ROCKY MOUNTAIN
JOURNAL OF MATHEMATICS
Volume 3, Number 2, Spring 1973

MORSE THEORY IN HILBERT SPACE

E. H. ROTHE

To the Memory of Heinz Hopf

1. **Introduction.** Let E be a real Hilbert space, and let f be a real valued function whose domain V is a subset of E to be specified later. Morse theory deals with the critical points of f, i.e., those points $x_0 \in V$ at which the Fréchet differential $df(x; h)$ (see e.g. [2]) of f is zero (identically in h). Since, for any $x \in V$, the differential $df(x; h)$ is a linear bounded functional in h, there exists a unique element $g = g(x)$ in E, called the gradient of f, such that

$$df(x; h) = \langle g(x), h \rangle, \tag{1.1}$$

where $\langle \cdot, \cdot \rangle$ denotes the scalar product in E. (In a finite-dimensional space this definition is easily seen to agree with the usual one of grad f as the vector whose components are the partial derivatives of f.) It follows that a point $x \in V$ is critical if and only if it satisfies the equation

$$g(x) = 0. \tag{1.2}$$

Related to the notion of a critical point is that of a critical level:

DEFINITION 1.1. A critical level (or value) of f is a real number c such that $f(x) = c$ for at least one critical point x.

Two problems arise naturally:

PROBLEM I. Describe the "nature" of a critical point, a local problem.

PROBLEM II. Obtain an estimate of the number of critical points in terms of geometrical (topological) properties of the domain V of f, a problem in the large.

Problems I and II will be discussed in §§2 and 3 respectively in detail. At the present moment we confine ourselves to some introductory and intuitive remarks concerning these problems in very simple cases. These remarks motivate the use of the more abstract notions employed later, in particular the use of the singular homology theory. A review of the relevant definitions and facts of this theory forms the last part of this introduction.

A simple case of Problem I. Let $E = E^2$ be the Euclidean plane of points $x = (x_1, x_2)$, let $V = D$ be a disc with center $\theta = (0, 0)$, and let f be a diagonalized quadratic form. We consider the cases

Received by the editors February 11, 1972.

AMS (MOS) subject classifications (1970). Primary 49F15, 49-01; Secondary 58E05.

$$(\alpha)\ f = x_1^2 + x_2^2, \quad (\beta)\ f = -x_1^2 + x_2^2, \quad (\gamma)\ f = -x_1^2 - x_2^2.$$

The three cases differ by the number j of negative squares, the index of the quadratic form. (A point mass situated at θ is stable in case (α), unstable in case (γ), and stable or unstable in case (β) according to the initial displacement. Thus the index j gives a measure of instability.) Obviously the point θ is a critical point of f furnishing a minimum in case (α), a saddle point in case (β), and a maximum in case (γ). For any real number a let

$$f_a = \{x \in V \mid f(x) < a\}, \tag{1.3}$$

and

$$A = f_a, \qquad B = f_a \cup \{\theta\}. \tag{1.4}$$

Then

$$B = \theta, \qquad A = \emptyset, \qquad \text{the empty set in case } (\alpha),$$

$$B = V, \qquad A = V - \{\theta\} \quad \text{in case } (\gamma)$$

while in case (β), A is the (open) shaded area, and B is this area augmented by the point θ. (See Fig. 1.)

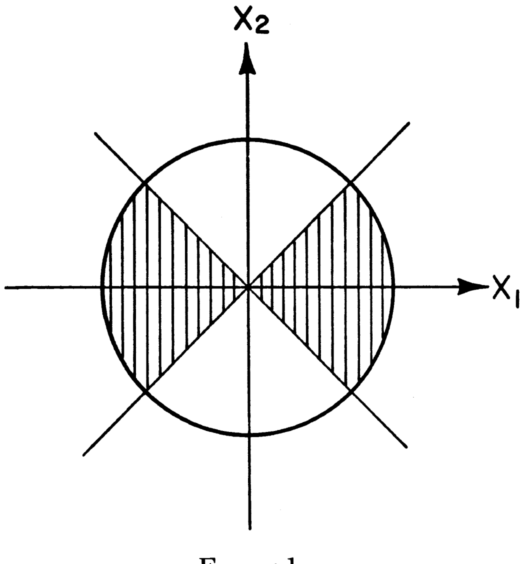

FIGURE 1

This simple example suggests that the nature of the critical point may be described by considering the couple (B, A) of the sets defined in (1.4) and by examining the change in topological properties connected with the transition from A to B. (Note e.g. that in case (β) the set B is connected but A is not.)

A simple case of Problem II. Let V be a torus tangent to the (x_1, x_2)-plane E^2 as indicated in Fig. 2, and let $x_3 = f(x)$ be the distance of the point $x = (x_1, x_2, x_3) \in V$ from E^2.

Obviously, p_1, p_2, p_3, p_4 are the critical points of f, and $c_i = f(p_i)$ $(i = 1, 2, 3, 4)$ are the critical levels. We make two remarks:

REMARK 1. The set f_a defined in (1.3) does not change its "homotopy type" if a varies without crossing one of the critical levels c_i. For instance for $c_2 < a < c_3$, f_a can be deformed into a cylinder; this is no more true if a leaves the interval $[c_2, c_3]$ by crossing one of its endpoints. The effect of crossing e.g. c_3 may be described by comparing the topological properties of the set f_{a_3} with those of the set f_{a_2} where $c_2 < a_2 < c_3 < a_3 < c_4$.

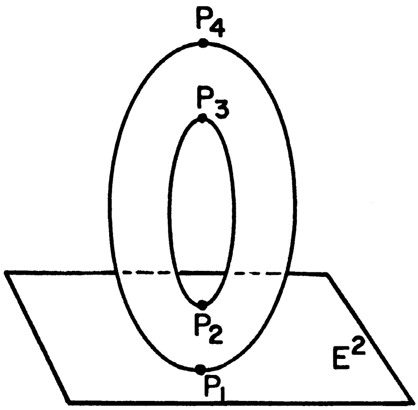

FIGURE 2

REMARK 2. In Fig. 2, p_1 is a minimum point for f, p_2 and p_3 are saddle points, and p_4 is a maximum. Consequently by the preceding discussion concerning Problem I, the indices j of these critical points are respectively 0, 1, 1, 2. Thus if M_j denotes the number of critical points of index j we see that $M_0 = 1$, $M_1 = 2$, $M_2 = 1$. But the integers at the right of these equalities are just the "Betti number" R_0, R_1, R_2 resp. of the torus (see e.g. [1, p. 212]). Thus

$$M_q = R_q, \qquad q = 0, 1, 2. \tag{1.5}$$

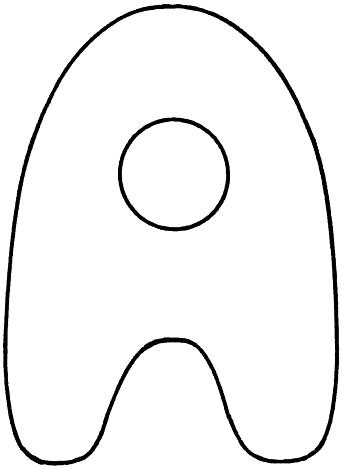

FIGURE 3

However, if we deform the torus to look like Fig. 3 [22, p. 8], then the Betti numbers remain unchanged while $M_0 = 2$, $M_1 = 3$, $M_2 = 1$. Thus

$$(1.6) \qquad M_q \geqq R_q, \qquad q = 0, 1, 2.$$

These inequalities, in contrast to (1.5), are true under rather general conditions as are the celebrated Morse inequalities

$$(1.7) \qquad \begin{aligned} M_0 &\geqq R_0, \\ M_1 - M_0 &\geqq R_1 - R_0, \\ M_q - M_{q-1} + \cdots &+ (-1)^q M_0 \\ &\geqq R_q - R_{q-1} + \cdots + (-1)^q R_0, \qquad q = 0, 1, 2, \cdots, \end{aligned}$$

which imply (1.6). They were proved by Morse under proper assumptions in the case that the domain V of f is of finite dimension n. In this case the equality holds for $q = n$. The validity of (1.6) and (1.7) will be discussed later on.

