A relation between the type numbers of a critical point and the index of the corresponding field of gradient vectors.

ERHARD SCHMIDT zum 75. Geburtstag gewidmet.

By ERICH H. ROTHE of Ann Arbor, Mich. (USA.).

(Eingegangen am 6. 3. 1950.)

1. Introduction.

Let $I = I(x)$, be a scalar, i. e., a real valued function of the point $x = (x_1, x_2, \ldots, x_n)$ of the real Euclidean n-space E^n $(n \geqq 3)$. We assume I to have second continuous derivatives in some neighborhood U_1 of the origin $o = (0, 0, \ldots, 0)$ of E^n. Moreover, we assume that o is an isolated critical point of I, i. e., that

$$\text{grad } I = 0 \quad \text{for } x = o \tag{1.1}$$

while $\text{grad } I \neq 0$ for all $x \neq o$ of some neighborhood U of o whose closure $\overline{U}$ is contained in U_1. The vector field $g(x) = \text{grad } I(x)$ has then o as an isolated singularity. Let j be the index of this singularity[1]) and m^r $(r = 0, 1, \ldots, n)$ be the r-th (Morse) type number[2]) of the critical point o of I. The object of this paper is the proof of the relation

$$j = \sum_{r=0}^{n} (-1)^r m^r \tag{1.2}$$

under the following

Hypothesis H. *There exists a neighborhood U_2 of o such that for all $x \neq o$ of the intersection $\overline{U}_2 \cap \{x \mid I(x) = I(o)\}$ the vectors $x - o$ and* $\text{grad } I(x)$ *are linearly independent*[3]).

Regarding the validity of the hypothesis H we make the following remarks: H is vacuously satisfied if o is a maximum or minimum since then (for a small enough neighborhood of o) the set $\{x \mid I(x) = I(o)\}$ contains only o (cf. section 6). If o is an "intermediate" critical point, i. e., neither maximum nor minimum point, it has been proved by A. B. BROWN and M. MORSE[1]) that H is always satisfied if I is an analytic function. Moreover it will be shown in a separate paper[2]) that H is also satisfied under the following circumstances: I has continuous derivatives up to and including order $p + 1$ where p is an integer $\geqq 2$; all derivatives of order less than p vanish at $x = o$ while the homogeneous form of degree p giving the p-th differential at $x = o$ is not degenerate in the algebraic sense[3]).

The proof of (1.2) is straightforward in the case of a maximum or minimum (section 6). The major part of the paper deals therefore with the proof in the "intermediate" case which is based on a geometric discussion of the "cylindrical neighborhood" of the critical point o as defined by SEIFERT and THRELFALL[4]) (sections 2 and 3), on the LEFSCHETZ fixed point formula[5]), and on HOPF's extension theorem[6]). To facilitate the exposition we give a rough outline of the proof in this case: without loss of generality we assume that

