ON DIFFERENTIABLE FUNCTIONS WITH ISOLATED CRITICAL POINTS

DETLEF GROMOLL† and WOLGANG MEYER

(Received 12 December 1968)

The purpose of this paper is to describe some quantitative aspects of a Morse theory for differentiable functions on a manifold which have only isolated but possibly degenerate critical points. We will show that around such points locally, a function splits into degenerate and non-degenerate parts, in terms of which the relative homology can be expressed in a certain way when passing a critical level. Our investigation originated in connection with a specific geometric problem, namely to prove the existence of infinitely many geometrically distinct periodic geodesics for a very large class of compact riemannian manifolds, see [3], but the results of this paper may be useful for other applications of Morse theory as well. We wish to thank Ralph Abraham and Alan Weinstein for helpful conversations.

To fix some notations and assumptions, we consider a riemannian manifold \( M \) modelled on a separable real Hilbert space \( H \), and a differentiable function \( f : M \to \mathbb{R} \), all data will be sufficiently smooth. Let \( f \) have only isolated critical points. From §2 on, we need the condition (C) of Palais and Smale for \( f \), then the critical values of \( f \) are isolated and \( f^{-1}[a,b] \) contains only finitely many critical points for arbitrary real numbers \( a < b \). Let \( \nabla f \) denote the gradient vector field of \( f \), \( H_p \) the hessian form of \( f \) on the tangent space \( M_p \) at a critical point \( p \in M \), which gives rise to a selfadjoint operator \( A \) on \( M_p \) defined by

\[
\langle Ax, y \rangle = H_p(x,y).
\]

We restrict ourselves to the case when \( A \) admits a decomposition \( A = I + k \) with the identity \( I \) and a compact operator \( k \). Then \( A \) is a Fredholm operator of index 0, the null space \( \ker A \) of \( H_p \) and coker \( A \) are finite dimensional and of same dimension, \( A \) is injective iff \( A \) surjective iff \( A \) bijective, as in finite dimensions. We refer to Palais [7] for all basic definitions.

§1. NORMALIZATION OF \( f \) IN A NEIGHBORHOOD OF A CRITICAL POINT

In non-degenerate Morse theory the so-called Morse lemma, which introduces special coordinates around a critical point or manifold, is the important starting point and key tool (see [5] p. 6, [7] p. 307, [4] p. 55). We are going to generalize this fundamental lemma to some extent and derive a "splitting lemma" in the case of arbitrary critical points.

† Miller Fellow.
It suffices to look at the special situation when $M$ a neighborhood of the origin $0$ in $H$, $0$ the only critical point of $f$, $f(0) = 0$. Let the operator $A$ be defined as in (1) with respect to $0$. For $N := \ker A$ we have the orthogonal complement $E$ of $N$ so that $H = E \oplus N$.

**Lemma 1.** Let $f$ satisfy the assumptions as above. Then there exists an origin preserving diffeomorphism $\Phi$ of some neighborhood of $0$ in $H$ into $H$ and an origin preserving differentiable map $h$ defined in some neighborhood of $0$ in $N$ into $E$ such that

$$f \circ \Phi(x,y) = \|Px\|^2 - \|(I - P)x\|^2 + f(h(y),y)$$

with an orthogonal projection $P : E \to E$.

**Note.** $y - f(h(y),y)$ has vanishing hessian at the isolated critical point $0$. It is also remarkable that $\Phi$ keeps the $y$-coordinate fixed, furthermore $h_0 = 0$ and $\Phi_0 | N = id_N$. If $0$ is not an isolated critical point we still obtain the same splitting formula (2), in that case all critical points of $f \circ \Phi$ lie on the subspace $N$ of $H$.

**Proof.** Define $\varphi : E \oplus N \to E \oplus N$ by $\varphi(x,y) = P_1 Vf_{(x,y)} \oplus y$, $P_1 : H \to E$ orthogonal projection, $\varphi(0) = 0$. Since $D(P_1 Vf_{(x,y)}) = P_1 D Vf_{(x,y)} = P_1 A$, the differential of $\varphi$ at $0$ has the form $D\varphi_{(0)} = P_1 A \oplus id_N$. $A$ is selfadjoint with kernel $N$, so $A = P_1 AP_1$. It follows now from $\dim \ker A = \dim \coker A$ that $A$ is invertible.