In our discussion of Problem I as well as in that of Problem II we were led to compare the members of a couple (B, A) of sets where A is a subset of B. This observation gives a plausible reason for the important role which the "relative homology groups $H_q(B, A)$ of B modulo A" play in the Morse theory (cf. Definition 2.1 and (3.1)), for — speaking crudely — these groups afford a topological comparison of the sets A and B.

Since in general there is no triangulation in Hilbert space we will use the *singular* relative homology groups. For the precise definition of these groups we have to refer to topology books (see e.g. [5, Chapter VII]). Here we confine ourselves to the following more or less intuitive remarks:

A singular q-simplex σ_q in E is a map into E of the simplex Δ_q whose vertices are the unit points on the coordinate axes of Euclidean $(q+1)$-space. The support $|\sigma_q|$ of σ_q is the point set $\sigma_q(\Delta_q) \subset E$. A singular q-chain with coefficient group G (a module over a principal ideal domain — integers and fields are the most frequent cases) is a finite "linear combination"

$$c_q = \sum_i g_i \sigma_q{}^i$$

of singular simplices $\sigma_q{}^i$ with coefficients $g_i \in G$. The support $|c_q|$ of c_q is defined by

$$|c_q| = \bigcup_i |\sigma_q{}^i|.$$

If S is a subset of E we often write "$c_q \subset S$" for $|c_q| \subset S$.

Starting from the elementary definition for the boundary $\partial\Delta_q$ of the simplex Δ_q, the boundary $\partial\sigma_q$ of the singular simplex σ_q is defined as a $(q-1)$-chain in such a way that

$$\partial\partial = 0 \tag{1.8}$$

(see [5, p. 186]), and this definition of ∂ is extended by linearity to singular chains.

We recall that a "couple" (B, A) in E consists of two subsets B and A of E with $A \subset B$. A singular q-chain z_q with $z_q \subset B$ and $\partial z_q \subset A$ is called a relative q-cycle on B modulo A. With an obvious definition of addition these relative cycles z_q form an abelian group $Z_q = Z_q(B, A)$.

A singular q-chain z_q is called bounding on B mod A if there exist a $(q+1)$-chain $c_{q+1} \subset B$ and a q-chain $a_q \subset A$ such that $\partial c_{q+1} = z_q + a_q$. It follows from (1.8) that such a z_q is necessarily a cycle on B modulo A. Thus the group B_q of chains bounding on B modulo A is a subgroup of Z_q.

The qth homology group $H_q(B, A)$ of B modulo A is then defined as the quotient group Z_q/B_q. An element $\hat{z}_q$ of this group is thus a coset modulo B_q of Z_q; i.e., it is a class of relative cycles on B modulo A any two elements $z_q{}^1, z_q{}^2$ of which satisfy the equivalence relation "$z_q{}^1 \sim z_q{}^2$ if and only if $z_q{}^1 - z_q{}^2$ is a bounding cycle on B modulo A".

We now state a few properties of the relative homology groups which are basic for their application to Morse theory.

Let (B, A) and (D, C) be two couples in E for which

$$(B, A) \supset (D, C) \tag{1.9}$$

(which means: $B \supset D$ and $A \supset C$), and let I denote the closed unit interval. We then say that a map

$$\delta(x, t) : (B \times I, A \times I) \to (B, A)$$

deforms (B, A) into (D, C) if

$$\left.\begin{array}{l}\delta(x, 0) = x \\ \delta(x, 1) \in D\end{array}\right\} \text{ for } x \in B, \text{ and } \delta(x, 1) \in C \text{ for } x \in A.$$

Property H_1 (*homotopy property*). Let $\delta(x, t)$ deform (B, A) into (D, C) in such a way that the restriction of δ to $(D \times I, C \times I)$ maps this couple into (D, C). Then $H_q(B, A) \approx H_q(D, C)$ where the symbol "$\approx$" denotes group isomorphism.

Remark to Property H_1. The assumption just made on the deformation δ is in particular satisfied if $\delta(x, t) = x$ for $(x, t) \in (D \times I)$. In this case (D, C) is called a deformation retract of (A, B) (see [5, p. 30]).

Property H_2 (*Addition theorem* (*see* [5, *Theorem* I. 13.2])). Let the subsets B_1, B_2 of E have a positive distance and suppose that $A_i \subset B_i$ for $i = 1, 2$. Then

$$H_q(B_i \cup B_2, A_1 \cup A_2) \approx H_q(B_1, A_1) \dotplus H_q(B_2, A_2),$$

where the symbol "$\dotplus$" denotes "direct sum."

Property H_3 (*Excision theorem* (*see* [5, §§VII.9 *and* IV.9])). $H_q(B, A) \approx H_q(B - U, A - U)$ if U is a set whose closure (in B) is contained in the interior of A.

Property H_4. $H_q(A, A) = 0$ for every subset A of E [5, I.8.1].

Property H_5. Let $A \subset B \subset C \subset E$ with a nonempty A. It is asserted:

(α) if $H_q(C, B) = 0$, then $H_q(C, A) \approx H(B, A)$;

(β) if $H_q(B, A) = 0$, then $H_q(C, B) \approx H_q(C, A)$ [5, I.10.4].

Property H_6. Let A, B, C be as in H_5. It is asserted:

(α) if B is a deformation retract of C, then $H_q(C, A) \approx H_q(B, A)$;

(β) if A is a deformation retract of B, then $H_q(C, B) \approx H_q(C, A)$.

This follows easily from H_1, H_4, and H_5.

2. **Problem I, the local problem.** Let B be a closed ball in the Hilbert space E with the zero element θ as center. Let f be a real valued function with domain B for which θ is the only critical point. For the sake of simplicity we assume that

$$f(\theta) = 0. \tag{2.1}$$

DEFINITION 2.1. The qth critical group $C_q(\theta)$ attached to the critical point θ of f is defined by

$$C_q(\theta) = C_q(\theta; f) = H_q(f_0 \cup \{\theta\}, f_0) \tag{2.2}$$

where f_0 is given by (1.3) (with $a = 0$ and $V = B$).

It is easily seen from the excision theorem H_3 of the introduction that this definition is independent of the radius of B as long as θ is the only critical point in B. Actually one could replace the spherical neighborhood B of θ by an arbitrary one.

Our first goal is to determine $C_q(\theta)$ if f is a "nondegenerate quadratic form" (Theorem 2.1). For later use we define the notion of form and nondegeneracy for the case when the degree of the form is an arbitrary integer $p \geqq 2$.

DEFINITION 2.2. Let $Q(x_1, x_2, \cdots, x_p)$ be a function from $E \times E \times \cdots \times E$ (a product of p factors) to the reals which is linear in each x_j, symmetric, and bounded in the sense that, for some constant K,

$$|Q(x_1, x_2, \cdots, x_p)| \leqq K\|x_1\| \, \|x_2\| \cdots \|x_p\| \tag{2.3}$$

where $\| \;\|$ denotes the norm in E. Then, for $x_1 = x_2 = x_p = x$,

$$Q(x) = Q(x_1, x_2, \cdots, x_p) \tag{2.4}$$

is called a p-form on E. If $p = 2$, Q is called quadratic.

