DOUBLE NORMALS OF CONVEX BODIES(1)

BY

NICOLAAS H. KUIPER

ABSTRACT

In this paper we study the set of double normals of a solid convex body in E^n (e.g. n-simplex). There are at least n double normals. The lengths form a set of measure zero in $\mathbb{R}$ for $n \leqq 3$, not necessarily so for $n > 3$.

1. **The problem.** A bounded compact convex set B with at least one interior point in euclidean n-dimensional vector space $E = E^n$ will be called a *convex body*. Its boundary is denoted by ∂B. A chord is a line segment $[p, q]$ with end points $p, q \in \partial B$. It is called a *double normal* in case, in terms of inner products of vectors in E,

$$(x - p)(q - p) \geqq 0 \text{ and } (x - q)(p - q) \geqq 0 \quad \text{for all } x \in B.$$

1. *What can be said about the set of double normals of a convex body?*
In particular:
2. *Must a convex body in E^n admit at least n double normals?* (2)

2. **Examples.** The polar coordinate values of a vector $v \in E$ are by definition $r = \sqrt{v^2}$ and $\omega = v/r$. The unit vector ω is a point of the unit sphere $S^{n-1} \subset E$. The antipodal equivalence class $z = \pi(\omega) = \{\omega, -\omega\}$ of the unit vector ω is a point of the real projective $(n-1)$-space P^{n-1}, of which the unit sphere S^{n-1} is covering space under the double covering $\pi : S^{n-1} \to P^{n-1}$.

If ω is a unit vector in the direction of a double normal $[p, q]$ of B, $\omega = (q - p)/|q - p|$ then so is $-\omega = (p - q)/|p - q|$. The set of such unit

Received August 18, 1964.

(1) a. Research partially supported by the N.S.F. b. Lecture in the conference on differential geometry, June 1964, at Oberwohlfach, Germany.

(2) This last question was problem 1 in "Unsolved problems in intuitive geometry" by V. Klee (Mimeographed notes, University af Washington (1960)). Some mathematicians, e.g. T. Ganea, have been aware that the solution of this problem is essentially contained in work of L. Lyusternik, L. Schnirelmann [5] and M. Morse. V. Klee and B. Grünbaum suggested that publication, in particular for the case that B is not known to be C^2, is desirable anyhow. For the treatment of an analogous C^∞-problem in a space with a Riemannian metric see Bos [3]

vectors is therefore invariant under the antipodal map and it defines the unique set of double normal directions $K(B) \subset P^{n-1}$ which it covers.

THEOREM 1. *If $f: P^{n-1} \to \mathbb{R}$ is any real C^2-function ($=$ with continuous second derivatives), then there exists a symmetric convex body B in E^n with centre 0, for which the set of double normal directions is*

$$K(B) = K(f) \subset P^{n-1},$$

where $K(f)$ is the critical set

$$K(f) = \{z \mid z \in P^{n-1},\ (\mathrm{d}f)_z = 0\}.$$

Proof. For $t \geqq 0, t$ sufficiently small, the point set defined in terms of polar coordinates (r, ω) by

$$r = 1 + tf(\pi(\omega))$$

is a hypersurface ∂B_t, boundary of a body B_t in E . For t converging to zero this hypersurface converges to the unit sphere S^{n-1}, and the continuous first and second derivatives are included in this convergence. By compactness of P^{n-1} it follows that $\varepsilon > 0$ exists such that the hypersurface ∂B_ε has at each point all normal curvatures positive, as has S^{n-1}, and is therefore strictly convex. Because B_ε is symmetric with respect to 0 and strictly convex, every chord that connects two tangent points on parallel tangent hyperplanes, passes through 0. All double normal directions are then found from the equation

$$dr = d[1 + \varepsilon f(\pi(\omega))] = 0, \quad \text{or} \quad df = 0. \tag{1}$$

Consequently $K(B_\varepsilon) = K(f)$ and the theorem is proved.

THE CONDITION C^{2-}. A function $f: \mathbb{R}^m \to \mathbb{R}$ will be called C^{2-} in case it is C^1 (continuous first derivatives) and for any x_0 in the domain there exist δ and N such that

$$|f(x+h) - f(x) - (df)_x(h)| < N \cdot |h|^2 \tag{2}$$

for all x with $|x - x_0| < \delta$ and $h < \delta$. $(df)_x$ is the derivative of f at x. It sends the vector h into $(df)_x(h)$. If a function f on a C^∞-m-manifold (like P^m) has for any C^∞-chart $\kappa: U \to \mathbb{R}^m$ a composition $f\kappa^{-1}$ which is C^{2-}, then f is called C^{2-}.

