THE CRITICAL POINTS OF A FUNCTION OF n VARIABLES*

BY

MARSTON MORSE

1. Introduction. This paper contains among other results a treatment of the critical† points of a real analytic function without restriction as to the nature of the critical points. Together with the results stated by the author elsewhere‡ on the removal of the boundary conditions it constitutes a complete treatment of the problem, the first of its kind.

In most of the paper the function considered is of class C'', and may have critical loci not even complexes, in fact an infinite set of such loci. Moreover even in the analytic case it is not assumed that the critical loci are complexes, a considerable advantage in any case, and the more so because no adequate proof exists that they are complexes.

Starting with a topological definition of type numbers in terms of ordinary neighborhoods of the critical sets, it ends with a most precise determination of these type numbers *in terms of regions bounded by closed analytic manifolds without singularity.* At no point is it necessary to break up regions more complicated than these into complexes.

All of the results on critical points known to the author, with one exception,§ follow as special cases. The results on isolated critical points obtained by Brown|| in his Harvard Thesis are the simplest of corollaries. The author's¶ previous results on non-degenerate critical points are obtained with more difficulty. It is shown that the definitions of type numbers given are justified by a kind of invariance under slight analytic deformations of the function.

The treatment will carry over to regular n-spreads in $(n+r)$-space. In it deformations predominate. It is essentially a generalization of the methods

* Presented to the Society, October 25, 1930; received by the editors October 7, 1930.

† A critical point of a function is a point at which all of the first partial derivatives of the function vanish. The value of the function at such a point is called a critical value.

‡ Morse II, Proceedings of the National Academy of Sciences, vol. 13 (1927), p. 813.

§ Whyburn, W. M., Bulletin of the American Mathematical Society, vol. 35 (1929), p. 701. Here the critical values are not necessarily finite in number.

|| Brown, American Journal of Mathematics, vol. 52 (1930), p. 251. See also Annals of Mathematics, vol. 31 (1930), p. 449.

¶ Morse I, these Transactions, vol. 27 (1925), p. 345.

found necessary in a treatment of n-dimensional critical loci in the calculus of variations* in the large where a reduction to complexes was not possible.

2. **The functions.** Let $(x)=(x_1, \cdots, x_n)$ be a point in euclidean n-space in a finite region Σ, bounded by a closed point set M consisting of a finite number of connected, regular,† non-intersecting $(n-1)$-spreads of class C'''.

Let $f(x)=f(x_1, \cdots, x_n)$ be a real function of class C'' defined on a region including Σ in its interior. On M we suppose that the directional derivative of f in the sense of the exterior normal to M is positive. As in Morse I§20 we can then alter the definition of f neighboring M so that the resulting function, which we will again call f, will take on an absolute maximum on M relative to its values on Σ. This can be done without introducing any new critical points.

We *assume* that the critical values of f are *finite* in number. This hypothesis is always fulfilled if f is analytic.

If a and b are any two ordinary (not critical) values of f, with no critical values between them, the domains $f \leqq a$ and $f \leqq b$ are homeomorphic (Morse I §7). When there are critical values between a and b this will not in general be so. We are concerned in what follows with the topological differences between the domains $f \leqq a$ and $f \leqq b$, and the manner in which these differences depend on the critical points of f.

We shall begin by supposing that there is just one critical value of f between a and b, and shall denote the domains $f \leqq a$ and $f \leqq b$ by A and B respectively, $a < b$.

3. **The neighborhoods of critical sets** g. By a *critical set* g will be understood any closed set of critical points on which f is constant, which is at a positive distance from other critical points. It may or may not be connected. In general it will not be a complex.

By a neighborhood N of g will be meant an open set of points which includes all points within a small positive distance of g. We admit only neighborhoods which lie on $B-A$ and are at a positive distance from other critical points of f. A neighborhood N^1 will be called smaller than N if it is on N and the distance between the boundaries of N and N^1 is positive. We always suppose N^1 smaller than N and in particular so small that any point on N^1 can be connected to g on N. This is always possible.

---

* Morse III, these Transactions, vol. 36 (1930), p. 599. See also Birkhoff, these Transactions, vol. 18 (1917), p. 240. Poincaré, Liouville's Journal, (4), vol. 1 (1885), pp. 167–244. Kronecker, Werke, vol. I, pp. 175–226, and vol. II, pp. 71–82.

† An $(n-1)$-spread is called regular and of class C''', if in the neighborhood of any one of its points it can be represented by giving one of its coördinates as a function of class C''' of the remaining coördinates.

Suppose $f=0$ on g. By $\overline{N}$ and $\overline{N}^1$ we shall mean those points of N and N^1 at which $f<0$.

We assume for the present that g is of such sort that for a proper choice of N and corresponding sufficiently small choice of N^1 the following sets of cycles exist.

(α) *A complete set* $(a)_k$ *of k-cycles on* $\overline{N}^1$, *independent on* $\overline{N}$, *dependent on* N^1.

(β) *A complete set* $(c)_k$ *of k-cycles on* N^1 *independent on* N *of the k-cycles on* $\overline{N}$.

We say that the set $(a)_k$ [substitute (c_k)] is *independent* of the choice of admissible neighborhoods N and N^1 if there exists a fixed neighborhood N^* with the following properties. The set $(a)_k[(c_k)]$ determined for any N smaller than N^* and sufficiently small N^1 corresponding to N is equivalent† on $\overline{N}[N]$ to the set $(a)_k[(c)_k]$ determined for any smaller N and corresponding sufficiently small N^1.

We assume that the sets $(a)_k$ *and* $(c)_k$ *are independent of the choice of N and* N^1 *in the preceding sense, and admit only neighborhoods N smaller than* N^*.

We shall show in this paper that our assumptions are always fulfilled in the analytic case and in certain other particular cases. In a later paper we shall show that these assumptions are always fulfilled for the most general critical set as defined above.