DEFINITION 2.3. The p-form $Q(x)$ is called nondegenerate if there exists a positive constant m such that, for all $x \in E$,

$$\|\operatorname{grad} Q(x)\| \geqq m \, \|x\|^{p-1}. \tag{2.5}$$

LEMMA 2.1. *Let $p = 2$, i.e., let Q be a quadratic form, and let* $\kappa(x) = \operatorname{grad} Q(x)$. *Then*

(α) *κ is a linear bounded symmetric map of E into E,*

(β) *Q is nondegenerate if and only if κ is nonsingular, i.e., if κ has a bounded everywhere-defined inverse.*

PROOF. The differential of Q at x with "increment h" equals $2Q(x, h)$. Therefore, by the definition of κ,

$$\langle \kappa(x), h \rangle = 2Q(x, h). \tag{2.6}$$

This equality together with (2.3), (2.4) and the symmetry of $Q(x, h)$ is easily seen to imply (α).

We turn to the proof of (β). That the nonsingularity of κ implies the nondegeneracy of Q is trivial since the boundedness of κ^{-1} obviously implies (2.6). To prove the converse part of (β), we remark first that the range $R(\kappa)$ of κ is closed as is easily seen from the fact that if

$Y_n = \kappa(x_n)$ is a Cauchy sequence then, by (2.6) and the linearity of κ, the sequence x_n is likewise a Cauchy sequence. But, as is well known, the closedness of $R(\kappa)$ together with the symmetry of κ implies that E is the direct sum of $R(\kappa)$ and the nullspace $N(\kappa)$ of κ. Now $N(\kappa) = \{\theta\}$ as is clear from (2.6). It follows that $E = R(\kappa)$ and that κ has an everywhere-defined inverse λ. The boundedness of λ follows again from (2.6).

We will need the following lemma which was proved by Hestenes ([**6**, Theorem 7.1]; see also [**12**, §2.5]).

LEMMA 2.2. *Let Q be a quadratic form on E. Then there exists a direct decomposition of E into the linear subspaces $E = E^+ + E^- + E^0$ with the following properties*:

$$\begin{aligned}
\langle x^0, x^+\rangle = \langle x^0, x^-\rangle = \langle x^+, x^-\rangle &= Q(x^+, x^-)\\
&= 0 \quad \textit{for } x^+ \in E^+, x^- \in E^-, x^0 \in E^0;\\
Q(x^+) &> 0 \quad \textit{for } x^+ \in E^+ - \theta;\\
Q(x^-) &< 0 \quad \textit{for } x^- \in E^- - \theta,\\
Q(x^-, x) &= 0 \quad \textit{for } x^- \in E^- \textit{ and all } x \in E.
\end{aligned}$$

If Q is nondegenerate then $E^0 = \{\theta\}$.

DEFINITION 2.4. The index j of the quadratic form $Q(x)$ is the maximal dimension of linear subspaces L of E such that $Q(x) < 0$ for $x \in L - \{\theta\}$. (j may be a finite integer or ∞.)

Without proof we state

LEMMA 2.3. *The index j of the quadratic form Q equals the dimension of the space E^- (see Lemma 2.2).*

The linear space E^- should not be confused with the set $Q_0 = \{x \in B \mid Q(x) < 0\}$. In Fig. 1, E^- is the x_1-axis while Q_0 is the shaded area.

THEOREM 2.1. *Let Q be a nondegenerate quadratic form on E and let j be its index. Then*

$$C_q(\theta; Q) = \begin{cases} 0 & \textit{if } q \neq j, \\ G & \textit{if } q = j, \end{cases} \tag{2.7}$$

where the left member is defined by (2.2) *and where G is the coefficient group.*

PROOF. By Definition 2.1 (with $f = Q$),

$$C_q(\theta; Q) = H_q(Q_0 \cup \{\theta\}, Q_0). \tag{2.8}$$

We set $B^- = B \cap E^-$ with E^- as in Lemma 2.2 and claim

$$H_q(Q_0 \cup \{\theta\}, Q_0) \approx H_q(B^- \cup \{\theta\}, B^-). \tag{2.9}$$

For the proof of (2.9) it will, by the homotopy theorem H_1, be sufficient to show that the couple occurring at the right member of (2.9) is a deformation retract of the couple occurring at the left. Let $x \in Q_0$. By Lemma 2.2,

$$x = x^- + x^+ \quad \text{where } x^- \in E^- \cap B \text{ and } x^+ \in E^+ \cap B.$$

For $0 \leqq t \leqq 1$ we set $\delta(x, t) = x^- + (1 - t)x^+$. Using Lemma 2.2, we see that

$$\begin{aligned} Q(\delta(x, t)) &= Q(x^-) + (1 - t)^2 Q(x^+) \\ &\leqq Q(x^-) + Q(x^+) = Q(x) < 0. \end{aligned}$$

This inequality together with the convexity of B obviously proves that $B^- = B \cap E^-$ is a deformation retract of Q_0. Since the deformation δ also retracts $Q_0 \cup \{\theta\}$ onto $B^- \cup \{\theta\}$ the validity of (2.9) is established. Now by Lemma 2.3, E^- is a Hilbert space of dimension j. Therefore we see from (2.8) and (2.9) that

$$C_q(\theta; Q) \approx H_q(B^j, B^j - \theta) \tag{2.10}$$

where $B^j = B \cap E^-$, a closed ball in E^- with center θ. But the sphere S forming the boundary of B^j is a deformation retract of $B^j - \theta$. Therefore (2.10) together with Property H_6 implies that

$$C_q(\theta; Q) \approx H_q(B^j, S). \tag{2.11}$$

Now in case of a finite j it is well known that the right member of (2.11) is isomorphic to the right member of (2.7) (see e.g. [5, §I, Theorem 16.4]).

It remains to consider the case $j = \infty$. Then B^j is a ball in an infinite-dimensional Hilbert space and the boundary S of B^j is a deformation retract of B^j as was proved by Kakutani [7] in the separable case; for the general case, see [3]. It follows that the right member of (2.11) is isomorphic to $H_q(S, S) = 0$ if $j = \infty$ (cf. Property H_4). Since q is finite and therefore $q \neq j$, the proof of (2.6) is complete.

The importance of Theorem 2.1 lies in the fact that, for a large class of functions f, a "first approximation" in the neighborhood of a critical point is given by a quadratic form. To make this point clear we recall Taylor's formula

$$\text{(2.12)} \qquad \begin{cases} f(x) - f(\theta) = \displaystyle\sum_{r=1}^{p} \frac{1}{r!} d^r f(\theta; x) + R_p, \\ R_p = \dfrac{1}{p!} \displaystyle\int_0^1 \ d^{p+1} f(tx; x)(1-t)^p \, dt \end{cases}$$

which holds in some neighborhood of θ if $f \in C^{p+1}(\theta)$, i.e., if f has continuous differentials up to and including order $p + 1$ in some neighborhood of θ (see e.g. [2, p. 186]).

Thus if θ is a critical point of a function $f \in C^{p+1}(\theta)$ and if $f(\theta) = 0$ then the quadratic form

$$\text{(2.13)} \qquad d^2 f(\theta; x)/2$$

is the first term in the Taylor formula. The problem arises to reduce the computation of the critical group $C_q(\theta; f)$ to that of $C_q(\theta; Q)$ where Q is the quadratic form (2.13). To discuss this problem we need

DEFINITION 2.5. The critical point θ of f is called nondegenerate if $f \in C^2(\theta)$ and if the quadratic form (2.13). is nondegenerate in the sense of Definition 2.3. The index of θ as critical point of f is then defined as the index of the quadratic form (2.13).

The following lemma, due to Morse, solves our problem for finite-dimensional spaces.