$$I(o) = 0 \tag{1.3}$$

and consider a small enough "cylindrical" neighborhood $C(\varepsilon, \varepsilon_1)$ (cf. section 2) of o where $\varepsilon < \varepsilon_1$ are positive numbers. The boundary ∂ of $C(\varepsilon, \varepsilon_1)$ turns out to be a simply closed connected complex consisting of a finite number of differentiable $(n-1)$-manifolds. In addition the intersection γ of ∂ with the level surface $I = -\varepsilon_1$ consists of a finite number of components $\gamma_1, \ldots, \gamma_\beta$ each of which is a differentiable $(n-1)$-manifold. We define for $i = 1, 2, \ldots, \beta$ subsets Γ_i of $C(\varepsilon, \varepsilon_1)$ as follows: through each point of γ_i we draw a "gradient line", i. e., a curve to which the vectors of the gradient field $g(x) = \text{grad } I(x)$ are tangential; Γ_i consists then of those points x of such a gradient line for which $-\varepsilon \geqq I(x) \geqq -\varepsilon_1$. In order to apply the Lefschetz fixed point formula we construct in section 4 a continuous map $f(x)$ of $C(\varepsilon, \varepsilon_1)$ into itself with the following properties: f is homotopic to the identity map; $f(x)$ has $\beta + 1$ fixed points namely o, and exactly one interior point o_i of Γ_i; f maps each Γ_i into itself; finally the index of the fixed point o of $f(x)$ is $(-1)^n$ times the index j of the singularity o of the gradient field $g(x)$. Since through each $x \neq o$ of $C(\varepsilon, \varepsilon_1)$ there passes exactly one gradient line $f(x)$ is easy to define in $C(\varepsilon, \varepsilon_1) - \sum_{i=1}^{\beta} \Gamma_i$ by displacing x by a proper amount along the gradient line through x in direction of decreasing I, and by setting $f(o) = o$. However in order to obtain a map of $C(\varepsilon, \varepsilon_1)$ into itself the definition has to be different in the set $\sum_{i=1}^{\beta} \Gamma_i$ which is "near" to the level surface $I = -\varepsilon_1$. It is here that we make essential use of the Hopf extension theorem (section 4).

The application of the Lefschetz theorem which states that for a map of a polyhedron into itself the algebraic sum of the fixed point indices is $(-1)^n$ times

---

1) For the definition of the index of the singularity of a vector field see [1], Kapitel XIV, 2 (Numbers in brackets refer to the bibliography p. 27).

2) [4], where also further literature about the critical point theory is to be found. Except for the coefficient domain, the definition used in this paper is the one given in [6], p. 29. It is stated explicitly at the beginning of section 5 of the present paper. For its agreement with the definition given in [2], p. 267 see p. 26, footnote 1 of the present paper.

3) For any point set U the closure is denoted by $\overline{U}$. The symbol $\cap$ denotes intersection. The symbol $\{y \mid P(y)\}$ denotes the set of all y having the property P. Sometimes we will shortly write $\{P\}$, e. g., $\{I \leqq 0\} = \{y \mid I(y) \leqq 0\}$.

1) [2], lemma 10; [4], p. 156, theorem 4.3.

2) E. H. ROTHE, A remark on isolated critical points. To appear in the Amer. J. Math.

3) If $p = 2$, o is obviously a non degenerate critical point in the sense of Morse ([4], p. 143). That in this case H is satisfied is shown in [4], p. 155, theorem 4.2. Moreover for a non degenerate critical point formula (1.2) (even in Hilbert space) was established in [5].

4) [6], p. 38.

5) See e.g. [1], Kapitel 14, especially p. 542, Satz Ia.

6) [1], p. 500.

the "Lefschetz number" of the map, is now as follows: since f is homotopic to the identity the Lefschetz number is the Euler characteristic. Applying the theorem to $C(\varepsilon, \varepsilon_1)$ and to each Γ_i [1]) we obtain the formulas

$$(1.4)\qquad (-1)^n \chi(C(\varepsilon, \varepsilon_1)) = (-1)^n j + \sum_{i=1}^{\beta} j_i, \quad (-1)^n \chi(\Gamma_i) = j_i \quad (i = 1, 2, \dots, \beta)$$

if j_i is the index of the fixed point $o_i \subset \Gamma_i$ of the map f, and if $\chi(P)$ denotes the Euler characteristic of the polyhedron P. Now $\chi(C(\varepsilon, \varepsilon_1)) = 1$ since $C(\varepsilon, \varepsilon_1)$ is contractible to o (lemma 3.7), and

$$\sum_{i=1}^{\beta} \chi(\Gamma_i) = \chi\left(\sum_{i=1}^{\beta} \Gamma_i\right) = \sum_{r=0}^{n-1} (-1)^r p^r$$

where p^r is the r-th Betti number of $\sum_{i=1}^{\beta} \Gamma_i$ [2]). This together with (1.4) gives