Let (N^0N) and (NN^1) be two admissible choices of N and N^1. Let $(a)_k$ and $(c)_k$ be the cycles described in (α) and (β) determined for the neighborhoods N and N^1.

4. Classification of cycles. Suppose now that g is the set of *all* critical points at which $f=0$.

We replace $(a)_k$ by an algebraically‡ equivalent set made up of two sets of k-cycles

$$(b)_k, \quad (l)_k, \tag{4.1}$$

so chosen that each cycle of $(l)_k$ bounds on $f<0$, and no linear combination of the cycles $(b)_k$ not null so bounds.

† We mean here that each cycle of the first set is dependent on cycles of the second set and vice versa. Dependent and independent are terms always understood with respect to bounding. Cycles and chains are here taken in the absolute sense or, with obvious changes, mod m with $m>1$ and prime. The phrase "a complete set of k-cycles etc." may be replaced by "a set containing the maximum number of k-cycles etc." Terms in analysis situs will in general be used in the senses defined by Alexander, *Combinatorial analysis situs*, these Transactions, vol. 28 (1926), p. 301. Chains will be understood, however, to be symbolic expressions for oriented complexes, singular or non-singular in the sense of Veblen, The Cambridge Colloquium, Part II, *Analysis Situs*. The Colloquium Lectures by Lefschetz were not out at the time of the writing of this article but the author knows that they will be helpful to the reader.

‡ By algebraically equivalent we mean that a non-zero multiple of each member of the first set algebraically equals a linear combination of members of the second set and vice versa. Various of these sets may be null.

Let l_{k-1} be any linear combination of cycles of the set $(l)_{k-1}$. We have

$$L_k \to l_{k-1} \text{ on } f < 0,$$

where L_k is a k-chain. Suitable multiples of the cycles (4.1) bound on N^1. Without loss of generality, we can suppose the cycles (4.1) themselves bound on N^1. Thus

$$u_k \to - l_{k-1} \text{ on } N^1$$

where u_k is a k-chain on N^1.

We now introduce the k-cycle

$$L_k + u_k = \lambda_k$$

which we will say *links* l_{k-1}. We shall term u_k and L_k the *upper* and *lower* parts respectively of λ_k. We shall denote by $(\lambda)_k$ the set of k-cycles which link the respective $(k-1)$-cycles of $(l)_{k-1}$.

On the cycles $(b)_k$, $f<0$, and hence is less than some negative constant e. These cycles are thus on the domain

$$f \leqq e. \tag{4.2}$$

They must form a subset of a complete set of k-cycles on (4.2). There then exist other k-cycles $(i)_k$ on (4.2) such that the two sets

$$(b)_k, \quad (i)_k$$

form a complete set of k-cycles on $f \leqq e$.

We can, without loss of generality, suppose the set $(i)_k$ lies on A, since it is homologous on (4.2) to such a set.

We introduce the following table of complete sets of cycles with appropriate terms:

$(c)_k$: *critical cycles*;
$(l)_k$: *linkable cycles*; $(\lambda)_k$: *linking cycles*;
$(i)_k$: *invariant cycles*; $(b)_k$: *newly bounding cycles*.

Any linear combination of k-cycles of any one of these sets will be called by the same name.

With each linkable cycle l_{k-1} we associate a linking k-cycle λ_k so that λ_k and l_{k-1} arise from the same linear combination of corresponding cycles of the sets $(\lambda)_k$ and $(l)_{k-1}$. We note that λ_k cannot be null unless l_{k-1} is null. For if λ_k were null, closed k-cells with points at which $f \geqq 0$ on u_k, the upper part of λ_k, would have to cancel among themselves and we would have l_{k-1} bounding on $\overline{N}^1$ contrary to the nature of the set $(a)_{k-1}$ upon which l_{k-1} depends.

5. **A complete set of k-cycles on B.** Let the orthogonal trajectories of the contour manifolds f constant (Morse I) be represented in the form $\dot{x}_i = -f_{x_i}$ where $\dot{x}_i$ stands for the derivative of x_i with respect to the time. At each critical point p we understand that there exists a trajectory coincident with p for all time.

The following lemma is fundamental.

DEFORMATION LEMMA. *Any cycle on z_k on B satisfies an homology $z_k \sim z_k' + z_k'' = z_k^*$ on B in which z_k' is a k-chain on N^1 and z_k'' a k-chain on $f<0$ while the boundary z_{k-1}' of z_k' is on $\overline{N}^1$.*

Let each point p of z_k be deformed along the orthogonal trajectory of f through p, starting at p when the time $t=0$, and moving along this trajectory until t equals a time t_0. Denote the resulting k-cycle by z_k^*. For t_0 sufficiently large z_k^* must consist of points either on N^1, or $f<0$, as one readily proves.

Let z_k' be the chain of k-cells of z_k^*, which, closed, are wholly on N^1, and z_k'' the chain of remaining k-cells of z_k^*. We see that z_k' lies on N^1. We see also that if z_k^* had been sufficiently finely subdivided, z_k'' would lie on $f<0$.

The lemma follows directly.

We call z_k' and z_k'' respectively the *upper* and *lower* parts of z^*.

We shall prove the following theorem.

THEOREM 1. *A complete set of k-cycles for B is formed by the sets $(i)_k$, $(\lambda)_k$, $(c)_k$.*

In other words a complete set of k-cycles for B is obtained by deleting the k-cycles of A which are newly bounding on B, keeping the invariant k-cycles of A, and adding the linking and critical k-cycles.

We shall prove this theorem with the aid of two lemmas.