LEMMA 2.4. *Let E be n-dimensional with points* $x = (x_1, x_2, \cdots, x_n)$ *and let* $f \in C^3(\theta)$ *be a real valued function for which* θ *is a critical point. Then for a small enough neighborhood N of* θ *the following statements hold:*

(α) *There exist functions* $a_{ik}(x)$, *symmetric in i, k, such that*

$$a_{ik}(\theta) = \frac{1}{2} \frac{\partial^2 f}{\partial x_i \partial x_k} \quad \text{and} \quad f(x) = \sum_{i,k}^{n}{}^{1} a_{ik}(x) x_i x_k.$$

(β) *Under the additional assumption that the critical point* θ *is nondegenerate, there exists a differentiable invertible map* $\psi : N \rightarrow E$ *with* $\psi(\theta) = \theta$ *such that*

$$\text{(2.14)} \qquad f(\phi(y)) = -\sum_{i=1}^{j} y_i^2 + \sum_{i=j+1}^{n} y_i^2 \quad \textit{where } \phi = \psi^{-1};$$

the index j of this quadratic form equals the index of the quadratic form (2.13).

(γ) *The critical group at* θ *of f is isomorphic to that of the quadratic form* (2.14).

That (γ) follows from (β) is obvious since ϕ is a homeomorphism. For the proof of (α) and (β) we refer the reader to [14, §§9 and 10]; we only note that the proof of (β) is based on (α) and a procedure which is quite analogous to the Lagrange method of bringing a quadratic form to the diagonal form by a *linear* transformation.

The above "Morse lemma" 2.4(β) was generalized to an arbitrary Hilbert space by Palais [13, pp. 307, 308] in the following form:

LEMMA 2.5. *Let the assumptions of Lemma* 2.4(β) *be satisfied with E being an arbitrary Hilbert space. Then there exists an orthogonal projection P in E, and N, ψ, ϕ of the properties described in Lemma* 2.4(β) *with the right member of* (2.14) *replaced by* $-\|y - P(y)\|^2 + \|P(y)\|^2$. *The dimension of* $(1 - P)E$ *equals the index of the form* (2.13).

There is another method of reducing the investigation of the critical group of f at the critical point θ to that of the quadratic form (2.13) which is based on the fact that the function (2.13) is a "good" approximation of f if θ is not degenerate. Such a method was employed by Seifert and Threlfall [22, p. 34] in the finite-dimensional case. Since this procedure is not based on diagonalization, we may admit nondegeneracy of higher order as defined in

DEFINITION 2.6. Let p be an integer $\geqq 2$, and let $f \in C^{p+1}(\theta)$. Then θ is called a nondegenerate critical point of order p for f if $df(\theta; x) = d^2f(\theta; x) = \cdots = d^{p-1}f(\theta; x) = 0$ while the p-form $d^pf(\theta; x)$ is nondegenerate (cf. Definition 2.3).

We note that, by (2.12),

$$(2.15) \qquad f(x) - f(\theta) = \frac{1}{p!} d^pf(\theta; x) + \int_0^1 \frac{d^{p+1}f(tx; x)(1-t)^p\,dt}{p!}$$

if θ is a nondegenerate critical point of order p. We also note the following:

LEMMA 2.6. *Let* $f \in C^{p+1}(\theta)$, *and suppose that* $f(\theta) = 0$. *Then* θ *is a nondegenerate critical point of order p for f if and only if there exist positive constants k, K and R such that*

$$k\|x\|^{p-1} \leqq \|g(x)\| \leqq K\|x\|^{p-1} \quad \textit{for } \|x\| < R$$

where $g = \operatorname{grad} f$ (*for a proof see* [21, *Lemma* 4.3]).

Our goal is to prove under suitable assumptions that

$$(2.16) \qquad C_q(\theta; f) = C_q(\theta; d^pf(\theta; x)/p!).$$

The proof of (2.16) may be based on N. H. Kuiper's generalization of the Morse lemma 2.4(β) to the case of nondegeneracy of order $p > 2$ [7a, p. 202, Corollary], or on the following approximation theorem.

THEOREM 2.2. *Let θ be a nondegenerate critical point of f of order p. Let $f(\theta) = 0$. Let m and R be two positive constants such that $\|g(x)\| \geqq 2m\|x\|^{p-1}$ for $\|x\| < R$ (such constants exist by Lemma 2.6). About the "approximation" ψ to f we make the following assumptions:*

(1) $\psi \in C^{p+1}(\theta)$, $\psi(\theta) = 0$,

(2) $\gamma = \operatorname{grad} \psi$ *has the following property: to each positive η there corresponds a positive $\rho = \rho(\eta) < R$ such that $\|g(x) - \gamma(x)\| < \eta\|x\|^{p-1}$ for $\|x\| < \rho$.*

Then θ is a nondegenerate critical point of order p for ψ, and

$$C_q(\theta; f) = C_q(\theta; \psi). \tag{2.17}$$

COROLLARY TO THEOREM 2.2. *The assumptions on f in Theorem 2.2 are sufficient for the validity of* (2.16).

The complicated proof of Theorem 2.2 is given in [21, §6]. It is based on a generalization to Hilbert space of the concept of a "cylindrical neighborhood" of an isolated critical point introduced by Seifert and Threlfall [22, §9] in the finite-dimensional case. For the proof of the corollary one has to show that $\psi = d^p f(\theta; x)/p!$ satisfies the assumptions made on ψ in Theorem 2.2. Note that, by (2.15) (with $f(\theta) = 0$), $|f(x) - \psi(x)|$ is of order $\|x\|^{p+1}$ from which it may be proved that $\|\operatorname{grad}(f - \psi)\| = \|g - \gamma\|$ is of order $\|x\|^p$. But for $\eta > 0$ given, $\|x\|^p < \eta\|x\|^{p-1}$ for $\|x\|$ small enough, and thus the inequality in assumption (2) of Theorem 2.2 is verified. For details see [21, §4, last paragraph].

In the above application of Theorem 2.2 the approximating function ψ was a Taylor approximation to f. Under more restrictive assumptions Theorem 2.2 can also be applied with ψ being a finite-dimensional approximation to f. More precisely, the following theorem holds.

THEOREM 2.3. *Let f satisfy the assumptions of Theorem 2.2. In addition it is assumed that, in some neighborhood of θ,*

$$f(x) = \|x\|^p/p + F(x), \qquad p \text{ even and} \geqq 2, \tag{2.18}$$

with $G(x) = \operatorname{grad} F(x)$ being completely continuous. For a given n-dimensional linear subspace E^n of E let x^n denote the orthogonal projection of the point x of E into E^n, and let

$$f_n(x) = \|x\|^p/p + F(x^n).$$

Then there exists a subspace E^{n_0} such that for $E^{n_0} \subset E^n$ the following statements hold:

(α) $C_q(\theta; f) = C_q(\theta; f_n)$,

(β) $C_q(\theta; f) = C_q(\theta; \tilde{f}_n)$,

where $\tilde{f}_n$ denotes the restriction of f_n to the projection of its domain into E^n.

Functions f of the above form appear in the theory of nonlinear integral equations. See e.g. [23, §21.2] and [21, beginning of §6]. For the proof of Theorem 2.3 we refer to [21, Theorems 7.1 and 7.2]. We only remark that the lengthy proof of (α) is based on showing that $\psi = f_n$ satisfies the assumptions of Theorem 2.2 while (β) is an easy corollary to (α).

We conclude this section on the local problem with a theorem which establishes a connection between the Morse numbers M_q, i.e. the ranks of the groups $C_q(\theta; f)$ (over the coefficient group G) and the Leray-Schauder index $d(g, \theta)$ [8, p. 54] of $g = \operatorname{grad} f$ at θ. By the original definition as given in [8, p. 54], the index $d(g, \theta)$ is defined only for $p = 2$ since, by (2.18), $g(x) = x\|x\|^{p-2} + G(x)$. However, by Lemma 2.6, $\|g(x)\|$ is bounded away from zero on every (small enough) sphere with center θ. This makes it possible to define the index in terms of the definition of the order as given in [15, p. 375] in spite of the fact that the factor of x in the expression for g vanishes at θ. Cf. [16, footnote on p. 459].