Let $D_E$ denote the partial differential for functions on $H$ with respect to the $E$-component. For $x \in E$ and $y \in N$ we define a continuous bilinear form $B_{xy}$ on $E$ by

$$B_{xy}(x_1,x_2) = \int_0^1 (1 - t)D_E^2(g \circ \psi^{-1})_{(tx,y)}(x_1,x_2) dt$$

and an operator $A_{xy} : E \to E$ by $\langle A_{xy} x_1,x_2 \rangle = B_{xy}(x_1,x_2)$, clearly $A_{00} = P_1 A | E$ is invertible. Using Taylor’s expansion we have

$$g \circ \psi^{-1}(x,y) = B_{xy}(x,y).$$

The remainder of the proof is a modification of Palais’ proof of the Morse lemma, compare [7] p. 307 and also (4) p. 55. Consider operators $C_{xy} := (A_{xy}^{-1} A_{00})^{1/2}$ in a neighborhood of $0$ in $H$. One checks that $\Psi$ given by $\Psi(x,y) = (C_{xy}^{-1} x) \oplus y$ is a local diffeomorphism and $g \circ \Psi^{-1}(x,y) = \langle A_{00} x, y \rangle$, so

$$f \circ \Psi^{-1}(x,y) = \langle A_{00} x, x \rangle + f(h(y),y).$$

Using operator calculus the first term of the right-hand side can be treated as in (7) in order to obtain the result: Let $T = |A_{00}|^{-1/2}$ so that $A_{00} T^2 = P - (I - P)$, where $P = \chi(A)$ and $\chi$ the characteristic function of $[0,\infty)$, then $\Phi$ defined by $\Phi(x,y) = \Psi^{-1}(TX,y)$ is the desired local diffeomorphism.

† Recently Mather has observed that given any two splittings of $f$ according to (2), then the corresponding non-degenerate and degenerate parts are equivalent up to local diffeomorphisms.
As a first application the splitting lemma helps to resolve degenerate critical points in a partially controlled way into finitely many non-degenerate critical points. Recall that the index of a critical point $p$ is the index of the hessian $H_p$, i.e. the supremum of the dimensions of all linear spaces where $H_p$ negative definite. Let $f$ be a function as in Lemma 1, then the index of the critical point $0$ is given by $\dim(I - P)H$. We assume, $f$ satisfies condition (C) of Palais [7], p. 313, and $0$ has finite index $\lambda \geq 0$. The function $y \rightarrow f(h(y),y)$ is defined in a neighborhood of $0$ in the finite dimensional vector space $N$ and has $0$ as an isolated completely degenerate critical point. Essentially by Sard's theorem the linearly perturbed function $y \rightarrow f(h(y),y) + \langle a, y \rangle$ is non-degenerate for almost every $a \in N$. Denote by $B'_p$, $B_p$ the open $p$-balls about the origins in $E$, $N$ such that $f \circ \Phi$ is defined on $B'_p \times B_p$. Consider differentiable functions $\varphi: R \rightarrow R$, $\varphi(t) = 1$ for $t \leq p/2$, $\varphi(t) = 0$ for $t \geq p$, and $\alpha: E \rightarrow R$, $\beta: N \rightarrow R$, $\alpha(x) = \varphi(\|x\|)$, $\beta(y) = \varphi(\|y\|)$. Then for almost every sufficiently small $a \in N$ the function

$$f: (x,y) \rightarrow f \circ \Phi(x,y) + \alpha(x)\beta(y)\langle a, y \rangle$$

coincides with $f \circ \Phi$ outside $B'_p \times B_p$ and has only finitely many non-degenerate critical points which are all contained in $0 \times B_{p/2}$. These points correspond to the non-degenerate critical points $y_1, \ldots, y_k$ of the function $y \rightarrow f(h(y),y) + \langle a, y \rangle$, where $y_i$ has some index $\lambda_i$. The critical points of $f$ are precisely $(0,y_i)$ with indices $\lambda_i + \lambda_i (i = 1, \ldots, k)$, this follows from the Morse lemma and (2). It seems to be an interesting problem whether $f$ may always be approximated arbitrarily close by a non-degenerate function without creating unbounded numbers of critical points. In the analytic case one can use approximations of the type (3), which do not seem to apply in general.

We return to our original situation and assume throughout the following, condition (C) holds for the function $f$. Let $M$ be complete and all critical points have finite index. Then the last conclusions lead to a finiteness result and some estimates for the homology $H_*(M^b, M^a)$, $M^a = f^{-1}(-\infty, a]$. This was known from Morse (6) in finite dimensions, where the splitting lemma is not needed for such purposes.

Let $a < b$ be regular values of $f$ and $p_1, \ldots, p_m$ the critical points of $f$ in $f^{-1}[a,b]$ with indices $\lambda_1, \ldots, \lambda_m$. By the above approximation method we may resolve each $p_i$ into finitely many non-degenerate critical points $p_{i1}, \ldots, p_{ik_i}$ of indices $\lambda_{i1} + \lambda_i, \ldots, \lambda_{ik_i} + \lambda_i$. Of course, $0 \leq \lambda_{ij} \leq \mu_i := \dim \ker A_i$, where the operator $A_i$ corresponds to the hessian $H_p$, according to (1). Set $\lambda := \min \lambda_i$, $\mu := \max \mu_i$. Now from non-degenerate Morse theory, compare (7), we obtain:

**Lemma 2.** $M^b$ is diffeomorphic to $M^a$ with a handle of index $\lambda \leq \lambda_{ij} + \lambda_i \leq \mu + \lambda$ attached for each critical point $p_{ij}$, $1 \leq i \leq m$, $1 \leq j \leq k_i$. The diffeomorphism can be chosen to keep $M^a$ fixed. In particular, $H_*(M^b, M^a)$ is of finite type, $H_*(M^b, M^a) = 0$ except possibly when $\lambda \leq n \leq \mu + \lambda$.