LEMMA 1. *Any k-cycle z_k on B is dependent on linking, critical, and invariant k-cycles.* (*We admit any integer $n \neq 0$ as coefficient of z_k.*)

We first deform z_k into z_k^* in accordance with the Deformation Lemma, obtaining thereby the cycle z_{k-1}' of the lemma. Since z_{k-1}' is on $\overline{N}^1$, and bounds on $f<0$, it satisfies an homology

$$nz_{k-1}' - l_{k-1} \sim 0 \quad \text{on} \quad \overline{N}, \; n \neq 0, \tag{5.1}$$

where l_{k-1} is a linkable $(k-1)$-cycle.

Let λ_k be the k-cycle linking l_{k-1}. The boundary of the lower part of the k-cycle,

$$nz_k^* - \lambda_k, \tag{5.2}$$

bounds on $\overline{N}$ according to (5.1), and its upper part is on N^1. Hence† the k-cycle (5.2) may be written as a sum of k-cycles on $f<0$ and k-cycles on N.

† The proof holds as stated even when $k=0$, regarding cycles with negative subscripts as null.

But on the one hand k-cycles on $f<0$ are dependent on B on invariant cycles. On the other hand k-cycles on N are dependent on N^0 on critical k-cycles and k-cycles on $\overline{N}^0$, that is, on $f<0$. But k-cycles on $f<0$ are dependent on B on invariant k-cycles.

Thus the k-cycle (5.2) is dependent on B on invariant and critical k-cycles, and the lemma is proved.

LEMMA 2. *The k-cycles of the theorem are independent on B.*

Suppose we had a relation of the form

$$D_{k+1} \rightarrow \lambda_k + c_k + i_k = D_k$$

where i_k, c_k, and λ_k are respectively invariant, critical, and linking k-cycles, and D_{k+1} a chain on B, and D_k its boundary.

We shall prove successively that λ_k, c_k, i_k are null.

We could deform D_{k+1} as in the Deformation Lemma into a set of points either on $f<0$ or N^1. Moreover, we could obtain the same result holding D_k fast, altering the deformation T of the lemma as follows.

Without loss of generality we can suppose c_k and λ_k are replaced by equivalent cycles null with the given cycles, and such that the new c_k, and u_k, the upper part of the new λ_k, are so near g that T will deform points within a sufficiently small positive distance r of c_k and u_k only through points on N^1.

Regard T as a movement depending continuously on the time t as t varies from 0 to 1. Let ρ be the initial distance of any point from D_k. We now perform T stopping the movement of a point initially within a distance r of D_k at a time $t=\rho/r$. We obtain thereby the desired deformation.

We can suppose that D_k bounds a chain D_{k+1} whose points lie on $f<0$ or else N^1.

Let d_{k+1} be a chain of all $(k+1)$-cells of D_{k+1} which, when closed, lie wholly on N^1. Let d_k be the boundary of d_{k+1}.

The points on D_{k+1} at which $f\geqq 0$ form a closed set on N^1. Accordingly if we suppose D_{k+1} sufficiently finely divided any closed k-cell of D_{k+1} possessing a point at which $f\geqq 0$ would be a k-cell of d_{k+1}, and would be on the boundary d_k if and only if it is on the boundary D_k of D_{k+1}. But boundary k-cells of D_k with points at which $f\geqq 0$ are found at most on c_k+u_k where u_k is the upper part of λ_k.

We see then that on the k-chain

$$d_k - c_k - u_k = e_k \tag{5.3}$$

the cells with points at which $f\geqq 0$ all cancel. Moreover e_k is on N^1, since d_k, c_k, and u_k are each on N^1. Thus e_k is on $\overline{N}^1$.

But the boundary of e_k is the boundary l_{k-1} of $-u_k$, since d_k and c_k are cycles. But a linkable cycle l_{k-1} does not bound on $\overline{N}$ unless null. Thus l_{k-1} is null. Accordingly λ_k is null as well as u_k.

We see then that

$$d_k = c_k + e_k.$$

But this is impossible unless c_k is null, for otherwise the critical cycle c_k would be homologous on N^1 to a cycle e_k on $\overline{N}^1$, since d_k bounds on N^1. Hence c_k is null.

Thus d_k is on $\overline{N}^1$. According to its origin it is homologous to i_k on $f<0$. Moreover, for some integer n not zero,

$$nd_k \sim l_k + b_k$$

on $\overline{N}$, where l_k and b_k are linkable and newly bounding cycles respectively. But l_k bounds on $f<0$. Hence $b_k \sim nd_k \sim ni_k$ on $f<0$. But this is impossible unless both b_k and i_k are null. Thus i_k is null.

The lemma is thereby proved.

The theorem follows at once from Lemmas 1 and 2.

6. **The associated ideal critical points.** With any critical set g we now associate a set of ideal critical points of *type k*. The number of points in this set will be denoted by M_k and called the *kth type number* of g.

The kth type number M_k of g shall be defined as the number of cycles in the sets $(a)_{k-1}$, and $(c)_k$ of §3.

This type number depends only on f and the topological properties of the neighborhoods of g. Accordingly, if $\bar{g}$ is a critical set composed of the sum of a finite number of distinct critical sets, the corresponding kth type number will be the sum of the kth type numbers of the component sets.

Let α and β be any two ordinary values of f $(\alpha<\beta)$. On the domain $f<\alpha$ there will be a complete set of $(k-1)$-cycles, independent on $f<\alpha$, but bounding on $f<\beta$. These we call *newly bounding relative* to the change from α to β. On $f<\beta$ there will be a complete set of k-cycles, independent on $f<\beta$, and independent on $f<\beta$ of k-cycles on $f<\alpha$. These we call *new k-cycles* relative to the change from α to β.

We shall evaluate the type number M_k.

The number of $(k-1)$-cycles in $(a)_{k-1}$ equals the number of newly bounding $(k-1)$-cycles and linking k-cycles in complete sets, as follows from the decomposition (4.1). The critical cycles $(c)_k$ are new k-cycles, and taken with the linking k-cycles form a complete set of new k-cycles as follows from Theorem 1.