THEOREM 2.4. *Let f satisfy the assumptions of Theorem 2.3. Moreover we assume that $f \in C^{p+2}(\theta)$ and that g is uniformly differentiable in some neighborhood of θ. Then*

$$d(g, \theta) = \sum_{q=0} (-1)^q M_q. \tag{2.19}$$

For the proof see [16, Theorems 6.1, 7.1], [2, Lemma 7.27] and [18, Lemma 2.3]. We note that the sum in (2.19) is finite as follows from the reduction to the finite-dimensional case established in Theorem 2.3 (β).

3. **Problem II, the problem in the large** (see the introduction). M. Morse started his investigations with the domain V of f being finite-dimensional. The following cases were treated (see [10] and the papers listed in the bibliography of that book):

(a) V is the closure of an open bounded domain in $E = E^n$ the dimension n being finite. It is supposed that $g = \operatorname{grad} f$ is outwardly directed at the boundary $\dot{V}$ of V (regular boundary condition).

(b) V is as in (a) but there are a finite number of points on $\dot{V}$ in which g has the direction of the interior normal (general boundary conditions).

(c) V is a finite-dimensional Riemannian manifold.

(d) The problem of the geodesics joining two given points p and q on a finite-dimensional Riemannian manifold V. Here the function $f(x)$ is the length of the curve x on $\tilde{V}$ joining p and q, and the domain V of f is the properly metrized space of such curves called the loop space. Thus V is infinite-dimensional. Morse treated this problem by an approximation procedure using finite-dimensional problems of type (c) above as approximations.

In all these cases the following "condition C" introduced by Palais and Smale (see e.g. [13, p. 300]) plays a basic rule for the generalization to a Hilbert space E:

DEFINITION 3.1. A real valued function f with domain $V \subset E$ satisfies condition C if for every set S whose closure $\bar{S}$ is contained in V the following is true: if on S, $|f|$ is bounded but $\|\text{grad}\, f\|$ is not bounded away from zero then grad f vanishes in some point of $\bar{S}$.

Condition C is automatically satisfied in the following special cases: (i) V is compact, therefore in particular in the finite-dimensional case if V is closed and bounded; (ii) V is a closed bounded set in the Hilbert space E and the gradient g of f is of the form $g(x) = x + G(x)$ with a completely continuous G; (iii) $V \subset E$ is a small enough neighborhood of the nondegenerate critical point θ of order $p \geqq 2$ (as can be seen easily from Lemma 2.6).

In all three cases (a), (b), (c), one wants to attach to an isolated critical level c a critical group $C_q(c)$ by setting

$$C_q(c) = H_q(\bar{f}_b, \bar{f}_a) \qquad (\bar{f}_\alpha = \{x \in V \mid f(x) \leqq \alpha\}), \tag{3.1}$$

where $a < c < b$ and where c is the only critical value in the closed interval $[a, b]$. The first task then is to show that (within the restrictions of a and b indicated) the right member of (3.1) is independent of the choice of a and b. The proof for this is, except for the use of condition C, essentially the same as the one given by Pitcher [14] in the finite-dimensional case. To be more concrete we will give an outline of the generalization to Hilbert space for case (a) which is the simplest among the three cases in question.

Let V be a bounded open set in the Hilbert space E, and let f be a real valued function whose domain is the closure $\bar{V}$ of V. The following assumptions (A)–(E) are made:

(A) This is an assumption on the geometry of V for whose technical details we refer to [21] and [20]. Suffice it to say that this assumption guarantees the existence of a unique "exterior" unit normal $n(x)$ to the boundary $\dot{V}$ of V at every point x of $\dot{V}$.

(B) f is not constant in any ball, and its differential $df(x;h)$ is uniformly bounded in V.

(C) The condition C of Definition 3.1 holds.

(D) The critical levels are isolated.

(E) $\langle n(x), g(x)\rangle > 0$ for $x \in \dot{V}$. ($g(x) = \operatorname{grad} f(x)$.)

THEOREM 3.1. *Let the levels a, b, c, be as specified in the two lines following* (3.1), *and let Assumptions* (A)–(E) *be satisfied. Then the right member of* (3.1) *is independent of a and b.*

Now it is easily seen that Theorem 3.1 is a consequence of

LEMMA 3.1. *If the closed interval $[\alpha, \beta]$ contains no critical level then $\bar{f}_\alpha$ is a deformation retract of $\bar{f}_\beta$ and*

$$H_q(\bar{f}_\beta, \bar{f}_\alpha) = 0. \tag{3.2}$$

To show that this lemma implies Theorem 3.1 let $a < a' < c < b' < b$ where a, b, c are as in Theorem 3.1 and apply the lemma to the intervals $[b', b]$ and $[a, a']$ which are free of critical levels. We then see from Property H_6 of the introduction that

$$H_q(\bar{f}_b, \bar{f}_a) = H_q(\bar{f}_{b'}, \bar{f}_a) = H_q(\bar{f}_{b'}, \bar{f}_{a'})$$

which proves Theorem 3.1.

We now sketch the proof of the first assertion of Lemma 3.1, the second one being a consequence. For $x_0 \in \bar{f}_b - \bar{f}_a$ let $x(t) = x(t, x_0)$ be the solution of the initial value problem

$$dx/dt = -\,g(x), \qquad x(0, x_0) = x_0. \tag{3.3}$$

It can be proved from Assumption (E) that $x(t, x_0) \in V$ for all positive t [21, Theorem 2.1]. Now by the chain rule, $df(x(t))/dt = \langle \operatorname{grad} f, dx/dt\rangle$, and therefore, by (3.3),

$$df(x(t))/dt = -\,\|g\|^2. \tag{3.4}$$

Thus $f(x(t)) \leqq f(x_0) \leqq b$ for $t \geqq 0$, and

$$f(x(t)) \in [a, b] \tag{3.5}$$

as long as $f(x(t)) \geqq a$. But by Assumption (D) the interval $[a, b]$ has a positive distance from the set of critical *levels*. From this it can be proved that an $x(t)$ satisfying (3.5) has a distance from the set of critical *points* which is not smaller than a positive constant $m =$

$m(a, b)$. But by condition C this implies the existence of a positive constant $m = m(a, b)$ such that $\|g(x(t))\| \geqq m$ for $x(t)$ satisfying (3.5). This allows us to conclude from (3.4) that $f(x(t)) \leqq f(x_0) - m^2t \leqq b - m^2t$. It follows the existence of a unique positive $T(x_0) \leqq (b - a)/m^2$ such that $f(x(t)) = a$ for $t = T(x_0)$. Therefore a deformation $\delta(x_0, t)$ retracting $\bar{f}_b$ onto $\bar{f}_a$ is given by

$$\delta(x_0, t) = \begin{cases} x(t, x_0), & \text{for } x_0 \in \bar{f}_b - \bar{f}_a, 0 \leqq t \leqq T(x_0), \\ x(T(x_0), x_0) & \text{for } x_0 \in \bar{f}_b - \bar{f}_a, T(x_0) < t \leqq (b - a)/m^2, \\ x_0 & \text{for } x_0 \in \bar{f}_a, 0 \leqq t \leqq (b - a)/m^2. \end{cases}$$

In the theorem just proved, the numbers a and b occurring in the definition of the critical group $C_q(c)$ (equation (3.1)) may be chosen arbitrarily close to c but different from c. It is natural to ask for conditions under which $C_q(c)$ can be expressed in terms containing only the level c. To this end we introduce

Assumption (F). Let $b > c$, and let c be the only critical value in $[b, c]$. Then there exists a deformation $\delta(x, t)$ of $\bar{f}_b$ into $\bar{f}_c$ for which points of $\bar{f}_c$ stay in $\bar{f}_c$ during the deformation. We say Assumption (F) is strictly satisfied if there exists such δ for which each point of $\bar{f}_c$ stays fixed during the deformation (i.e. if $\bar{f}_c$ is a deformation retract of $\bar{f}_b$).