We should mention that by this type of arguments for example, general deformation problems on the loop space level of a riemannian manifold can often be treated more directly without analysing degeneracy questions.
§2. LOCAL INVARIANTS

We will associate to a critical point $p$ of $f$ homological invariants which do only depend on suitable arbitrarily small neighborhoods of $p$, assume $f(p) = 0$ after normalization. Let $g$ be a real valued differentiable function defined in some open neighborhood of $p$, $g \geq 0$, $g(p) = 0$, and $p$ non-degenerate minimum of $g$. Such a $g$ will be called a cut function at $p$. We identify $M_p$ with $\mathbb{H}$ and choose a local parametrization $\Psi$ of $M$ around $p$ defined on a neighborhood of 0 in $\mathbb{H}$, $\Psi(0) = p$, according to Lemma 1,

$$g \circ \Psi(x) = \|x\|^2$$

in some open ball $B_\rho$ of radius $\rho$ about the origin in $\mathbb{H}$.

Start with $\rho > 0$ small such that $\|\nabla g\| < \alpha$ on $U := \Psi(B_\rho)$ and $\|\Psi^{-1} \nabla f\|$ bounded away from zero on $B_\rho - B$, for all $0 < \tau < \rho$. Fix numbers $0 < \rho_0 < \rho_1 < \rho$,

$$0 < \delta < \frac{\beta \rho_1^2 - \rho_0^2}{2\alpha}, \quad \text{where} \quad \|\nabla f|_q\| \geq \beta > 0$$

for all $q \in g^{-1}[\rho_0^2, \rho_1^2]$, and consider a function $d$ on $U$, $d(q) := \frac{\rho_1^2 - \rho_0^2}{2\delta} f(q) + g(q)$.

Now we define admissible regions

$$W: = f^{-1}[-\delta, \delta] \cap d^{-1}(-\infty, \frac{\rho_1^2 + \rho_0^2}{2}], \quad W^-: = f^{-1}(-\delta) \cap W,$$

note that $W^-$ may be void. Set

$$\mathcal{H}(f, p) := H_*(W, W^-), \quad \mathcal{H}_k(f, p) = H_k(W, W^-),$$

we use singular homology with arbitrary coefficients. It will be shown, $\mathcal{H}(f, p)$ does not depend on the choices involved in the definition of $W$, so we obtain a local homological invariant of the function $f$ at any critical point $p$. In fact $\mathcal{H}(f, p)$ was already defined by Morse in [6] for finite dimensions, however, in view of some applications our modified construction seems to be useful.

**Proposition.** $\mathcal{H}(f, p)$ is well defined.

**Proof.** Let $\tilde{W}$ be another admissible region with corresponding data $\tilde{g}$, $\tilde{\Psi}$, $\tilde{U}$, $0 < \tilde{\rho}_0 < \tilde{\rho}_1 < \tilde{\rho}$, $0 < \tilde{\delta}$, $\tilde{\beta}$. There is an admissible region $\tilde{W}$ with $\tilde{U} \subset W \cap \tilde{W}$, hence it suffices to assume $\tilde{W} \subset \tilde{U} \subset W$. We have the vector field $X := -\nabla f/\|\nabla f\|^2$ on $U - \{p\}$ and the maximal flow $\Delta: U_0 \rightarrow U$ of $X$, $U_0 \subset \mathbb{R} \times U$ open neighborhood of $0 \times U$, $\Delta_q: J_q \rightarrow U$ maximal integral curve of $X$ with $\Delta_q(0) = q$, $J_q \times q = (\mathbb{R} \times q) \cap U_0$, $\Delta_q(t) = \Delta(t, q)$. Setting

$$V_1 := (f^{-1}[-\delta, \delta] \cap W) \cup \left(f^{-1}[-\delta, -\delta] \cap d^{-1}\left(\frac{\rho_1^2 + \rho_0^2}{2}\right)\right), \quad V_1^- := f^{-1}[-\delta, -\delta] \cap V_1,$$

we show that

$$H_*(W, W^-) = H_*(V_1, V_1^-).$$

Put $V_2 := f^{-1}[-\delta, -\delta] \cap W$, then $[0, \delta + f(q)] \subset J_q$, $\Delta_q[0, \delta + f(q)] \subset V_2$, and $\Delta_q(\delta + f(q)) \in W^-$ for $q \in V_2$. Observe that $\Delta_q(t) \in d^{-1}\left(\frac{\rho_1^2 + \rho_0^2}{2}\right)$, $0 \leq t \leq \delta + f(q)$, at
most if \( t = 0 \), for \((f \circ \Delta_q)' = -1\), \((d \circ \Delta_q)' < 0\) because \((d \circ \Delta_q)' = -\langle \nabla d, \nabla f/\|\nabla f\|^2 \rangle \circ \Delta_q\) and \(\langle \nabla d, \nabla f/\|\nabla f\|^2 \rangle = \frac{\rho_1^2 - \rho_0^2}{2\delta} - \frac{\|\nabla g\|}{\|\nabla f\|} > 0\) by (5).