With these definitions *we may replace in Theorem* 1 *C^2-function f by C^{2-}-function f, with the same conclusions.*

A special example is obtained as follows. If $x_1, \cdots, x_n$ are orthonormal coordinates for E^n and $g = \sum_{j=1}^n j x_j^2$, then the restriction of g to S^{n-1} has $2n$ critical points and it is invariant under the antipodal map. The well defined function $f = f\pi^{-1}: P^{n-1} \to \mathbb{R}$, which is the composition of the relations π^{-1} and f, has exactly n critical points, $y_1, \cdots, y_n$.

For any n given distinct points $z_1, \cdots, z_n$ on P^{n-1} there exists a C^∞-diffeomorphism τ carrying these points onto the n points $y_1, \cdots, y_n$ respectively. But then $f\tau : P^{n-1} \to \mathrm{R}$ is a function with $z_1, \cdots, z_n$ as critical points. With Theorem 1 we obtain for the corresponding convex bodies:

COROLLARY. *Given n arbitrary distinct straight lines through $0 \in E^n$ ($n \geqq 2$), for example for $n = 3$ all in one 2-plane, there exists a convex body B with exactly n double normals, and such that these are on the n given lines.*

REMARK. If all double normal directions are close to each other, then the boundary of the body lies between two concentric spheres with radii-ratio close to one.

EXERCISE. Consider also the geometry of the case defined by $g = \sum_{j=1}^m x_j^2$ with $m < n$.

3. A converse to theorem 1.

THEOREM 2. *Given a convex body B in E^n, there exists a centrally symmetric convex body B' with C^{2-}-boundary $\partial B'$, and there exists a C^{2-}-function $f: P^{n-1} \to \mathbb{R}$ such that*

$$K(B) = K(B') = K(f).$$

Proof. For any sets B and C in a vectorspace, let

$$-B = \{x \mid -x \in B\} \text{ and } B + C = \{x + y \mid x \in B, y \in C\}.$$

Consider a convex body B in E^n. Take any hyperplane in E^n and call all parallel hyperplanes "horizontal". Define the symmetric convex body B'' by $B'' = B + (-B)$ The tangent hyperplanes at $u \in \partial B''$ and $-u$ are parallel. If and only if these hyperplanes are horizontal, then there exist p and q in ∂B with $u = 2(p - q)$, at which the tangent planes are horizontal. The corresponding chords in B and B'' are double normal if and only if $p - q$ is vertical. Consequently a vector has the direction of a double normal of B if and only if it has the direction of a double normal of B'', and $K(B) = K(B'')$. Observe that the length of a double normal of B'' is twice the length of a corresponding double normal of B.

Let B' be the set of all points of E that have a distance smaller or equal to *one* to the convex body B''. B' is a symmetric convex body, whose boundary $\partial B'$ is "parallel" to $\partial B''$. Clearly $K(B') = K(B'')$. The length of a double normal of B' is 2 more than the length of the corresponding double normal of B''.

Now if $z \in \partial B'$, than there exists $y \in \partial B''$ such that $z - y$ has length 1. $\partial B'$ has no point in common with the open ball with radius 1 and centre y. B' also contains no point at the other side from $T(z)$, the hyperplane through z orthogonal to $z - y$. Then we conclude that $T(z)$ is the uniquely defined tangent hyperplane of $\partial B'$ at z. The surface $\partial B'$ is differentiable at every point, and it is the boundary of a convex body. Then it follows (see Bonnesen-Fenchel, Theorie der konvexen Körper, p. 26) that $\partial B'$ is a C^1-*hypersurface*. Equivalently the unit normal vector at $z \in \partial B'$ is a continuous function of z. If $\partial B'$ is given in polar-coordinates by $r = r(\omega) = r(-\omega)$ then the function $f: P^{n-1} \to \mathbb{R}$ required in theorem 2 is defined by $r(\omega) = f(\pi(\omega))$. f is clearly a C^1-function.