$$(1.5)\qquad j = -(p_0 - 1) + \sum_{r=2}^{n} (-1)^r p^{r-1}.$$

Now the r-th type number m^r is defined as the r-th Betti number of the (singular) complex $K = C(\varepsilon, \varepsilon_1) \cap \{I \leqq 0\}$ modulo $K - o$ (cf. section 5). From this definition the following relations are easily established in the intermediate case (section 5):

$$(1.6)\qquad m^0 = 0, \quad m^1 = p^0 - 1, \quad m^r = p^{r-1} \quad (r \geqq 2).$$

Combining (1.5) with (1.6) we obtain (1.2).

In all that follows the assumptions of this introduction are supposed to hold with the exception that the hypothesis H is not required for sections 2 and 6.

2. The cylindrical neighborhood $C(\varepsilon, \varepsilon_1)$ [3]).

Through each point $x^0 = (x_1^0, x_2^0, \dots, x_n^0) \neq o$ of the neighborhood U of the introduction passes one and only one gradient line, i. e., one and only one solution $x(t) = (x_1(t), x_2(t), \dots, x_n(t))$ of the differential equations

$$(2.1)\qquad \frac{dx_\nu}{dt} = \frac{I_\nu}{\sum_{\nu=1}^{n} I_\nu^2} \qquad \left(I_\nu = \frac{\partial I}{\partial x_\nu};\ \nu = 1, 2, \dots, n\right).$$

Along such a gradient line it is possible to introduce I instead of t as parameter [4]). The part of the gradient line $x = x(I)$ given by $a \geqq I \geqq b$ is called the segment between a and b. If a is positive and the gradient line $x = x(I)$ is defined for all positive $I \leqq a$ and if $\lim_{I \to 0} x(I) = o$ we call the set of points x on this gradient line for which $0 < I \leqq a$ plus the point o the segment below a ending at o. The term "segment above b starting at o" for negative b is defined correspondingly.

We formulate as lemma 2.1 the following important property of a gradient line the proof of which is given in [6], § 9, section 3:

[1]) Lemmas 3.9 and 3.10 assert that $C(\varepsilon, \varepsilon_1)$ and the Γ_i are polyhedra.
[2]) $p^n = 0$ as the n-th Betti number of a polyhedron in E^n.
[3]) Cf. [6], p. 38. [4]) [6], p. 36.

Lemma 2.1. *Let $x^0 \neq o$ be a point of U and let $x = x(I)$ be the gradient line through x^0 such that $x^0 = x(I_0)$ where, by definition, $I_0 = I(x^0)$. Then one (and only one) of the following two cases must take place if we consider the gradient line $x = x(I)$ for $I \leqq I_0$: case 1: $x(I) \subset U$ for all $I \leqq I_0$; in this case we must have $I(x_0) > 0$ and $\lim_{I \to 0} x(I) = o$; case 2: there exists a $I_1 < I_0$ such that $x(I_1)$ is a boundary point of $\overline{U}$. The corresponding statements are true if we consider $x = x(I)$ for $I \geqq I_0$.*

The following notation will be used throughout: $\|x\|$ denotes the norm $\left(\sum_{i=1}^{n} x_i^2\right)^{\frac{1}{2}}$ of the point $x = (x_1, x_2, \dots, x_n)$; for any $R > 0$, V_R is the solid sphere $\|x\| < R$, $\overline{V}_R$ its closure, and S_R its boundary $\|x\| = R$. With these conventions let $R_0 > 0$ be such that