We thus have the following theorem.

THEOREM 2. *If a and b are two ordinary values of f, $a<b$, between which lies just one critical value, the kth type number of the corresponding critical set will equal the number of newly bounding $(k-1)$-cycles plus the number of new k-cycles in complete sets, taking these sets relative to a change from $f\leqq a$ to $f\leqq b$.*

Let g be the critical set of the theorem. We shall divide the M_k ideal critical points of type k associated with g, into two sets of points in number m_k^+ and m_k^-, called respectively critical points of *increasing* or *decreasing* type.

The number m_k^+ shall be the number of linking k-cycles and critical k-cycles in complete sets, and the number m_k^- shall be the number of newly bounding $(k-1)$-cycles in a complete set.

We now have the following corollary of Theorems 1 and 2. See Morse I §18.

COROLLARY. *The kth Betti number of $f\leqq b$ minus the kth Betti number of $f\leqq a$ affords a difference given by the formulas*

$$\begin{aligned}\Delta R_k &= \overset{+}{m}_k - \overset{-}{m}_{k+1},\\ M_k &= \overset{+}{m}_k + \overset{-}{m}_k \qquad (k = 0, \cdots, n),\end{aligned}\tag{6.1}$$

where m_k^+ and m_k^- are respectively the numbers of ideal critical points associated with the critical set g, of increasing and decreasing k-type respectively. Here $m_0^- = m_{n+1}^- = 0$.

The preceding corollary also clearly holds if a and b are any two ordinary values of f. If then we eliminate either the integers m_k^- or m_k^+ from (6.1) we obtain the following theorem.

THEOREM 3. *Let a and b be any non-critical constants, $a<b$. Then between the changes in the Betti numbers as we pass from the domain $f\leqq a$ to the domain $f\leqq b$, and the sums of the type numbers of the critical sets with critical values between a and b the following relations hold:*

$$\begin{aligned}M_0 - M_1 + \cdots + (-1)^i M_i &= \Delta(R_0 - R_1 + \cdots + (-1)^i R_i) + (-1)^i \overset{-}{m}_{i+1},\\ M_0 - M_1 + \cdots + (-1)^i M_i &= \Delta(R_0 - R_1 + \cdots + (-1)^{i-1} R_{i-1}) + (-1)^i \overset{+}{m}_i,\end{aligned}$$

where $i=0, \cdots, n$ and $m_{n+1}^- = 0$.

Many inequalities and other equalities can be deduced from these relations. See, for example, Morse III §14 and §15, also Morse I §19. Note the interesting relation obtained from (6.1)

$$M_k = \Delta R_k + \overset{-}{m}_k + \overset{-}{m}_{k+1}.$$

A similar formula can be obtained involving the increasing type numbers.

It will be convenient for the next proof to replace the words "the number of k-cycles in a complete set of k-cycles" by the words "the *count* of k-cycles."

We now state a generalization of Theorem 2.

THEOREM 4. *If a and b are any two ordinary values of f the sum M_k of the kth type numbers of the critical sets with critical values between a and b will exceed or equal the count u of new k-cycles plus the count v of newly bounding $(k-1)$-cycles relative to a change from $f \leqq a$ to $f \leqq b$, $a < b$.*

Let c change from a to b taking on successively between a and b a set of ordinary values, separating the critical values. Let c_1 and c_2 be two such successive values.

Let h be the count of $(k-1)$-cycles on $f \leqq c_1$, independent on $f \leqq c_1$, bounding on $f \leqq c_2$. Let h^1 be the count of the subset of such cycles dependent on $f \leqq c_1$ on cycles of $f \leqq a$. We have $h^1 \leqq h$. Summing for all such changes of c,

$$v = \sum h^1 \leqq \sum h.$$

Now let m be the count of k-cycles on $f \leqq c_2$, independent on $f \leqq c_2$ of cycles on $f \leqq c_1$. Let m^1 be the count of the subset of such cycles independent on $f \leqq b$ of cycles on $f \leqq c_1$. Summing for all changes of c we have

$$u = \sum m^1 \leqq \sum m.$$

Combining these results we have

$$u + v \leqq \sum h + \sum m = M_k,$$

and the theorem is proved.

7. **The (ϕf) trajectories.** Let ϕ be a function of (x) of class C'' in the neighborhood of a point p. Suppose p is an ordinary point of both f and ϕ and that the gradients of f and ϕ at p are not parallel. By the (ϕf) *vector field* we mean the set of vectors, at each point q near p, obtained by projecting the gradient of ϕ on the tangent $(n-1)$-plane of the manifold $f=c$ at q.

By the (ϕf) *trajectories* we mean regular curves of class C' tangent at each point to vectors of the (ϕf) vector field.

The condition that the gradients of f and ϕ be not parallel is the following:

$$A(x) = \sum_{ik} [\phi_i f_k - \phi_k f_i]^2 \neq 0 \qquad (i, k = 1, \cdots, n) \tag{7.1}$$

where the subscripts i and k indicate partial differentiation with respect to the variables x_i and x_k.

The differential equations of the (ϕf) trajectories have the form*

* The summation convention of tensor analysis is used throughout.

$$\frac{dx_i}{dt} = \rho(f_k f_k \phi_i - \phi_k f_k f_i) = X_i(x), \qquad \rho \neq 0, \tag{7.2}$$

where ρ is a function of (x) of class C' near p. The right hand members of (7.2) do not all vanish near p as follows from (7.1).

On the (ϕf) *trajectories* f *is constant.*

For along such a trajectory we have

$$\frac{df}{dt} = f_i \frac{dx_i}{dt} = f_i X_i \equiv 0. \tag{7.3}$$

If we choose ρ as the reciprocal of

$$f_k f_k \phi_i \phi_i - \phi_k f_k f_i \phi_i = A(x), \tag{7.4}$$

then $d\phi/dt$ will be one along each trajectory. The function (7.4) is not zero since it equals the function $A(x)$ of (7.1) as indicated.