THEOREM 3.2. *Let Assumptions* (A)–(F) *be satisfied, the last one strictly. Let* $\sigma(c)$ *denote the set of critical points at level* c. *Then*

$$C_q(c) = H_q(\bar{f}_c, \bar{f}_c - \sigma(c)) = H_q(f_c \cup \sigma(c), f_c). \tag{3.6}$$

We omit the proof (see [**21**, Theorem 3.4 and remark to Theorem 3.4 on p. 37]). However we will give an outline of the proof of the next theorem.

THEOREM 3.3. *Let assumptions* (A)–(E) *be satisfied and suppose that the set* $\sigma(c)$ *defined in Theorem* 3.2 *consist of a finite number of points. Then Assumption* (F) *is strictly satisfied.*

The proof is based on the solution $x(t) = x(t, x_0)$ of the initial value problem

$$\frac{dx}{dt} = \frac{-(f(x_0) - c)g(x)}{\|g(x)\|^2}, \qquad x(0, x_0) = x_0 \in (\bar{f}_b - \bar{f}_c). \tag{3.7}$$

Similar to the derivation of (3.4) from (3.3) we derive from (3.7) the relation $df(x(t))/dt = -(f(x_0) - c)$, or $f(x(t)) = f(x_0) - (f(x_0) - c)t$

which implies that $\lim_{t\to 1^-} f(x(t)) = c$. But since the denominator of the right member of (3.7) approaches 0 as x approaches a critical point it is not obvious that $\lim_{t\to 1^-} x(t)$ exists. However this limit relation can be proved under the assumptions of our theorem [**19**, Theorem 5.1]. Using this fact one obtains the desired retracting deformation $\delta(x_0, t)$ by setting

$$\delta(x_0, t) = \begin{cases} x(t, x_0) & \text{for } x_0 \in \bar{f}_b - \bar{f}_c,\ 0 \leqq t < 1, \\ \lim\limits_{t\to 1^-} x(t, x_0) & \text{for } x_0 \in \bar{f}_b - \bar{f}_c,\ t = 1, \\ x_0 & \text{for } x_0 \in f_c,\ 0 \leqq t \leqq 1. \end{cases}$$

This establishes Theorem 3.3. (For the proof that $\delta(x_0, t)$ is continuous jointly in (x_0, t) see [**21**, Appendix].) Under the assumptions of this theorem,

$$\sigma(c) = \bigcup_{\rho=1}^{r(c)} \sigma_\rho(c) \tag{3.8}$$

where $\sigma_1(c), \sigma_2(c), \cdots, \sigma_{r(c)}(c)$ denote the critical points at level c. Our next goal is to find the connection between the critical group $C_q(c)$ at level c and the critical groups $C_q(\sigma_\rho(c))$ of the critical points $\sigma_\rho(c)$ as defined in Definition 2.1 (with θ replaced by $\sigma_\rho(c)$ and the level 0 by the level c).

Our first step is to "localize" the expression (3.6) for $C_q(c)$; let $W(c)$ be an open neighborhood of $\sigma(c)$ whose closure contains no other critical points than those of $\sigma(c)$. The existence of such W follows easily from [**21**, Lemma 2.3]. We excise the set $f_c \cup \sigma(c) - W(c)$ from the couple at the right member of (3.6). Using the excision property H_3 of the introduction we see from (3.6) that

$$C_q(c) \approx H_q(f_c \cap W(c) \cup \sigma(c), f_q \cap W(c)). \tag{3.9}$$

Let now $W_\rho(c)$ be a spherical neighborhood of $\sigma_\rho(c)$ whose closure contains no critical point except $\sigma_\rho(c)$. Then, for $W(c) = \bigcup_{\rho=1}^{r(c)} W_\rho(c)$, (3.9) holds. Assuming the $\overline{W_\rho(c)}$ to be disjoint we see from the addition theorem H_2 of the introduction that

$$C_q(c) = \dot{\sum_{\rho=1}^{r(c)}} H_q(f_c \cap W_\rho(c) \cup \sigma_\rho(c), f_c \cap W_\rho(c))$$

where the symbol $\dot{\Sigma}$ denotes the direct sum. Comparison with (2.2) yields

$$(3.10) \qquad C_q(c) = \sum_{\rho=1}^{r(c)} C_q(\sigma_\rho(c)).$$

This formula gives the desired relation between the group attached to a critical level and the groups attached to the critical points at that level provided there are only a finite number of such critical points. We now make the assumption that this is true for every critical level. Now under Assumptions (A)–(E) the number of critical levels can be proved to be finite (see [**21**, Lemma 2.3]). We denote them by $c_1, c_2, \cdots, c_N$, and set

$$(3.11) \quad r_i = r(c_i), \qquad \sigma_\rho{}^i = \sigma_\rho(c_i), \qquad W_\rho{}^i = W_\rho{}^i(c_i), \qquad \rho = 1, 2, \cdots, r_i.$$

We now introduce "Morse numbers" $M_q(\sigma_\rho{}^i)$, $M_q{}^i$, M_q by setting

$$(3.12) \quad M_q(\sigma_\rho{}^i) = \operatorname{rank} C_q(\sigma_\rho{}^i), \quad M_q{}^i = \operatorname{rank} C_q(c_i), \quad M_q = \sum_{i=1}^{N} M_q{}^i.$$

(Here "rank" means rank over the coefficient group.)

The following theorem gives an interpretation of the number $M_q{}^i$ and M_q in the "nondegenerate" case:

THEOREM 3.4. *Let Assumptions* (A)–(E) *be satisfied. In addition it is assumed that each critical point of f in V is nondegenerate of order $p = 2$* (*see Definitions* 2.5 *and* 2.6). *Then*

(α) *the number of critical points in V is finite* (*the proof shows this to be true if the critical points are nondegenerate of arbitrary order $p \geqq 2$*),

(β) *$M_q{}^i$ equals the number of critical points in V of index q at level c_i*,

(γ) *M_q equals the number of critical points in V of index q.*

PROOF. The set of critical points in V is compact (see [**21**, Lemma 2.3]). On the other hand, a nondegenerate critical point is isolated as is easily seen from Lemma 2.6. Moreover there are no critical points on the boundary $\dot{V}$ of V. These three facts together imply assertion (α) in an obvious manner.

But on account of Theorems 3.3 and 3.2, assertion (α) implies the validity of (3.6) and, therefore, of (3.10). It also implies that the assumption made in the paragraph above (3.11) is satisfied. Thus definitions (3.11) and (3.12) make sense. But by the Corollary to Theorem 2.2, $C_q(\sigma_\rho{}^i)$, the qth critical group at $\sigma_\rho{}^i$ for f, is isomorphic to the qth critical group of the second differential of f at $\sigma_\rho{}^i$, and by Theorem 2.1 this group equals the coefficient group G if q equals the index $j_\rho{}^i$ of this differential as quadratic form and is 0 otherwise. Its rank is therefore 1 if $q = j_\rho{}^i$, and zero otherwise. Since by Definition 2.5, $j_\rho{}^i$ is the index of $\sigma_\rho{}^i$ as critical point of f we see from (3.12) that

$$M_q(\sigma_\rho{}^i) = \begin{cases} 1 & \text{if } q = \text{index of } \sigma_\rho{}^i, \\ 0 & \text{if } q \neq \text{index of } \sigma_\rho{}^i. \end{cases}$$

Summing over ρ from 1 to r_i we obtain (β) by using (3.10) and (3.12). Assertion (γ) is an obvious consequence of (β) and (3.12).

We now turn our attention to the inequalities (1.6) of the introduction.

THEOREM 3.5. *Let the Assumptions* (A)–(F) *be satisfied. Let R_q denote the Betti numbers of V (i.e. the ranks of the groups $H_q(V)$) and let M_q be defined by* (3.12). *Then* (1.6) *holds.*

In this theorem no nondegeneracy assumptions are made. Thus $M_q{}^i$ and M_q may be infinite numbers.