$$(2.2)\qquad V_{R_0} \subset \overline{V}_{R_0} \subset U$$

and let

$$(2.3)\qquad Z(V_{R_0}) = V_{R_0} \cap \{I = 0\}.$$

For any pair of positive numbers $\varepsilon, \varepsilon'$ we define the sets $G_{\varepsilon,\varepsilon'}(V_{R_0})$, $G_\varepsilon(V_{R_0})$, $G_{-\varepsilon'}(V_{R_0})$ as follows: $G_{\varepsilon,\varepsilon'}(V_{R_0})$ is the set of all points on the segments $-\varepsilon' \leqq I \leqq \varepsilon$ of those gradient lines which pass through points of $Z(V_{R_0})$; $G_\varepsilon(V_{R_0})$ is the set of all points on the segments below ε ending at o, and $G_{-\varepsilon'}(V_{R_0})$ the set of all points on the segments above $-\varepsilon'$ starting at o. Then it is known [1]) that there exists a positive ε_1 such that the union $G_{\varepsilon,\varepsilon'}(V_{R_0}) + G_\varepsilon(V_{R_0}) + G_{-\varepsilon'}(V_{R_0})$ is contained in U for any couple $\varepsilon, \varepsilon'$ satisfying

$$(2.4)\qquad 0 < \varepsilon < \varepsilon' \leqq \varepsilon_1.$$

This union is, by definition, a cylindrical neighborhood of o. We denote it by $C(\varepsilon, \varepsilon')$ and have, under the assumption (2.4),

$$(2.5)\qquad C(\varepsilon, \varepsilon') \subset U.$$

Lemma 2.2. *There exists an $R_1 > 0$ such that*

$$(2.6)\qquad \overline{V}_{R_1} \subset C(\varepsilon, \varepsilon_1) \quad \textit{and} \quad 0 < R_1 < R_0.$$

For the proof we refer to [6], p. 96, footnote 16.

Lemma 2.3. *Let $\zeta = \zeta(\varepsilon, \varepsilon_1)$ be the set of points on the segments $-\varepsilon_1 \leqq I \leqq \varepsilon$ of the gradient lines through the points of $S_{R_1} \cap \{I = 0\} = S_{R_1} \cap Z(V_{R_0})$. Let ζ^+, ζ^- be the subsets of ζ lying on the segments $0 \leqq I \leqq \varepsilon$ and $-\varepsilon_1 \leqq I \leqq 0$ respectively of these gradient lines such that*

$$(2.7)\qquad \zeta = \zeta^+ + \zeta^-, \quad \zeta^+ \cap \zeta^- = S_{R_1} \cap \{I = 0\}.$$

Finally let

$$(2.8)\qquad \gamma^-(\varepsilon, \varepsilon_1) = \gamma^- = C(\varepsilon, \varepsilon_1) \cap \{I = -\varepsilon_1\}, \quad \gamma^+(\varepsilon, \varepsilon_1) = C(\varepsilon, \varepsilon_1) \cap \{I = \varepsilon\},$$

and $\partial = \partial C(\varepsilon, \varepsilon_1)$ the boundary of $C(\varepsilon, \varepsilon_1)$. Then

$$(2.9)\qquad \partial = \zeta + \gamma^- + \gamma^+.$$

[1]) [6], p. 38.

Proof. We prove first that each point of $\zeta + \gamma^- + \gamma^+$ is a point of ∂. Let then $x \subset \gamma^+$. Since by definition $\gamma^+ \subset C(\varepsilon, \varepsilon_1)$, x is a point of $C(\varepsilon, \varepsilon_1)$. On the other hand since $I = \text{const} = \varepsilon$ on γ^+, the components of grad I tangent to γ^+ are zero which together with the fact that grad $I \neq 0$ at x shows that the derivative of I normal to γ^+ is different from zero. Consequently in each neighborhood of x there are points x' in which $I(x') > I(x) = \varepsilon$, i. e., points x' which are not in $C(\varepsilon, \varepsilon_1)$. This proves that a point x of γ^+ is a boundary point, and the proof for points of γ^- is obviously quite analogous.