Now from our choice of \( \rho \) and the completeness of \( H \) we get \([0, \delta + f(q)] \subset J_q\). The deformation \( V_2 \times [0,1] \to V_2\), \((q,s) \to s(\delta + f(q))\), defines a retract \( V_2 \to W^- \) for \( s = 1 \), hence the inclusion \((W, W^-) \subset (V_1, V_2)\) induces an isomorphism

\[
H_*(W, W^-) = H_*(V_1, V_2).
\]

We construct another deformation \((W, V_2) \times [0,1] \to (W, V_2)\) which provides a retract \((W, V_2) \to (V_1, V_1^-)\) and an isomorphism \(H_*(W, V_2) = H_*(V_1, V_1^-)\). Combining this with (8) we obtain (8).

For \( q \in V_2 \) there is exactly one \( 0 \leq t_q \in J_q \) with \( \Delta_q(t_q) \in V_1^- \). Consider again the functions \( f \circ \Delta_q, d \circ \Delta_q \). To prove the existence of \( t_q \), one may use the intermediate value theorem, uniqueness follows by a monotony argument as above. Now the map \( q \to t_q \) is continuous on \( V_2 \) : Given \( q \) we find \( 0 > \tau \in J_q \) such that \( \tau \in J_q \) and \( \Delta_q(\tau) \in U - V_2 \) for all \( \tilde{q} \) in some neighborhood of \( q \) in \( U \). Suppose \( q \to q_v \), then \( 0 \geq t_{q_v} \geq \tau \) for almost all \( v \) and \( 0 \geq t'_{q_v} \geq \tau \), \( t'_{q_v} \in J_q \), when \( t'_{q_v} \) a limit point of \( t_{q_v} \), hence \( \Delta(t'_{q_v}, q) \in V_1^- \) and \( t'_{q_v} = t_{q_v} \) by uniqueness. Set \( t_q := 0 \) for \( q \in V_1^- \), the map \((q, s) \to \Delta(st_q, q)\) furnishes the desired deformation.

For technical reasons we introduce another admissible region \( \tilde{W} \) with the same data as \( W \) except that \( \tilde{\rho}_1 \) satisfies \( \tilde{\rho}_1 > \rho_1 < \tilde{\rho}_1 \), in particular \( \tilde{\delta} = \delta \). Using essentially arguments as before we will deform \((V_1, V_1^-)\) onto \((\tilde{W} \cup V_1^-, V_1^-)\) and \((\tilde{W}, \tilde{W}^-)\) onto \((\tilde{W} \cup \tilde{W}^-, \tilde{W}^-)\) at the same time, so one has isomorphisms

\[
H_*(\tilde{W} \cup V_1^-, V_1^-) = H_*(V_1, V_1^-), \quad H_*(\tilde{W} \cup \tilde{W}^-, \tilde{W}^-) = H_*(\tilde{W}, \tilde{W}^-).
\]

Excision yields \(H_*(\tilde{W} \cup \tilde{W}^-, \tilde{W}^-) = H_*(\tilde{W} \cup V_1^-, V_1^-)\), hence

\[
H_*(\tilde{W}, \tilde{W}^-) = H_*(V_1, V_1^-).
\]

Clearly the assertion of the proposition is contained in (8) and (9).

It remains to construct the deformation \((V_1, V_1^-) \times [0,1] \to (V_1, V_1^-)\). Associate to \( q \in V_1^- \) the unique \( 0 \leq t_q \in J_q \) such that \( \Delta_q(t_q) \in \left( \tilde{W} \cap \left[ f^{-1}(\delta) \cup d^{-1}\left( \tilde{\rho}_1^2 + \tilde{\rho}_0^2 \right) \right] \right) \cup \left( V_1^- - \tilde{W} \right)\), put \( t_q := 0 \) for \( q \in \tilde{W} \), and let \((q, s) \to \Delta(st_q, q)\). This completes the proof.

