-deformation $D(p)$.

II. If p is a point of C below c the upper-reducibility of J at p implies the following: there exists a spherical M-neighborhood $V(p)$ of p and a constant $a(p)<c$ such that $V(p)$ admits a weak J-deformation $D(p)$ which is a proper deformation of the subset of $V(p)$ above $a(p)$.

Let $U(p)$ and $R(p)$ be spherical M-neighborhoods of p with radii one-third and one-sixth that of $V(p)$, respectively. Since C is M-compact there exists a finite set of neighborhoods $R(p)$, say neighborhoods $R_1, \cdots, R_n$ of points $p_1, \cdots, p_n$, respectively, covering C. For points p_i at the level c, set $\mu=\min c(p_i)$. For points p_i below c set $\nu=\max a(p_i)$. We shall apply Lemma 2.1 setting $s=p_i$, $B=J^\mu$ and e is equal to the radius of $U_i=U(p_i)$. With this choice of s, B and e we infer the existence of a weak J-deformation H_i of J^μ which deforms $U_i\cdot J^\mu$ as does $D(p_i)$.

Let r equal the minimum of the radii r_i of the spheres R_i. Let θ_i be a weak J-deformation defined by taking the time interval in H_i so short that the maximum M-displacement under θ_i is less than r/n. The product deformation $\Delta=\theta_n\cdots\theta_1$ is a weak J-deformation of J^μ and displaces no point an M-distance in excess of $r\leqq r_i$. In particular we can conclude that

$$\theta_{i-1}\cdots\theta_1(R_i\cdot J^\mu)\subset U_i\cdot J^\mu. \tag{3.3}$$

Let S_i be the set defined by the left number of (3.3) and let C_i be the subset of S_i above ν. The deformation θ_i is proper on C_i since $\theta_i\equiv D(p_i)$ on C_i. If $C_i\neq 0$, the J-height of C_i is accordingly diminished under θ_i by a positive constant η_i. If $C_i=0$, set $\eta_i=0$. Let 2η be a positive constant less than each $\eta_i\neq 0$ and such that $c+\eta<\mu$ (recall that $c<\mu$). Then Δ deforms the set

$$(R_1+\cdots+R_n)\cdot J^{c+\eta} \tag{3.4}$$

into a set with a J-height at most $\max(\nu, c-\eta)<c$.

The set (3.4) contains a J-neighborhood of C and the proof of the theorem is complete.

A class K of k-cycles on L, either relative or absolute, will be said to form a *J-class* if any cycle into which a cycle u of K can be carried by a weak J-deformation is also in K. With this understood we have the following corollary of the theorem.

COROLLARY 3.1. *Suppose J is upper-reducible at each point of an M-compact set C. If there exists a J-class K of k-cycles such that K includes a cycle in each J-neighborhood of C but no cycle with a J-height less than that of C there exists at least one homotopic J-critical point with the J-height of C.*

In applying this corollary one is led to the two following specializations.

COROLLARY 3.2. *Let K be a homology class of absolute k-cycles, non-bounding on L, and let c be the greatest lower bound of J-heights of cycles of K. If J is upper-reducible at each point of J^c there exists a homotopic J-critical point of J at the level c.*

Corollary 3.2 follows from Corollary 3.1 upon taking C as J^c.

COROLLARY 3.3. *Let K be a homology class composed of k-cycles* mod J^a, *non-bounding* mod J^a *on L, and let c be the greatest lower bound of J-heights of cycles of K. If $c>a$ and if J is upper-reducible at each point of J^c there exists a homotopic J-critical point at the level c.*

REFERENCES

1. E. J. McShane, *Semi-continuity of integrals in the calculus of variations*, Duke Math. J. vol. 2 (1936) pp. 597–616.

2. Marston Morse, *Functional topology and abstract variational theory*, Mémorial des Sciences Mathématiques, no. 92, 1939, Gauthier-Villars, Paris.

3. ———, *The calculus of variations in the large*, Amer. Math. Soc. Colloquium Publications, vol. 18, 1934.

THE INSTITUTE FOR ADVANCED STUDY