We can then so choose the parameter t *on the* (ϕf) *trajectories that at each point we have* $\phi = t$, *and this choice we suppose made.*

There is a (ϕf) trajectory through each point near p, and one through each point of the contour manifolds $\phi = c$ near p. The intersections of these trajectories and manifolds will vary continuously with the constants c and the trajectories.

We shall term the ordinary orthogonal trajectories of the contour manifolds, ϕ constant, the ϕ trajectories. We take their differential equations in the form

$$\frac{dx_i}{dt} = \frac{\phi_i}{\phi_k \phi_k} = Y_i(x). \tag{7.5}$$

So taken we may suppose $t = \phi$ at each point of a trajectory.

8. **Neighborhood functions.** The existence of neighborhood functions ϕ, as we shall define them, will enable us to express the type numbers in the simplest possible form.

Let g be any connected critical set of f on which $f=0$. Relative to g we shall call a function $\phi(x)$ a *neighborhood function* if it satisfies the following conditions.

(a) It is of class C'' in the neighborhood of g.

(b) It takes on a relative minimum zero on g.

(c) At points near g but not on g it is ordinary.

(d) At points near g but not on g at which $f=0$ the gradients of f and ϕ are not parallel.

We shall exhibit neighborhood functions ϕ in certain important cases beginning with the analytic case.

THEOREM 5. *If the function f is analytic, the function* $\phi=f_if_i$ *is an admissible neighborhood function* ϕ *relative to the critical set g of f.*

The function ϕ clearly satisfies (a) and (b). We shall finish by proving the following lemma.

LEMMA. *If f is analytic, any analytic function which takes on a proper relative minimum zero on g, is an admissible neighborhood function* ϕ.

Such a function ϕ satisfies (a) and (b). It must then satisfy (c). For g is a set of critical points of ϕ, and if ϕ were not ordinary near g the critical set g would be a subset of a larger connected critical set. But on all connected critical loci an analytic function is constant. Thus ϕ would be zero at some points near g not on g, contrary to the nature of a proper minimum. Thus (c) holds.

Now (d) could fail only at points not on g at which

$$A(x) = 0, \quad f = 0, \tag{8.1}$$

when $A(x)$ is given by (7.1).

But (8.1) is satisfied on g. Suppose it were satisfied on a larger analytic locus γ connected with g. Let h be any regular curve along which (8.1) is satisfied. On $h, f=0$ so that

$$f_i\frac{dx_i}{dt} = 0. \tag{8.2}$$

I say that on h

$$\phi_i\frac{dx_i}{dt} = 0.$$

This is certainly true on g and follows from $A(x)=0$ for points on h not on g. Thus ϕ is constant on h and accordingly on γ. It must then be zero on γ. From (b) we see that $\gamma=g$. Thus (d) holds.

The lemma and theorem are thereby proved.

We return now to the non-analytic case.

We term a critical set at which f takes on a relative maximum or minimum a *maximizing* or *minimizing* set respectively. We then state the following theorem, an immediate consequence of our definition of a neighborhood function ϕ.

THEOREM 6. *For a minimizing or maximizing set at which* $f=0$, *the functions f and* $-f$ *are respectively admissible neighborhood functions.*

In this connection we note the following.

If the critical values are isolated there are at most a finite number of minimizing or maximizing critical sets.

In fact the number of distinct contour manifolds on $f=c$ cannot vary as c approaches a critical value from either side. But the number of minimizing sets or maximizing sets cannot exceed the total number of these manifolds, and hence is finite.

There may however be an infinite number of distinct critical sets not minimizing or maximizing.

A neighborhood function ϕ always exists in the non-degenerate case as stated in the following theorem.

THEOREM 7. *If $(x)=(a)$ is a non-degenerate critical point of f the function*

$$(x_i - a_i)(x_i - a_i) = \phi$$

is an admissible neighborhood function.

The function clearly satisfies all the requirements, possibly excepting the one regarding gradients.

Suppose $(a)=(0)$. The relations of gradients will be unaltered if we use an orthogonal transformation to bring f to the form

$$f = \lambda_k x_k x_k + \eta \qquad (k = 1, \cdots, n),$$

where λ_k is a constant not zero, and η is of at least the second order with respect to the distance ρ to the origin.

If we can show that the function (7.4) does not vanish for real points on $f=0$ and for $\rho\neq 0$ in some neighborhood of the origin the proof will be complete.

Omitting terms of at least the fifth order this function (7.4) is seen to be

$$16\left[\lambda_k^2 x_k x_k x_i x_i - \lambda_k x_k x_k \lambda_i x_i x_i\right].$$

But on $f=0$ this becomes, up to the terms of at least the fifth order,

$$16\lambda_k^2 x_k x_k x_i x_i .$$

The ratio of the last expression to ρ^4 is positive and bounded away from zero for $\rho\neq 0$.

The function (7.4) is accordingly positive everywhere desired and the theorem is proved.

THEOREM 8. *If f is analytic and $(x)=(0)$ is an isolated critical point, the function $\phi=x_i x_i$ is an admissible neighborhood function.*

The theorem follows from the lemma under Theorem 5.

9. **The radial trajectories.** We shall now prove the existence of a set of trajectories termed *radial* trajectories. They lead away from g somewhat after the fashion of rays emanating from a point. The theorem is

THEOREM 9. *If ϕ is a neighborhood function for g, then on the domain*

$$R: \qquad 0 < \phi \leqq r, \ f \leqq 0, \tag{9.1}$$

where r is a small positive constant, there exists a radial field of trajectories, one through each point of R, satisfying differential equations of the form

$$\frac{dx_i}{dt} = B_i(x), \quad B_iB_i \neq 0,$$

where the functions $B_i(x)$ are of class C' on R. These trajectories reduce to (ϕf) trajectories on $f=0$. On them t may be taken equal to ϕ. As t increases they pass out of R only by reaching $\phi = r$.