We should point out that \( \mathcal{H}(f, p) \) is of finite type. This follows for example using the approximation methods of the preceding section. Furthermore, \( \mathcal{H}(f, p) \) is invariant under transformations of \( f \) by a local diffeomorphism \( \Psi \) about \( p \), \( \mathcal{H}(f, p) = \mathcal{H}(f \circ \Psi, \Psi^{-1}(p)) \), for the transform of a cut function at \( p \) will be a cut function at \( \Psi^{-1}(p) \). Of course, one could also define \( \mathcal{H}(f, p) \) for regular points \( p \) of \( f \) in a similar way, but then obviously \( \mathcal{H}(f, p) = 0 \). In the case of a complete manifold \( M \) we are now ready to derive the basic localization result.

**Lemma 3.** Let \( M \) be complete and \( c \) the only critical value of \( f \) in \([c - \varepsilon, c + \varepsilon]\) for some \( \varepsilon > 0 \). If \( p_1, \ldots, p_m \) are the critical points of \( f \) in \( f^{-1}(c) \), then

\[
H_*(M^{c+\varepsilon}, M^{c-\varepsilon}) = \sum_{i=1}^m \mathcal{H}(f, p_i).
\]
Proof. After normalization assume \( c = 0 \). Recall that \((M^4, M^{-4})\) is canonically diffeomorphic with \((M', M'^{-4})\) whenever \(0 < \delta \leq \varepsilon\). This follows using standard deformations along the gradient lines of \(f\). On the homology level, excision in connection with a deformation yields

\[
H_*(M^4, M^{-4}) = H_*(f^{-1}[-\delta, \delta], f^{-1}(-\delta)).
\]

When \(m = 1\) one may proceed as follows: Take an admissible region \(\tilde{W}\) at \(p = p_1\) with \(\delta = \delta < \varepsilon\), then \(\mathcal{H}(f, p) = H_*(\tilde{W}, \tilde{W}^-)\). On the other hand, \(H_*(\tilde{W}, \tilde{W}^-) = H_*(f^{-1}[-\delta, \delta], f^{-1}(-\delta))\) following the proof of the Proposition if one replaces \((W, W^-)\) by \((f^{-1}[-\delta, \delta], f^{-1}(-\delta))\). In general choose mutually disjoint admissible regions \(\tilde{W}_1, \ldots, \tilde{W}_m\) with \(\delta = \delta_1, \ldots, \delta_m < \varepsilon\), and the reader will easily modify the proof.

Next we will show that \(\mathcal{H}(f, p)\) is invariant under certain homotopies of \(f\).

**Lemma 4.** Let \(p \in M\) and \(f : M \times [0, 1] \to \mathbb{R}\) be differentiable such that the partial function \(f_t : M \to \mathbb{R}\) has \(p\) as the only critical point and satisfies condition (C) for all \(t \in [0, 1]\). Then \(\mathcal{H}(f_t, p) = \mathcal{H}(f_0, p), \ t \in [0, 1]\).

**Proof.** We may assume \(f_t(p) = 0\). It suffices to determine a neighborhood \(J\) for any \(t_0\) in \([0, 1]\) such that \(\mathcal{H}(f_t, p) = \mathcal{H}(f_0, p), \ t \in J\). Choose a cut function \(g\) at \(p\) and an admissible region \(W\) for \(f_0\) with data \(g, \rho_0, \delta\). Now fix numbers \(0 < \tilde{\rho}_0 < \tilde{\beta} < \rho_0, g^{-1}[0, \rho_0^2] \subset \text{int} \ W\), such that \(\frac{\partial f}{\partial t}\) and \(\frac{\partial}{\partial t} \nabla f_t\) are bounded on \(g^{-1}[0, \rho_0^2] \times [0, 1]\). This is possible since \(\frac{\partial f}{\partial t}\), \(\left\| \frac{\partial}{\partial t} \nabla f_t \right\|\) continuous and \([0, 1]\) compact. Let \(\varphi : \mathbb{R} \to \mathbb{R}\) be a differentiable function with \(\varphi(t) = 1\) for \(|t| \leq \tilde{\rho}_0^2\), \(\varphi(t) = 0\) for \(|t| \geq \tilde{\beta}^2\), and define \(f_t = (\varphi \circ g)f_t + (1 - \varphi \circ g)f_0\). Clearly \(f_t(q) = f_0(q)\) for \(q \in g^{-1}[0, \tilde{\beta}^2]\), \(f_t(q) = f_t(q)\) for \(q \notin g^{-1}[0, \rho_0^2]\). We find a neighborhood \(J\) of \(t_0\) such that condition (C) holds for \(f_t\), \(t \in J\). Note that condition (C) is already satisfied outside \(g^{-1}[\tilde{\rho}_0^2, \tilde{\beta}^2\]), on this set we consider \(V f_t = (f_t - f_0) \nabla (\varphi \circ g) + (1 - \varphi \circ g) \nabla f_0 + (\varphi \circ g) \nabla f_t\). From our choice of \(\tilde{\beta}\) it follows that \(\lim_{t \to t_0} f_t = f_0\) and \(\lim_{t \to t_0} \nabla f_t = \nabla f_0\) uniformly on \(g^{-1}[0, \rho_0^2]\). Since \(\| \nabla (\varphi \circ g) \|, \| \nabla f_0 \|\) bounded away from zero on \(g^{-1}[\tilde{\rho}_0^2, \tilde{\beta}^2]\), the same will hold for \(\| f_t \|\) if \(t\) sufficiently close to \(t_0\).