A proof of this theorem based on ideas of Seifert and Threlfall [22, §5] is given in [21, §3].

We turn to the Morse inequalities (1.7) with R_q, M_q defined as in Theorem 3.5. Assuming that the critical values $c_1, c_2, \cdots, c_N$ are ordered by magnitude, let a_i be numbers such that

$$(3.13) \quad a_0 < c_1 < a_1 < \cdots < a_{i-1} < c_i < a_i < \cdots < a_{N-1} < c_N < a_N$$

where $a_N > \sup_{x \in V} f(x) > \inf_{x \in V} f(x) > a_0$ and let $A_i = \bar{f}_{a_i} \cap V$. (As already mentioned, Assumptions (A)–(E) imply that N is a finite number. Moreover f is bounded [21, Lemma 2.1].) Now the purely algebraic proof given by Pitcher [14, §11] in the finite-dimensional case for (1.7) shows that (1.7) holds if each group A_i is finitely generated provided N is finite and $f(x)$ is bounded so that the above choice of a_N and a_0 is possible. Since these two conditions are realized under assumptions (A)–(E) we can state

LEMMA 3.2. *Let Assumptions* (A)–(E) *be satisfied, and suppose that the groups $A_0, A_1, \cdots, A_N$ defined above are finitely generated. Then* (1.7) *holds.*

We now make the following assumptions (α), (β) and (γ) in addition to (A)–(E).

(α) Each critical point is nondegenerate of some order $p \geqq 2$.

As already mentioned this implies that the number of critical points is finite. We denote again by $\sigma_\rho{}^i$ ($\rho = 1, 2, \cdots, r_i$) the critical points at level c_i.

(β) Each critical point $\sigma_\rho{}^i$ $(i = 1, 2, \cdots, N;\ \rho = 1, 2, \cdots, r_i)$ is non-degenerate of even order $P_{i\rho}$, and in some neighborhood $W_\rho{}^i$ of $\sigma_\rho{}^i$, f is of the form

$$f(x) = (p_{i\rho})^{-1}\|x - \sigma_\rho{}^i\|^{p_{i\rho}} + F_{i\rho}(x - \sigma_\rho{}^i)$$

where the gradient of $F_{i\rho}$ is completely continuous (cf. the assumption for Theorem 2.3).

(γ) $H_q(V)$ is finitely generated.

THEOREM 3.6. *Let Assumptions* (A)–(E) *and* (α)–(γ) *be satisfied. Then the Morse relations* (1.7) *hold.*

PROOF. By Lemma 3.2 it will be sufficient to prove that the groups $H_q(A_i)$ are finitely generated for $i = 0, 1, \cdots, N$ and all q. Now $H_q(A_N) = H_q(V)$ is finitely generated by assumption (γ). Moreover $H_q(A_n, A_{n-1})$ is finitely generated. For by (3.1), (3.13) and (3.10),

$$H_q(A_N, A_{N-1}) = C_q(c_N) = \sum_{\rho=1}^{r_N} C_q(\sigma_\rho{}^N).$$

But it follows from Theorem 2.3(β) that each $C_q(\sigma_\rho{}^N)$ is finitely generated. Thus $H_q(A_N)$ and $H_q(A_N, A_{N-1})$ are finitely generated for all q. But this implies that $H_q(A_{N-1})$ is finitely generated as can be shown from the exactness of the homology sequence

$$H_{q+1}(A_N, A_{N-1}) \rightarrow H_q(A_{N-1}) \rightarrow H_q(A_N).$$

Replacing N by $N - 1$ one proves that $H_q(A_{N-2})$ is finitely generated. Continuation of this procedure proves our assertion.

So far we treated the generalization to Hilbert space only in the first of the cases (a)–(d) listed at the beginning of §3. We now continue with some remarks concerning the remaining cases.

REMARK ON THE CASE (b) (general boundary conditions). We make the following change in the assumptions made in case (a); Assumption (A) is strengthened by some additional differentiability conditions on the boundary $\dot{V}$ of V. Moreover, the function f is defined in an open bounded domain V_2 containing V and satisfying the strengthened Assumption (A) made for V. The boundaries $\dot{V}$ and $\dot{V}_2$ are supposed to have a positive distance.

Assumption (B) holds in V_2 and is strengthened by the requirement that the second differential of f is also uniformly bounded in V_2.

Assumption (C) is replaced by the stricter condition that $g(x) = \operatorname{grad} f(x) = x + G(x)$ with completely continuous G in V_2. Moreover if, for $x \in \dot{V}$, $g_t(x)$ denotes the component of g tangential to $\dot{V}$ then $\|g_t\|$ is supposed to be bounded away from 0 on any closed set $S \subset \dot{V}$ which contains no zero of g_t.

Assumption (E) is replaced by the condition that $g(x) \neq \theta$ on $\dot{V}$, and there exist a finite number of points $x_1, x_2, \cdots, x_r$ on $\dot{V}$ at which g has the direction of the interior normal to $\dot{V}$.

In addition we assume that all critical points of f in V are nondegenerate of order 2, and that the same is true of the points $x_1, x_2, \cdots, x_r$ considered as critical points of the "boundary function" ϕ, i.e. of f restricted to V. (The critical points of ϕ are the zeros of g_t [20, Lemma 6.3].) Finally, it is assumed that $H_q(V)$ is finitely generated.

THEOREM 3.7. *Under the above assumptions the Morse relations* (1.7) *hold if, in these relations, M_q is replaced by $M_q{}^1 = M_q + M_q{}^-$ where $M_q{}^-$ denotes the number of those of the points $x_1, x_2, \cdots, x_r$ defined above which are of index q.*

As in the finite-dimensional case [11], the proof consists in reducing the present case (b) to case (a) by constructing a modification f_1 of f which satisfies regular boundary conditions.

In fact it can be proved [20] that

there exists an open set V_1 with $\bar{V} \subset V_1 \subset \bar{V}_1 \subset V_2$, and a function f_1 of the following properties:

(i) $f_1(x) = f(x)$ for $x \in \bar{V}$;

(ii) f satisfies regular boundary conditions on $\dot{V}_1$, i.e. $\langle n_1(x), g_1(x)\rangle > 0$ for $x \in \dot{V}_1$ where $n_1(x)$ is the exterior unit normal to $\dot{V}_1$, and where $g_1 = \operatorname{grad} f_1$;

(iii) to each of the points $x_\rho \in \dot{V}$ defined above corresponds a unique point $x_\rho{}^* \in V_1 - \bar{V}$ which is a critical point of f_1, and these points $x_\rho{}^*$ are the only critical points of f_1 in $V_1 - \bar{V}$;

(iv) each of the critical points $x_\rho{}^*$ is nondegenerate, and the index of $x_\rho{}^*$ equals the index of x_ρ as a critical point of the boundary of function ϕ.

(v) $H_q(V) \approx H_q(V_1)$.

It is obvious from these properties that the function f_1 with domain $\bar{V}_1$ satisfies the assumptions of Theorem 3.6. Therefore by this theorem the relations (1.7) hold if the Morse numbers M_q of f are replaced by the Morse numbers $M_q{}^1$ of f_1. Theorem 3.7 now follows from properties (iv) and (v) above together with Theorem 3.4(γ).

Generalization of case (c) (see the beginning of §3). Here we confine ourselves to a short description of some of the methods and results of Palais' paper [13] frequently referring to analogies with the above treatment of case (a). Let V be a "complete Riemannian-Hilbert manifold" and let f be a real valued function with domain V. The term "Hilbert manifold" refers to the fact that the parameter spaces U which parametrize small enough open sets $D \subset V$ can be chosen to be (separable) Hilbert spaces. For such a fixed representation (or chart) $x = \phi(u)$ where $x \in D$, $u \in U$, differentials of $f \circ \phi(u)$ can be defined. Since the first differential is invariant under change of parameters the notions of critical point and gradient can be defined as in the Introduction. The second differential is not invariant in general. However it is invariant at points which correspond to critical points. Thus nondegeneracy of a critical point can be defined analogously to Definition 2.5.