The ϕ trajectories themselves would do except for the fact that they cross $f=0$ in general. We shall alter the ϕ trajectories neighboring $f=0$ so that they will suffice.

The $(f\phi)$-trajectories ζ emanating from $f=0$ on R in general form a field only for a short distance from $f=0$ depending upon how near ϕ is to zero. We shall be more precise and say that we can determine a negative function $h(\alpha)$ of class C', for $0<\alpha\leqq r$, such that the field persists on a trajectory ζ on which $\phi=\alpha$, and on which f decreases from zero to $h(\alpha)$.

We can in fact define $h(\alpha)$ successively on the intervals with end points

$$r, \ r/2, \ r/4, \cdots,$$

and so define $h(\alpha)$ on its entire interval.

Now let $M(z)$ be a function of (z) of class C' for $z\geqq 0$, identically one for $z>1$ and zero for z zero, otherwise positive.

Our radial trajectories will now be defined by (7.5), except for the above points on ζ where f decreases from zero to $h(\phi)$. At these points the differential equations of the radial trajectories shall have the form

$$\frac{dx_i}{dt} = X_i(x) + M\left[\frac{f(x)}{h(\phi(x))}\right][Y_i(x) - X_i(x)] \tag{9.2}$$

where X_i and Y_i are the functions of (7.2) and (7.5) respectively.

On $f=0$ the radial trajectories reduce to the (ϕf) trajectories (7.2). For $f=h(\phi)$ they take the form (7.5). Moreover on them

$$\frac{d\phi}{dt} = \phi_iX_i(1-M) - Y_i\phi_iM = 1 - M + M = 1.$$

This shows that we can take $t=\phi$ on the radial trajectories. It also shows that the right hand members of (9.2) are not all zero at any one point on R.

The theorem follows at once.

10. **The type numbers in terms of neighborhood functions.** By a *radial* deformation we shall mean one in which each point moves on a radial trajectory, and points for which ϕ is constant are deformed into points for which ϕ is constant.

The domains of points satisfying $\phi = e$ or $0 < \phi \leqq e$, where e is a small positive constant, less than the constant r of Theorem 9, will be respectively denoted by ϕ_e and ϕ_e^0. The points on these same domains at which $f < 0$ will be denoted by $\bar{\phi}_e$ and $\bar{\phi}_e^0$.

From the existence of the radial trajectories we infer the following.

(1) For any two constants e and η less than r the domains $\bar{\phi}_e$ and $\bar{\phi}_\eta$ are homeomorphic.

(2) If $e < \eta$ the domain $\bar{\phi}_\eta^0$ can be radially deformed onto the domain $\bar{\phi}_e^0$ leaving ϕ_e^0 fixed and never increasing ϕ.

(3) For any closed point set on $\bar{\phi}_\eta^0$ there exists a radial deformation of $\bar{\phi}_\eta^0$ that leaves $\bar{\phi}_\eta$ fixed and deforms the point set onto $\bar{\phi}_\eta$.

We can now prove the following theorem.

THEOREM 10. *If a neighborhood function ϕ exists, the sets of cycles $(a)_k$ and $(c)_k$ of* §3 *exist, and are independent of the choices of admissible neighborhoods N and N^1.*

As a choice of the fixed neighborhood N^* of the definition of independence of §3 we take the domain $\phi \leqq r$. If N be any neighborhood on $\phi \leqq r$ let η be a positive constant so small that the points on $\phi \leqq \eta$ lie on N. Corresponding to N a sufficiently small choice of the neighborhood N^1, as we shall see, will be any neighborhood N^1 on $\phi \leqq \eta$. Let e be a positive constant so small that the domain $\phi \leqq e$ consists of points on N^1.

It appears, then, that relative to N and N^1, the cycles $(a)_k$ may be taken as a complete set on $\bar{\phi}_e$ independent on $\bar{\phi}_e$, bounding on $\phi \leqq e$, while the k-cycles $(c)_k$ may be taken as a complete set on $\phi \leqq e$ independent on $\phi \leqq e$ of the cycles on $\bar{\phi}_e$. We have then the following theorem.

THEOREM 11. *The type number M_k of a critical set is the number of cycles in the following two sets:*

(a) *a complete set of $(k-1)$-cycles on $\bar{\phi}_e$ independent on $\bar{\phi}_e$, bounding on $\phi \leqq e$;*

(b) *a complete set of k-cycles on $\phi \leqq e$ independent on $\phi \leqq e$ of the cycles on $\bar{\phi}_e$.*

The number of cycles in these sets is independent of the choice of the constant e for e positive and sufficiently small.

A cycle of any set (a) has been termed linkable if it bounds a chain L_k on $f<0$. If it so bounds it must bound outside of $\phi<e$ as well. For the part of L_k on $\bar{\phi}_e^0$ can be radially deformed on $\bar{\phi}_e^0$ onto $\bar{\phi}_e$.

We summarize and complete these results as follows.

The linkable $(k-1)$*-cycles are those on* $\bar{\phi}_e$*, independent on* $\bar{\phi}_e$*, bounding on* $\phi\leqq e$*, and bounding on* $f<0$ *outside of* $\phi<e$.

The newly bounding $(k-1)$*-cycles are those on* $\bar{\phi}_e$ *independent on* $\bar{\phi}_e$*, independent on* $f<0$ *outside of* $\phi<e$*, but bounding on* $\phi\leqq e$.

The critical k-cycles are the cycles (b).