Moreover

$$K(f) = K(B') = K(B'') = K(B)$$

For each point $z \in \partial B'$ it follows from the fact that some neighborhood of z in $\partial B'$ is pinched between a hyperplane and a sphere both tangent to $\partial B'$ at z, that the condition for C^{2-} (see (2)) is fulfilled for f. This proves theorem 2. In the next section we study $K(f)$ and obtain conclusions which, applied to $K(B)$, give among others the affirmative answer to the problem in the introduction.

4. **Critical points of a continuous (or C^1) function on a closed topological (or C^1) manifold.** Let M be a closed (compact and without boundary) n-manifold and $f: M \to \mathbb{R}$ a C^0 (continuous) or C^1-function. A point $z \in M$ is called *non-critical* if there exists a C^0 or C^1 (resp.) coordinate system $\kappa: U(z) \to \mathbb{R}^n$ covering some neighborhood $U(z)$ of z, such that the last coordinate is the restriction of f to $U(z)$. A point is called *critical* if it is not noncritical. Let $K(f)$ be the set of all critical points of f. $K(f)$ is a closed set in M, because its complement is open. In the C^1-case the critical points z are those for which $(df)_z = 0$.

DEFINITION. The relative *Lyusternik-Schnirelmann* (L.S.-)*category)* $\gamma(A, M)$ of a closed subset A in M, is the minimal number of open contractible (in M) sets of M that can cover A. The absolute L.S.-category of M is $\gamma(M) = \gamma(M,M)$.

THEOREM 3. *Let Γ be a component of the set of critical points in a level set of the C^0- or C^1-function f on the compact closed manifold M: Γ is a component of $K(f) \cap \{z \mid f(z) = c\}$ for some c. Then*

$$\Sigma_\Gamma \gamma(\Gamma, M) \geqq \gamma(M),$$

where the summation extends over all components of each critical set in each level set.

As $\gamma(M)$ is finite, the conclusion is valid in case the summation extends over an infinite number of non-zero terms. Hence we may exclude this and assume that there are only a finite number of critical levels, and that each level set has a finite number of components.

LEMMA. *Let c be the only critical value of f in the half open interval $(b, e] \subset R$. Suppose $f_b = \{z \mid f(z) \leqq b\}$ is covered by p contractible (in M) open sets $V_1, \cdots, V_p$ with union V, and suppose the critical set at level c is covered* *by q contractible in (M) open sets $V_{p+1}, \cdots, V_{p+q}$ with union V'. Then $f_e = \{z \mid f(z) \leqq e\}$ can be covered by $p+q$ contractible open sets.*

Proof of the theorem from the lemma. Applying the lemma inductively, while starting from a value smaller than the minimal value of f as first example of b, we can go up to any value smaller than the maximum m of f. But the maximal value critical set is covered by a given minimal number of open contractible sets, which cover $M - f_{m-2\varepsilon}$ for some $\varepsilon > 0$. After going up to the value $m - \varepsilon$ we can just add these open sets to cover M completely. This implies Theorem 3.

Proof of the lemma. (Compare Kuiper [4]). If b is a critical value, and the compact set f_b is covered by some open sets, then these sets also cover $f_{b+\varepsilon}$ for some $\varepsilon > 0$. Hence we can replace b by $b + \varepsilon$ and we then can assume that b is not a critical value. This we will do now.

For every non-critical point $z \in M$ and neighbourhood V_z there exists a homeomorphism h_z of M and a neighborhood U_z of z, with $U_z \subset \bar{U}_z \subset V_z$, such that

$$\begin{aligned} f(h_z(u)) &\geqq f(u) && \text{for } u \in M \\ h_z(u) &= u && \text{for } u \notin V_z \\ f(h_z(u)) &> f(u) && \text{for } u \in U_z \end{aligned}$$

With the local coordinate system in which the last coordinate is f, this can be seen from the following model for such a homeomorphism in coordinate space $\mathbb{R}^n$ with coordinates $x_1, \cdots, x_n$ at the point 0 and $f = x_n$. Let $:\mathbb{R} \to \mathbb{R}$ be a C^∞-function for which

$$\begin{aligned} &\phi(s) = 1 && \text{for } |s| \leqq 1 \\ &0 \leqq \phi(s) \leqq 1 && \text{for } s \in \mathbb{R}. \\ &\phi(s) = 0 && \text{for } |s| > 2. \end{aligned}$$

Consider the C^∞-map h_0 given by

$$h_0(x_1, \cdots, x_n) = \left(x_1, \cdots, x_{n-1}, x_n + t\phi\left(\lambda \sum_{i=1}^{n} x_i^2\right)\right).$$

For λ large the support of this map is in a small ball about 0 as centre. For t small enough h is a diffeomorphism, and it lifts any point at most to a x_n-level which is t higher.