Now let x' be a point of ζ^+. Then $x' \subset C(\varepsilon, \varepsilon_1)$ and we have to show that each neighborhood U_1 of x' contains points $\bar{x}$ not in $C(\varepsilon, \varepsilon_1)$. Now x' lies on the segment $0 \leqq I \leqq \varepsilon$ of the gradient line through some point x^0 of $S_{R_0} \cap \{I = 0\}$. It follows from classical theorems concerning the continuous dependence of a solution of the system (2.1) of ordinary differential equations on the initial values that there is a neighborhood U_0 of x^0 such that if $x = x(I)$ is a gradient line with $x(0) \subset U_0 \cap \{I = 0\}$, then $x(I_1) \subset U_1$ where $I_1 = I(x')$. If we now choose a point $\bar{x}(0) \subset U_0 \cap \{I = 0\}$ with $\|\bar{x}(0)\| > R_0 = \|x^0\|$, then the point $x = \bar{x}(I_1)$ on the gradient line through $\bar{x}(0)$ will be in U_1 and outside $C(\varepsilon, \varepsilon_1)$. Since the proof for a point of ζ^- is obviously analogous we have finished the proof that each point of $\zeta + \gamma^- + \gamma^+$ is a point of ∂.

We now have to show that these are the only points of ∂, i. e., that every other point x' of $C(\varepsilon, \varepsilon_1)$ is an interior point. That $x' = o$ is an interior point follows from lemma 2.2. To deal with the case $x' \neq o$ we first state the following:

Lemma 2.4. *Let $x' \neq o$ be a point in $C(\varepsilon, \varepsilon_1)$ which is not on ζ. Then there exists a positive ϱ of the following property: if W_ϱ is the spherical neighborhood of x' of radius ϱ and if $x = x(I)$ is a gradient line through a point of $w_\varrho = W_\varrho \cap \{I(x) = I(x')\}$, then either $\lim_{I \to 0} x(I) = o$, or the gradient line $x = x(I)$ intersects $\{I = 0\}$ in a point x^0 with $\|x^0\| < R_0$.*

We postpone the proof of this lemma and show first that it implies that x' is an interior point if $x' \not\subset \zeta + \gamma^+ + \gamma^-$. Since the argument is essentially the same in the two cases $I(x') \geqq 0$ and $I(x') \leqq 0$, we restrict ourselves to the case in which

$$I(x') \geqq 0. \tag{2.10}$$

First of all it follows from lemma 2.4 and the definition of $C(\varepsilon, \varepsilon_1)$ that all points of w_ϱ are in $C(\varepsilon, \varepsilon_1)$. Moreover since $x' \subset C(\varepsilon, \varepsilon_1)$ but not on γ^+ it follows from (2.10) that $0 \leqq I(x') < \varepsilon$. Consequently we may choose an h such that $0 < h < \min(\varepsilon_1, \varepsilon - I(x'))$. Then $-\varepsilon_1 < I(x') - h < I(x') + h < \varepsilon$. Therefore all points on the segments $I(x') - h \leqq I \leqq I(x') + h$ of gradient lines through points of w_ϱ will still be in $C(\varepsilon, \varepsilon_1)$. However the set of all these points is a "cylindrical neighborhood of a non critical point" (namely x') in the sense of Seifert and Threlfall[1]), and such a cylindrical neighborhood is known to contain a spherical neighborhood of x' [2]).

We return to the proof of lemma 2.4 restricting ourselves again to the case (2.10) without loss of generality. We distinguish 2 cases:

(i) The gradient line through x' meets $\{I = 0\}$ in a point $x^0 \neq o$. By definition of $C(\varepsilon, \varepsilon_1)$ we have then $0 < \|x^0\| \leqq R_0$, and since $x' \not\subset \zeta$, we have even $0 < \|x^0\| < R_0$. Therefore we can choose a $\sigma > 0$ such that $\|x\| < R_0$ for all x of the intersection v_σ of the spherical neighborhood about x^0 of radius σ with $\{I = 0\}$. Since the coordinates of a point on a gradient line for $I = 0$ depend continuously on the coordinates of the values taken at $I = I_1$, we can choose a $\varrho > 0$ such that the gradient lines through the points of the intersection w_ϱ of $\{I(x) = I(x')\}$ with the spherical neighborhood about x^0 of radius ϱ go through points of v_σ. This proves that $w_\varrho \subset C(\varepsilon, \varepsilon_1)$.