Finally we take an admissible region \(\tilde{W}_t\) for any function \(f_t, t \in J\), satisfying \(\tilde{W}_t \subset g^{-1}[0, \rho_0^2] \subset W\). Then \(\tilde{W}_t\) is also admissible for \(f_t\), moreover \(W\) is admissible for \(f_0\) and \(f_t\). Hence by the Proposition, we obtain \(\mathcal{H}(f_t, p) = \mathcal{H}(\tilde{f}_t, p) = \mathcal{H}(f_t, p), t \in J\).

§3. THE SHIFTING THEOREM

In this section we discuss the announced result that the local invariant \(\mathcal{H}(f, p)\) of \(f\) at the critical point \(p\) can be calculated in a simple way from an invariant \(\mathcal{H}^2(f, p)\) which measures only the degenerate behavior of \(f\). Let \(f_0(p) = 0\), according to the splitting Lemma 1, there exists a local parametrization \(\Phi\) of \(M\) defined in some open neighborhood \(B\) of \(0\) in \(M_p \cong H, \ \Phi(0) = p\), such that \(f \circ \Phi(x, y) = \|Px\|^2 - \|(I - P)x\|^2 + f_0(y), \ H = E \oplus N_p\) orthogonal direct sum, \(N_p\) null space of the hessian \(H_p, f_0 : B \cap N_p \to \mathbb{R}\) differentiable.
function with isolated completely degenerate critical point 0, \( \Phi_{\circ} \mid N_p = id_{N_p} \). Call \( N = \Phi(B \cap N_p) \) a characteristic submanifold of \( M \) for \( f \) at \( p \) with parametrization \( \Phi \), clearly \( N_p \) is the tangent space of \( N \) at \( p \) and the null space of the hessians of \( f \) and \( f \circ \Phi \) as well. Now \( \mathcal{H}(f \mid N, p) = \mathcal{H}(f_0, 0) \) is well defined.

**Theorem.** Let \( N \subset M \) be a characteristic submanifold for \( f \) at \( p \) and \( \lambda \geq 0 \) the index of \( p \). Then

\[
\mathcal{H}_{k+\lambda}(f, p) = \mathcal{H}_k(f \mid N, p)
\]

for all integers \( k \).

In particular, \( \mathcal{H}(f \mid N, p) = : \mathcal{H}^\circ(f, p) \) does not depend on the choice of \( N \), we will call \( \mathcal{H}^\circ(f, p) \) the characteristic invariant of \( f \) at \( p \). The shifting theorem is an immediate consequence of the following two lemmas.

**B** open ball centered about the origin in \( \mathbb{H}, f : B \to \mathbb{R} \) differentiable function satisfying condition (C), 0 isolated critical point of \( f, f(0) = 0 \).

**Lemma 5.** If \( \mathcal{H} = E^+ \oplus E \) and \( f(x, y) = \|x\|^2 + f_0(y) \), then

\[
\mathcal{H}(f, 0) = \mathcal{H}(f_0, 0).\]

**Proof.** In \( B \) we may use \( g(x, y) = \|x\|^2 + \|y\|^2 \) as a cut function at 0. Fix an admissible region \( W \subset B \) for \( f \) with data \( \delta, \rho_0, \rho_1 \). Clearly \( W_0 = W \cap E \) is an admissible region for \( f_0 \) since \( \forall f \) tangent to \( E \). We have only to check that \( H_\ast(W, W^-) = H_\ast(W_0, W_0^-) \). Now

\[
W = \{x \oplus y \mid -\delta \leq \|x\|^2 + f_0(y) \leq \delta, \frac{\rho_1^2 - \rho_0^2}{2\delta} (\|x\|^2 + f_0(y)) + \|x\|^2 + \|y\|^2 \leq \frac{\rho_1^2 + \rho_0^2}{2} \}.
\]

Define continuous real valued functions \( c_1, c_2 \) on \( B \cap E \), \( c_1(y) = \max\{0, -\delta - f_0(y)\} \), \( c_2(y) = \min\{\delta - f_0(y), \frac{2\delta}{\rho_1^2 - \rho_0^2 + 2\delta} \left( \frac{\rho_1^2 + \rho_0^2}{2} - \frac{\rho_1^2 - \rho_0^2}{2\delta} f_0(y) - \|y\|^2 \right) \} \), so \( W = \{x \oplus y \mid c_1(y) \leq \|x\|^2 \leq c_2(y)\} \).