Any compact set $W \subset M - K(f)$ can be covered by a finite number of neighbourhoods U_z, each obtained as above with some non-critical point z, neighbourhood V_z and homeomorphism h_z. Let h be the product of these homeomorphisms in some arbitrary order. By a suitable choice (small) of the parameter t in each homeomorphism h_z, it can be attained that $0 \leqq f(h(u)) - f(u) < \delta$ for some arbitrarily given $\delta > 0$ and every $u \in M$. Finally let:

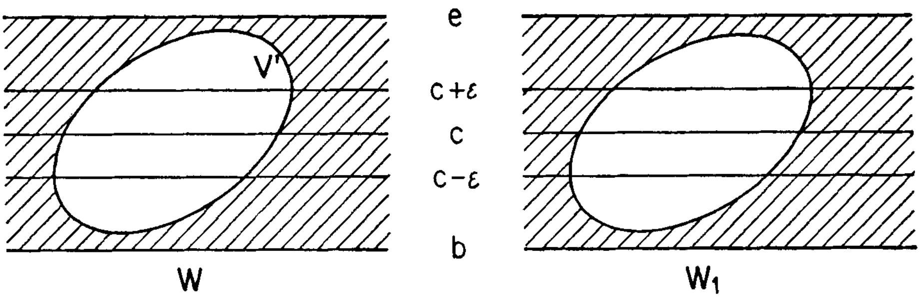
$$(3) \qquad 2\varepsilon = \inf_{u \in W} [f(h(u)) - f(u)] > 0.$$

We consider two examples for W to be called W and W_1 (See figure). W is the set $W = f_e - \operatorname{int} f_b - V'$, with the corresponding homeomorphism h, and $\varepsilon = \varepsilon(h)$. The choice is made such that $\varepsilon < \delta < \frac{1}{2}(c-b)$, so that

$$(4) \qquad h(f_b) \subset f_{c-\varepsilon}$$

The other example depends on ε and is as follows:

$$W_1 = (f_e - \operatorname{int} f_{c+\varepsilon}) + (f_{c-\varepsilon} - \operatorname{int} f_b).$$



The corresponding homeomorphism is h_1, and $2\varepsilon_1 = 2\varepsilon(h_1) = \inf_{u \in W} (f(h_1(u)) - - f(u))$. For every $z \in M$ we have

$$f(h(z)) \geqq f(z) \text{ and } f(h_1(z)) \geqq f(z)$$

If $z \in f_{c-\varepsilon}$, then $h_1^{-1}(z) \in f_{c-\varepsilon}$. Suppose moreover $h_1^{-1} z \notin f_b$, then

$$f(h_1^{-1} z) \leqq f(z) - 2\varepsilon_1.$$

Consequently one finds that

$$h_1^{-\alpha}(f_{c-\varepsilon}) \subset f_b \quad \alpha \text{ integer and } 2\alpha\varepsilon_1 > c-b$$

Or, as $V \supset f_b$,

$$(5) \qquad h_1^{\alpha}(V) \supset h_1^{\alpha}(f_b) \supset f_{c-\varepsilon} \quad \text{for } 2\alpha\varepsilon_1 > c-b$$

Analogously

$$(6) \qquad h_1^{\beta}(f_{c+\varepsilon}) \supset f_e \qquad \text{for } 2\beta\varepsilon_1 > e-c$$

If $hz \in f_{z+\varepsilon} - f_{c-\varepsilon}$ then $z \in f_{c+\varepsilon} - f_b$ by (4). If moreover $hz \notin hV'$, that is $z \notin V'$, then $f(hz) - f(z) \geqq 2\varepsilon$ hence $z \in f_{c-\varepsilon}$. Consequently for any point $hz \in f_{c+\varepsilon}$:

either $hz \in f_{c-\varepsilon}$, *then* $z \in f_{c-\varepsilon}$;

or $h_z \in hV'$, *then* $z \in V'$;

or $hz \in f_{c+\varepsilon} - f_{c-\varepsilon}$ and $hz \notin hV'$ — *then* $z \in f_{c-\varepsilon}$.