The number M_0 equals the number of critical 0-cycles. It is one for connected minimizing sets, and null for all other connected sets. It is of increasing type. The number M_n is the number of newly bounding $(n-1)$-cycles. This is true if we are operating in euclidean n-space or on a portion, not all, of a connected regular n-spread. For there are then no non-bounding n-cycles, in particular no linking or critical n-cycles. Unless a connected critical set is a maximizing set, M_n is null, for there are then no $(n-1)$-cycles in (a). It is obviously one for connected maximizing sets.

The following corollary of Theorems 8 and 11 brings all of Brown's results on isolated critical points under the results of the present paper.

COROLLARY. *Suppose* $(x)=(0)$ *is an isolated critical point in the analytic case. Then if we set* $\phi=x_ix_i$ *the kth type number* M_k*, for k not zero, is given by the formula*

$$M_k = R_{k-1} - \delta_1{}^k$$

where R_k *is the kth Betti number of the region on the* $(n-1)$*-sphere* $\phi=e$ *on which f is negative. In the case of a minimum* $M_0=1$. *Otherwise* M_0 *is null.*

This follows from Theorem 11, noting that in (a) of Theorem 11 all cycles on the $(n-1)$-sphere $\phi=e$ are bounding on the interior $\phi\leqq e$ excepting a point 0-cycle. From this exception the Kronecker delta $\delta_1{}^k$ arises. In (b) the complete set is null except in the case of a minimum, and then the origin may serve as a complete 0-set.

Because of its signal importance we state the following as a separate theorem.

THEOREM 12. *In the analytic case Theorem* 11 *holds without exception with* $\phi=f_if_i$.

From Theorems 6 and 11 we have the following theorem, true in the most general non-analytic case. In it we suppose $f=0$ on g.

THEOREM 13. *For a connected minimizing or maximizing set g, the type numbers of g always exist.*

For a minimizing set, M_k is the kth Betti number of the domain $f \leqq e$, neighboring g.

For a maximizing set, M_k is the number of cycles in the following two sets:

(a) *a complete set of $(k-1)$-cycles on $f=-e$, independent on $f=-e$, bounding on $f \geqq -e$, neighboring g;*

(b) *a complete set of k-cycles on $f \geqq -e$ independent on the same domain of the k-cycles on its boundary.*

Consider a connected maximizing set.

In euclidean n-space the set (b) is empty. Let the domain $f \geqq -e$ neighboring g be denoted by S. If R_i and β_i denote Betti numbers of S and its boundary β, respectively, we have*

$$R_{i-1} + R_{n-i} = \beta_{i-1} \qquad (i = 1, \cdots, n).$$

Now there are R_{n-i} independent $(i-1)$-cycles on the residue of S in n-space. To add S to its residue is then to diminish the $(i-1)$st Betti number by R_{n-i}. There must then be at least R_{n-i} $(i-1)$-cycles independent on β, and newly bounding. But there cannot be more than R_{n-i} such cycles independent on β bounding on S, since R_{i-1} cycles on β do not bound on S. Thus $M_i = R_{n-i}$.

11. The type number of a non-degenerate critical point. This case is the most important in the many geometric applications.

Suppose $(x)=(0)$ is a non-degenerate critical point at which $f=0$. Let the quadratic form $f_{ij}x_ix_i$, in which the partial derivatives are evaluated at the origin, be carried by a linear transformation into a form with squared terms only. The number of terms with negative coefficients thereby resulting is called the *index* of the critical point.

We wish to establish anew the results of Morse I. Our problem is primarily to prove the following theorem.

THEOREM 14. *If k is the index of a non-degenerate critical point, the kth type number $M_k = 1$, while all other type numbers are zero.*

It follows from our initial definition of type numbers that such a number, if it exists, will be independent of any one-to-one transformation of the neighborhood of the critical points of class C''. That the type number exists in this case is a consequence of the existence of the neighborhood function $x_ix_i = \phi$, as affirmed in Theorem 7.

* Apply Alexander's duality relations to β (these Transactions, vol. 23 (1922), p. 348).

Now we have shown in Morse I that in the neighborhood of the origin the function can be carried by a transformation of the above sort into the form

$$f = -y_1^2 - \cdots - y_k^2 + y_{k+1}^2 + \cdots + y_n^2. \tag{11.1}$$

The function y_iy_i is again an admissible neighborhood function. We see from Theorem 13 that the theorem is true if $k=0$ or n.

Suppose then that $0<k<n$. From Theorem 11 we have the following lemma.

LEMMA. *The ith type number M_i equals the $(i-1)$st Betti number R_{i-1} of the domain*

$$y_jy_j = 1, \ f < 0 \qquad (i, j = 1, \cdots, n), \tag{11.2}$$

minus one if $i=1$.

Our problem is then to determine the Betti numbers of the spherical region (11.2).

We shall prove that the domain (11.2) can be deformed on itself into the $(k-1)$-sphere

$$y_1^2 + \cdots + y_k^2 = 1, \ y_{k+1}^2 + \cdots + y_n^2 = 0, \tag{11.3}$$

by a deformation T that leaves (11.3) fixed. Accordingly the Betti numbers of (11.2) will be those of the $(k-1)$-sphere, and the theorem will follow from the lemma.

The deformation T will now be given.

Corresponding to any point $(x)=(a)$ on (11.2) there is a unique angle α such that

$$a_1^2 + \cdots + a_k^2 = \cos^2\alpha, \ a_{k+1}^2 + \cdots + a_n^2 = \sin^2\alpha, \ 0 \leqq \alpha < \frac{\pi}{4}, \tag{11.4}$$

and every point (a) satisfying (11.4) is on (11.2).