The deformation \( W \times [0, 1] \to W \) given by \( (x, y, t) \to \left((1-t)x + t \sqrt{c_1(y)} \frac{x}{\|x\|^2}, y\right) \) for \( x \neq 0 \), \( (0, y, t) \to (0, y) \), provides a retract \( (W, W^-) \to (W_0 \cup W^-, W^-) \). Therefore \( H_\ast(W, W^-) = H_\ast(W_0 \cup W^-, W^-) \).

We would like to remove \( W^- \to W_0^- \), however, excision does not apply immediately. Note that \( H_\ast(W_0, W_0^-) = H_\ast(W_0, W_0 \cap f^{-1}[-\delta, -\delta']) \) for \( 0 < \delta' < \delta \). Here \( W_0 \cap f^{-1}[-\delta, -\delta'] \) can be deformed onto \( W_0^- \) along the gradient lines of \( f_0 \) as described several times in the proof of our proposition. This also yields a deformation of \( (W_0 \cap f^{-1}[-\delta, -\delta']) \cup W^- \) onto \( W^- \), we have \( H_\ast(W_0 \cup W^-) = H_\ast(W_0 \cup W^-) \). By excision \( H_\ast(W_0 \cup W^-) = H_\ast(W, W^-) = H_\ast(W_0, W_0^-) \).

**Lemma 6.** If \( \mathcal{H} = E^+ \oplus E \) and \( f(x, y) = -\|x\|^2 + f_0(y) \), then

\[
\mathcal{H}_{k+\lambda}(f, 0) = \mathcal{H}_k(f_0, 0) \quad \text{for} \quad \dim E^- = \lambda < \infty,
\]

\[
\mathcal{H}(f, 0) = 0 \quad \text{for} \quad \lambda = \infty.
\]

† A relation similar to (12) was obtained by Rothe in [8] under restrictive conditions on the function \( f \).
Proof. Let \( W \) be an admissible region for \( f \), so \( W_0 = W \cap E \) admissible for \( f_0 \). Denote by \( P \) the orthogonal projection \( W \rightarrow E \) and consider \( V = P^{-1}f^{-1}[\delta', \delta] \cap f^{-1}[-\delta, \delta] \) for some \( 0 < \delta' < \delta \). If \( \Delta \) the maximal flow of \( -\nabla f/\|\nabla f\|^2 \) and \( \Delta_q(0) = q \in W \), observe that \((f_0 \circ P \circ \Delta_q)' = -\|\nabla f_0\|^2 \circ P/\|\nabla f\|^2 \circ \Delta_q \). Then using the technique in the proof of our proposition, we can deform the pair \((W, W^-) \) onto \((P^{-1}W_0 - V) \cup W^-, W^- \) and at the same time \((P^{-1}W_0, P^{-1}W_0 \cap W^-) \) onto \((P^{-1}W_0 - V, P^{-1}W_0 \cap W^-) \). Furthermore we have \( H_\ast((P^{-1}W_0 - V) \cup W^-, W^-) = H_\ast(P^{-1}W_0 - V, P^{-1}W_0 \cap W^-) \) by excision. Therefore \( H_\ast(W, W^-) = H_\ast(P^{-1}W_0, P^{-1}W_0 \cap W^-) \).

The pair \((P^{-1}W_0, P^{-1}W_0 \cap W^-) \) is homotopy equivalent to \((W_0 \times D^1, W_0 \times \partial D^1 \cup W_0^- \times D^1) \). Now for \( \lambda < \infty \) and \( W_0^- = \emptyset \) our assertion is a consequence of the Thom isomorphism theorem. If \( \lambda < \infty \) and \( W_0^- \neq \emptyset \), we obtain by induction \( H_\ast(W_0 \times D^1, W_0 \times \partial D^1 \cup W_0^- \times D^1) = H_\ast(\Sigma^1 W_0, \Sigma^1 W_0^-) \), where \( \Sigma^1 \) denotes the \( \lambda \)-fold suspension without base point, but \( H_\ast(\Sigma^1 W_0, \Sigma^1 W_0^-) = H_\ast(W_0, W_0^-) \). A Künneth formula can be used as well. In the case \( \lambda = \infty \) we apply Bessaga’s theorem saying that \( D^\infty - 0 \) is diffeomorphic with \( D^\infty \), see [2]. Thus \((W_0 \times D^\infty, W_0 \times \partial D^\infty \cup W_0^- \times D^\infty) \) homeomorphic with \((W_0 \times (D^\infty - 0), W_0 \times \partial D^\infty \cup W_0^- \times (D^\infty - 0)) \), and this in turn is homotopy equivalent to \((W_0 \times \partial D^\infty, W_0 \times \partial D^\infty) \), which completes the proof.

Of course, in the shifting theorem we have \( \mathcal{H}(f, p) = 0 \) in general by (12) and (14) when the index of \( p \) is infinite, so critical points of this type are homologically invisible. A proof of (11) by approximation methods as in section 1 seems to be more complicated.