This implies

$$h(f_{c-\varepsilon} \cup V') \supset f_{c+\varepsilon} \tag{7}$$

From (5), (6) and (7) follows:

$$f_e \subset h_1^\beta(f_{c+\varepsilon}) \subset h_1^\beta h(f_{c-\varepsilon} \cup V') \subset h_1^\beta h(h_1^\alpha V \cup V')$$

and f_e is covered by the $p+q$ contractible open sets

$$h_1^\beta h h_1^\alpha(V_i) \qquad \text{for } i = 1, \cdots, p$$

and

$$h_1^\beta h h(V_j') \qquad \text{for } j = 1, \cdots, q \qquad \text{q.e.d.}$$

If M is the real projective $n-1$-space P^{n-1}, then it is known that $\gamma(M) = \gamma(P^{n-1}) = n$ and so for a function $f: P^{n-1} \to \mathbb{R}$ we have:

COROLLARY

$$\sum_\Gamma \gamma(\Gamma, P^{n-1}) \geqq n.$$

In particular if $\gamma(\Gamma, P^{n-1}) = 1$ for each Γ, for example if each Γ is one point, then the total number of these components is greater or equal to n.

5. **Application to the double normals problem.** Combining the results of section 4 with those of section 3 we get:

THEOREM 4. *A convex body B in E^n has at least n double normals. If $K(B) \subset P^{n-1}$ consists of a finite number of components Γ each belonging to a set of double normals of constant length, then*

$$\sum_\Gamma \gamma(\Gamma, P^{n-1}) \geqq n \tag{8}$$

For example if B is an ellipsoid in E^3 with *two* equal axes, then $K(B) \subset P^2$ consists of a point and a projective line and the left hand side in (8) is $1 + 2$ which is $\geqq 3$.

PROBLEM. Let $\mathcal{O}$ be the family of all convex symmetric bodies B in euclidean 3-space E^3, that have all double normals in a plane. Is there an upper bound to the ratio between the largest and smallest width of B for B in $\mathcal{O}$? How much is it? The ratio is

$$\frac{\max g(\omega)}{\min g(\omega)}$$

in case $r \leqq g(\omega)$ defines B in polar coordinates r and ω.

6. On the length of the double normals of a convex body.

THEOREM 5. *For $n \geqq 4$ there exists a convex body in E of non-constant width, such that every width (a real value) is attained as the length of some double normal. The body can be chosen such that there is an arc in P^{n-1} consisting of directions of double normals connecting a minimal width double normal direction to a maximal width double normal direction.*

Proof. Whitney [7] has given examples of C^{n-2}-functions f on the $n-1$-cube $I^{n-1} < \mathbb{R}^{n-1}$ with values in $I = \{t \mid 0 \leqq t \leqq 1\}$, such that df is zero at each point of an (non rectifiable) arc connecting two points with different f-values. This function can be carried over to P^{n-1} by imbedding I^{n-1} in P^{n-1} and extending the function suitably (Whitney [8])(3). The construction of section 2 then gives the convex body required in the theorem.

PROBLEM. In Theorem 4 it remains open whether the set of all double normal directions (not decomposed in constant length parts) obeys an analogous relation If Γ represents a component of this set, is then again $\sum_\Gamma \gamma(\Gamma, P^{n-1}) \geqq n$? A special problem in the same direction is:

Is there on the two-sphere a C^1-function f such that the set $\{x \mid (df)_x = 0\}$ is an arc?

For $n = 3$ we get a conclusion different from that in Theorem 5:

THEOREM 6. *The lengths of the double normals of a convex body B in E^3 form a set of measure zero in $\mathbb{R}$.*

Proof. This mainly consists in an application of the theorem of A. P. Morse and Sard. As in section 3 we replace B by B'' and then by B', and we denote the latter again by B.