It is supposed that all critical points are nondegenerate, that condition C (Definition 3.1) holds, and that f is at least three times differentiable. Under these assumptions our previous Assumption (D) that the critical values are isolated is automatically true, and each critical level contains only a finite number of critical points [13, p. 314]. Moreover the analogue to Lemma 3.1 holds [13, p. 310]. Consequently the critical group $C_q(c)$ at the critical level c can again be defined by (3.1). If $M_q(c)$ denotes the number of those critical points at level c whose index is finite then $C_q(c)$ equals the direct sum of the coefficient group taken $M_q(c)$ times [13, p. 336]. (As to the reason for the fact that the critical points of infinite index do not contribute to $C_q(c)$, see the last paragraph of the proof for Theorem 2.1.)

Let X be a topological space, B a ball, S its boundary and g a map: $S \to X$. Then the space obtained from X by "attaching B with attaching map g" is by definition constructed as follows: Take the topological sum of X and B and identify y in S with $g(y)$ in X.

If a, b, c are as in the definition of $C_q(c)$ (equation (3.1)) then the following relation holds between the sets $\bar{f_b}$ and $\bar{f_a}$: Let $j_1, j_2, \cdots, j_r$ denote the indices of those critical points at level c whose index is finite and let $f_a{}^+$ denote the space $\bar{f_a}$ with r balls of dimensions $j_1, j_2, \cdots, j_r$ disjointly attached. Then $f_a{}^+$ is a deformation retract of $\bar{f_b}$ [13, p. 336]. (Cf. [9, Theorem 3.2] for the finite-dimensional case.)

Turning to the Morse relations (1.7) we remark first that under the present conditions the number of critical levels is not necessarily finite. However since these levels are isolated there can be only a finite number in any given finite interval $[a, b]$, and since there are only a finite number of critical points at each critical level, the number $M_q(a, b)$ of critical points at levels in $[a, b]$ is finite.

Suppose now that a and b are not critical levels. Then the Morse relations (1.7) hold if the M_q are replaced by $M_q(a, b)$ and the R_q by the Betti numbers $R_q(a, b)$ of the couple $(\bar{f}_b, \bar{f}_a)$ [13, p. 338].

Case (d), *the problem of geodesics on a complete finite-dimensional Riemannian manifold* $\tilde{V}$. Let p and q be fixed points on $\tilde{V}$, and let $(V; p, q)$ denote the set of all curves $x = x(t)$ $(0 \leqq t \leqq 1)$ on $\tilde{V}$ joining p and q. Then, under certain smoothness assumptions on the manifold $\tilde{V}$ and the curves $x(t)$, the "loop space" $(\tilde{V}; p, q)$ can be interpreted as a complete Riemannian-Hilbert manifold V ([4] and [13, §13]). Let now the function f on V be defined by

$$f(x) = \int_0^1 \|x'(t)\|^2 \, dt. \tag{3.14}$$

Then it can be shown that the assumptions for the validity of the theory described in the preceding case (c), in particular condition C, are satisfied [13, §14].

Moreover the point $x \in V$ is a critical point of f if and only if the curve x on $\tilde{V}$ is a geodesic parametrized proportionally to arc length [13, p. 330, Corollary]. Thus the existence problem for geodesics on V is identical with the existence problem for critical points on the Hilbert manifold V of the function f defined by (3.14).

References

1. P. S. Alexandroff and H. Hopf, *Topologie*, Springer, Berlin, 1935.

2. J. Dieudonné, *Foundations of modern analysis*, Pure and Appl. Math., vol. 10, Academic Press, New York, 1960. MR **22** #11074.

3. J. Dugundji, *An extension of Tietze's theorem*, Pacific J. Math. **1** (1951), 353–367. MR **13**, 373.

4. J. Eells, Jr., *On the geometry of function spaces*, Sympos. internacional de topología algebraica, Universidad Nacional Autónoma de México and UNESCO, Mexico City, 1958, pp. 303–308. MR **20** #4878.

5. S. Eilenberg and N. E. Steenrod, *Foundations of algebraic topology*, Princeton Univ. Press, Princeton, N. J. 1952. MR **14**, 398.

6. M. R. Hestenes, *Applications of the theory of quadratic forms in Hilbert space to the calculus of variations*, Pacific J. Math. **1** (1951), 525–581. MR **13**, 759.

7. S. Kakutani, *Topological properties of the unit sphere of a Hilbert space*, Proc. Imp. Acad. Tokyo **19** (1943), 269–271. MR **7**, 252.

7a. N. H. Kuiper, *C^r-equivalence of functions near isolated critical points*, Ann. of Math. Studies, no. 69, Princeton Univ. Press, Princeton, N. J., 1972, pp. 199–218.

8. J. Leray and J. Schauder, *Topologie et équations fonctionnelles*, Ann. Sci. École Norm. Sup. (3) **51** (1934), 45–78.

9. J. Milnor, *Morse theory*, Ann. of Math. Studies, no. 51, Princeton Univ. Press, Princeton, N. J., 1963. MR **29** #634.

10. M. Morse, *Calculus of variations in the large*, Amer. Math. Soc. Colloq. Publ., vol. 18, Amer. Math. Soc., Providence, R. I., 1934.

11. M. Morse and G. B. Van Schaack, *The critical point theory under general boundary conditions,* Ann. of Math. (2) **35** (1934), 545–571.

12. R. Nevanlinna, *Über metrische lineare Räume.* III. *Theorie der Orthogonalsysteme,* Ann. Acad. Sci. Fenn. Ser. A I Math.-Phys. No. 115 (1952), 27 pp. MR **14**, 658.

13. R. S. Palais, *Morse theory on Hilbert manifolds,* Topology **2** (1963), 299–340. MR **28** #1633.

14. E. Pitcher, *Inequalities of critical point theory,* Bull. Amer. Math. Soc. **64** (1958), 1–30. MR **20** #2648.

15. E. H. Rothe, *The theory of topological order in some linear topological spaces,* Iowa State Coll. J. Sci. **13** (1939), 373–390. MR **1**, 108.

16. ———, *Leray-Schauder index and Morse type numbers in Hilbert space,* Ann. of Math. (2) **55** (1952), 433–467. MR **14**, 185.

17. ———, *Correction to the paper "Leray-Schauder index and Morse type numbers in Hilbert space",* Ann. of Math. (2) **58** (1953), 593–594. MR **15**, 236.

18. ———, *A remark on isolated critical points,* Amer. J. Math. **74** (1952), 253–263. MR **13**, 755.

19. ———, *Some remarks on critical point theory in Hilbert space,* Proc. Sympos. Nonlinear Problems (Madison, Wis., 1962), Univ. of Wisconsin Press, Madison, Wis., 1963, pp. 233–256. MR **28** #2424.

20. ———, *Critical point theory in Hilbert space under general boundary conditions,* J. Math. Anal. Appl. **11** (1965), 357–409. MR **32** #8361.

21. ———, *Critical point theory in Hilbert space under regular boundary conditions,* J. Math. Anal. Appl. **36** (1971), 377–431.

22. H. Seifert and W. Threlfall, *Variationsrechnung im Grossen,* Teubner, Leipzig, 1938.

23. M. M. Vaĭnberg, *Variational methods for the study of non-linear operators,* GITTL, Moscow, 1956; English transl., Holden-Day, San Francisco, Calif., 1964. MR **19**, 567; MR 31 #638.

University of Michigan, Ann Arbor, Michigan 48104