Now hold each such point (a) and corresponding α fast. In the required deformation the point (y) shall not move if it is initially on (11.3). If (y) is initially at a point (a) not on (11.3) it shall move as follows:

$$\begin{aligned} y_i &= \frac{\cos(\alpha t)}{\cos\alpha} a_i && (i = 1, \cdots, k), \\ y_j &= \frac{\sin(\alpha t)}{\sin\alpha} a_j && (j = k+1, \cdots n), \ \alpha \neq 0, \end{aligned} \tag{11.5}$$

as t varies from 1 to zero.

It is easily seen that this affords the required deformation except possibly for continuity of movement of points near (11.3). But for these points α is near zero, and we have from (11.5) that

$$y_{k+1}^2 + \cdots + y_n^2 = \sin^2(\alpha t),$$

so that for α near zero the variables $y_{k+1}, \cdots, y_n$ are uniformly near zero.

Thus the deformation is continuous.

The theorem now follows from the lemma.

The principal theorems on type number relations in Morse I now follow.

12. **A justification of the definitions of type numbers.** We shall investigate how the type numbers of a function change with variation of the function.

Let $F(x, \mu)$ be a function of (x) and a set of parameters (μ), analytic in (x) and (μ) for (x) on Σ and (μ) neighboring (0), and such that

$$f(x) \equiv F(x, 0).$$

For (μ) sufficiently near (0), $F(x, \mu)$ will satisfy our boundary conditions and possess critical points lying only in arbitrarily small neighborhoods of the critical sets of g.

We state the following lemma.

LEMMA 1. *If a and b are any ordinary values of f with $a<b$, then for (μ) sufficiently near (0) the domains $f \leqq b$ and $F \leqq b$ are homeomorphic under a transformation that makes $f \leqq a$ and $F \leqq a$ correspond.*

The lemma is easily established by using the orthogonal trajectories of f. A deformation can be set up along these trajectories which affords the homeomorphism, moving only those points which are very near $f=a$ and $f=b$. See Morse I §7.

We note that to prove this lemma we need only to have $F(x, \mu)$ of class C''.

Let g be a connected critical set of f. Recall that $\phi = f_i f_i$ is a neighborhood function for g. We now give a lemma which enables us to avoid critical sets with the same critical value.

LEMMA 2. *Corresponding to the critical set g of $F(x, 0)$ there exists a function $\psi(x)$ of class C'' throughout Σ with the following properties:*

(1) *Except when $\phi < e$, $\psi(x) \equiv 0$, where e is a small positive constant.*

(2) *When $\phi < e_1$, $\psi(x) \equiv \rho$, where ρ is an arbitrarily small positive constant and e_1 is a positive constant less than e.*

(3) *For (μ) sufficiently near (0), $F+\psi$ has no other critical points than those of F.*

Let $H(z)$ be a function of z of class C'' for $z \geqq 0$, one for $z < e_1$ and zero for $z > e$. The function $\psi(x)$ will now be defined as zero except when $\phi < e$, and when $\phi < e$ it will be defined by the equation

$$\psi(x) = \rho h[\phi(x)].$$

One sees that ψ has all the required properties, except possibly (3). But (3) could fail only when

$$e_1 < \phi < e. \tag{12.1}$$

For $\rho = 0$ and $(\mu) = (0)$ we have $F + \psi = f$. Moreover $F + \psi$ is of class C'' in (x), (μ), and ρ. On the domain (12.1) the gradient of f is not null. Accordingly for ρ and (μ) sufficiently near $\rho = 0$ and $(\mu) = (0)$ respectively, the gradient of $F + \psi$ is not null.

Thus ψ satisfies (3) and the lemma is proved.

Our definition of type numbers is justified by the following theorem.

THEOREM 15. *If (μ) be sufficiently near $(\mu) = (0)$, the critical points of $F(x, \mu)$ will appear only in sets arbitrarily near the critical sets of f. Corresponding to each critical set g of f the critical points of F which lie in g's neighborhood have a kth type number sum at least as great as the kth type number of g.*

By virtue of Lemma 2 we will lose no generality if we suppose the critical values at the different critical sets of f are all different. The addition of ψ in Lemma 2 did not change the position or type numbers of the critical points of F for (μ) sufficiently near (0).

Let g be any critical set of f whose critical value is separated from the other critical values of f by constants a and b. Let γ be the set of critical points of $F(x, \mu)$ which lie in the neighborhood of g. Let (μ) be taken so near (0) that the critical values of F on γ lie between a and b, and so that the homeomorphism of Lemma 1 holds.

The type number M_k of the set g relative to f will equal the number N_k of newly bounding $(k-1)$-cycles and new k-cycles in complete sets, relative to a change from the domain $f \leqq a$ to $f \leqq b$. By virtue of the homeomorphism of Lemma 1 the number N_k will be the same relative to F. But according to Theorem 4 the kth type number sum of the critical points γ of F will exceed or equal N_k, that is, the type number M_k of g.

The theorem is thereby proved.

Our work is further justified by the following theorem.

THEOREM 16. *If f is analytic there exists a set of constants (μ) arbitrarily near (0) such that the critical points of the function*

$$F(x, \mu) = f(x) + \mu_i x_i \qquad (i = 1, \cdots, n)$$

are non-degenerate and lie in arbitrarily small neighborhoods of the critical points of f.

Corresponding to each connected critical set g of f the critical points of $F(x, \mu)$ which lie in g's neighborhood have a kth type number sum at least as great as the kth type number of g.

The condition that $F(x, \mu)$ have no degenerate critical points is that the equations

$$f_{x_i} + \mu_i = 0$$

have no solution at which the Hessian of f vanishes. That a choice of the constants (μ) can be so made arbitrarily near $(\mu) = (0)$ follows from a general theorem formulated by Kellogg.*

The second statement in the theorem now follows from Theorem 15.

---

* Kellogg, *Singular manifolds among those of an analytic family*, Bulletin of the American Mathematical Society, vol. 35 (1929), p. 711.

Harvard University,
Cambridge, Mass.