Let this symmetric body B be defined by the inequality in polar coordinates $r \leqq f(\omega)$. Some neighbourhood of any boundary point $z \in \partial B$ in ∂B is pinched

(3) By a suitable modification of the example of Whitney [7] (see also Besicovith-Schoenberg [1]) one can obtain for $\delta > 0$ a real C^1 function f on $I^2 = \{(x_1, x_2) \in \mathbb{R}^2 \mid 0 \leqq x_1 \leqq 1, 0 \leqq x_2 \leqq 1\}$ such that the set of critical values

$$f(K) = f(\{x \mid (df)_x = 0\})$$

is the interval $I = \{t \mid 0 \leqq t \leqq 1\}$ but for which f is a function rather close to being C^2 in the following sense. There is a closed set J (in the examples it is an arc) and for $x \in J$, $y \in I^2$ one has, writing $r = \sqrt{(y-x)^2}$, $\Delta f = f(y) - f(x)$, that

$$\frac{\Delta f}{r^2(\ln r)^{2+\delta}} \text{ is bounded } (\delta > 0)$$

and f is C^∞ outside J.

PROBLEM. Does there exist an example $\delta = 0$?

between a unit ball and the tangent plane, mutually tangent at z, as we observed in section 3. Let $n(\omega)$ be the unit outside normal vector at the point $(r,\omega) = (f(\omega),\omega)$ of ∂B. From the assumptions about B it follows geometrically that

$$\frac{|n(\omega_1) - n(\omega_2)|}{|\omega_1 - \omega_2|}$$

is bounded. The function $n: S^2 \to S^2$ which assigns to ω the value $n(\omega)$ is then totally differentiable almost everywhere by a theorem of Rademacher (see Saks [9]). By a theorem of Federer (see Whitney [8]) for any $\varepsilon > 0$ there exists then a "big" closed set $Q \subset S^2$ such that $S^2 - Q$ has measure $< \varepsilon$, and such that $n(\omega)$ is a C^1-function on Q.

Geometrically one finds:

$d[f(\omega)\cdot\omega] = f(\omega)\cdot(n d\omega)\omega +$ orthogonal component. The left hand side is: $f d\omega + (df)\omega$

The inner product with ω yields, with $\omega^2 = 1$ and $d\omega^2 = 2\omega d\omega = 0$,

$$df = f(\omega)(n d\omega)$$

The coefficient $f(\omega)\cdot n(\omega)$ of $d\omega$ is C^1 for $\omega \in Q$.

Hence f is C^2 for $\omega \in Q \subset S^2$, and it can be extended to a C^2-function $g: S^2 \to S^2$ with

$$g(\omega) = f(\omega) \text{ for } \omega \in Q.$$

By the theorem of A. P. Morse-Sard [6, 10] the critical values of g form a set of measure zero. The critical points of f in Q therefore give a contribution zero to the measure of the set of critical values of f.

On the other hand the critical points of f outside Q have (Hausdorff-)measure $< \varepsilon$. Hence they can be covered by circular small discs with critical points as centres and with the sum of the areas smaller than 2ε.

Any such disc of small radius δ in ∂B is again pinched between a tangent plane and a ball with radius one (section 3), and it contributes at most $2(1 - \cos\delta) < 2\delta^2$ to the set of values of widths of the body. The sum of these contributions is then smaller than $2/\pi$ times the sum of the areas of the discs m, hence $< \varepsilon$. This being true for any ε, the theorem follows.

References

1. A. S. Besicovitch and I. J. Schoenberg, *On Jordan arcs and Lipschitz classes of functions defined on them*, Acta Math. **106** (1961), 114–136.

2. T. Bonnesen and W. Fenchel, *Theorie der konvexen Körper*, Berlin 1933.

3. W. Bos, *Kritische Schnen auf Riemannschen Elementarraumstücken*, Math. Ann. **151** (1963), 434–451.

4. N. H. Kuiper, *A continuous function with two critical points*, Bull. Amer. Math. Soc. **67** (1961), 281–285.

5. L. Lyusternik and L. Schnirelmann, *Méthodes topologiques dans les problèmes variationels*, Paris, 1934.

6. A. P. Morse, *The behaviour of a function on its critical set*, Ann. Math. **40** (1939), 62–70.

7. H. Whitney, *A function not constant on a connected set of critical points*, Duke Math. J. **1** (1935), 514–517.

8. H. Whitney, *On totally differentiable and smooth functions*, Pacific J. Math. **1** (1951), 143–159.

9. S. Saks, *Theory of the Integral*, New York, 1937, p. 311.

10. A. Sard, *Images of critical sets*, Ann. Math. **68** (1953), 247–259.

UNIVERSITY OF AMSTERDAM