A characteristic submanifold \( N \) for \( f \) at \( p \) is by no means unique, we do not know how to determine all such \( N \) by reasonable conditions. There are submanifolds \( N \) of \( M \) such that \( N_p \) the null space of the hessian and \( p \) isolated critical point of \( f \mid N \), but \( \mathcal{H}(f \mid N, p) \) no local invariant of \( f, \mathcal{H}(f \mid N, p) \neq \mathcal{H}_0(f, p) \). The simplest example is furnished by the function \( f : \mathbb{R}^2 \rightarrow \mathbb{R} \) with \( f(x, y) = x^2 - y^4 \). Here the line \( x = 0 \) is characteristic, whereas the parabola \( x = 2y^2 \) is not. The problem looks rather delicate, the restriction of \( f \) even to \( C^\infty \)-approximations of a characteristic submanifold may not have an isolated critical point at \( p \). A sufficient condition for a submanifold \( N \subset M \) to be characteristic at \( p \) is that the hessian has nullity \( \dim N \) and the gradient \( \nabla f \) is tangent to \( N \). In fact a characteristic submanifold \( N \) with these properties always exists at least in finite dimensions, one may take the center manifold of the vector field \( \nabla f \) through \( p \), compare [1]. The following criterion, which generalizes the above condition, was important in [3].

Lemma 7. \( p \) isolated critical point of \( f \), \( M \) closed Hilbert submanifold of \( M \) such that \( \tilde{M}_p \) contains the null space of the hessian of \( f \). Assume, \( \nabla f_q, q \in \tilde{M} \) for all \( q \in \tilde{M} \). If \( N \subset \tilde{M} \) sufficiently small and characteristic for \( \tilde{f} := f \mid \tilde{M} \) at \( p \), then \( N \) is also characteristic for \( f \), in particular

(15) \[ \mathcal{H}_0(f, p) = \mathcal{H}_0(\tilde{f}, p). \]

Proof. There is a local parametrization \( \tilde{\Phi} \) of \( \tilde{M} \) defined in a neighborhood \( \tilde{B} \subset \tilde{M}_p = E \oplus N_p \) about the origin, \( \tilde{\Phi}(0) = p, \tilde{\Phi}(B \cap N_p) = N, \) and \( f \circ \tilde{\Phi}(x, y) = \|\tilde{P}x\|^2 - \|(I - \tilde{P})x\|^2 + f_0(y) \). We may assume, the normal bundle \( v(\tilde{M}) \) of \( \tilde{M} \) in \( M \) restricted to \( \tilde{\Phi}(\tilde{B}) \) is trivial.
Using the exponential map of $M$ on $v(\hat{M})$, one easily extends $\hat{\Phi}$ to a parametrization $\Psi$ of a neighborhood of $N$ in $M$ by some neighborhood $B \subset M_p = \hat{M}_p \perp \hat{M}_p$ such that $\Psi_{|\psi(0,y)} \hat{M}_p^\perp$ is the orthogonal complement of $\hat{M}_{\psi(0,y)}$.

Consider the splitting $M_p = E \oplus N_p$. Clearly $f \circ \Psi|E \oplus y$ has a critical point at $0 \oplus y$ for all $y \in B \cap N_p$ which is non-degenerate for $y = 0$. Following the explicit construction in the proof of Lemma 1 we find a local diffeomorphism $\Phi$ of $M_p$, $\Phi(0) = 0$, such that $f \circ \Psi \circ \Phi = \|Px\|^2 - \|(I - P)x\|^2 + f_0(y)$, observe that in our special situation $\Phi(0,y) = (0,y)$.

We conclude with some further remarks. We have not been interested in discussing the weakest order of differentiability necessary for our constructions, for example in the splitting lemma, if all given data are $C^r$, $r \geq 2$, then $\Phi$ is of class $C^{r-2}$. The admissible domains $W$ used for the definition of the local invariant $\mathcal{H}(f, p)$ are in fact smoothable topological manifolds with boundary for almost all sufficiently small data $\rho_0, \rho_1, \delta$. This follows using essentially transversality arguments, which in infinite dimensions depend heavily on the finiteness of the nullity. This will also imply the main assertions of Lemma 2.

An interesting question in connection with the shifting theorem is the following: Consider more generally functions $f_0: E_0 \to \mathbb{R}, f_1: E_1 \to \mathbb{R}$ with isolated critical points at the origins and $f: E_0 \oplus E_1 \to \mathbb{R}, f(x, y) = f_0(x) + f_1(y)$. Does the formula $\mathcal{H}(f, 0) = \mathcal{H}(f_0, 0) \otimes \mathcal{H}(f_1, 0)$ hold?

Under additional assumptions, many of our procedures may be extended to study differentiable functions with more complicated critical sets, say critical submanifolds. In particular, the case of isolated critical orbits with respect to the equivariant action of a compact Lie group, can be handled with these methods.

REFERENCES


University of California,
Berkeley

Universität Bonn