(ii) If $x = g'(I)$ is the gradient line through x', then $\lim_{I \to 0} g'(I) = o$. In this case we consider the sphere V_{R_1} with the property given in lemma 2.2. If $x' \subset V_{R_1}$, this lemma shows that x' is an interior point of $C(\varepsilon, \varepsilon_1)$. Suppose then $\|x'\| \geqq R_1$. Let $\bar{R}$ be a positive number $< R_1$. Because $\lim_{I \to 0} g'(I) = 0$ the segment below $I(x')$ ending at o of our gradient line intersects the sphere $S_{\bar{R}}$. Let $\bar{I}$ be the greatest I-value (in the interval $I(x'), 0$) for which $\bar{x} = g'(\bar{I})$ is a point of $S_{\bar{R}}$. Let $v = V_{R_1} \cap \{I(x) = I(\bar{x})\}$. By lemma 2.2 we have that

$$v \subset C(\varepsilon, \varepsilon_1) - \partial \quad \text{and} \quad I(x) = I(\bar{x}) > 0 \quad \text{for} \quad x \subset v. \tag{2.11}$$

By an argument similar to the one used above we can choose a spherical neighborhood of x' such that its intersection w_ϱ with $\{I(x) = I(x')\}$ has the following property: the gradient lines $x = g(I)$ through points of w_ϱ go through v. It follows from (2.11) and the definition of $C(\varepsilon, \varepsilon_1)$ that for each such gradient line we have either $\lim_{I \to 0} g(I) = o$ or that $g(0) \neq o$ with $\|g(0)\| < R_0$. Consequently all points of w_ϱ satisfy the assertion of lemma 2.4.

3. Consequences of the hypothesis H.

From now on, except in section 6, we will always suppose that o is an intermediate critical point and that all assumptions of the introduction, including hypothesis H, are satisfied. The present section is a discussion of the geometric nature of $C(\varepsilon, \varepsilon_1)$ under these assumptions.

Lemma 3.1. *Let Z_0 denote the set $Z(V_{R_0})$ defined in (2.3) minus the point o. Then each component of Z_0 is a differentiable $(n-1)$-manifold and contains points x with $\|x\| = r$ for all r satisfying*

$$0 < r \leqq R_0. \tag{3.1}$$

Proof. Since o is an intermediate point, i.e., neither maximum nor minimum for I, it follows from (1.3) that each neighborhood of o contains points at which $I < 0$ and also the points at which $I > 0$, and therefore also points $x \neq o$ at which

$$I(x) = 0. \tag{3.2}$$

This shows that Z_0 is not empty. Let, then, x^0 be a point of Z_0, and let $\varphi_1, \varphi_2, \ldots, \varphi_{n-1}$ be a system of coordinates on the unit sphere which together with $r = \|x\|$ form a "spherical" coordinate system of E^n, regular in some neighborhood of x^0, and let $r^0, \varphi_1^0, \varphi_2^0, \ldots, \varphi_{n-1}^0$ be the coordinates of x^0 in this sytem. We claim that at least one of the derivatives $\frac{\partial I}{\partial \varphi_i}$ $(i = 1, 2, \ldots, n-1)$ is different from 0 at x^0. Indeed otherwise we would have not only $\frac{\partial I}{\partial r} \neq 0$ at x^0 (since grad $I \neq 0$ at $x = x_0$) but it would also follow that the r-component

[1]) [6], p. 37. [2]) [6], p. 95, footnote 14.