

REGULAR POINTS OF LIPSCHITZ FUNCTIONS¹

BY

ALEXANDER D. IOFFE

ABSTRACT. Let f be a locally Lipschitz function on a Banach space X , and S a subset of X . We define regular (i.e. noncritical) points for f relative to S , and give a sufficient condition for a point $z \in S$ to be regular. This condition is then expressed in the particular case when f is C^1 , and is used to obtain a new proof of Hoffman's inequality in linear programming.

1. Introduction. Let F be a C^1 -mapping from a Banach space X (or from an open set in X) into another Banach space Y . It is usually called *regular* at z if $F'(z)$ (the derivative of F at z) maps X onto Y . A well-known theorem of Ljusternik [9] says that the tangent space to $Q = \{x | F(x) = F(z)\}$ at z coincides with the null space of $F'(z)$ if F is regular at z . It is shown in [7] that a slight change of the original Ljusternik proof leads to a stronger conclusion: if F is regular at z , then there are $\varepsilon > 0$, $r > 0$ such that $d_Q(x) \leq r \|F(x) - F(z)\|$ if $\|x - z\| \leq \varepsilon$ ($d_Q(x)$ denotes the distance from x to Q). In many situations, this latter property is what is really needed, not regularity in the original sense. This suggests the following general definition.

DEFINITION. Let X be a Banach space, $z \in S \subset X$, and let f be a real-valued function defined in a neighbourhood of z . Denote $Q = \{x | x \in S, f(x) = f(z)\}$. We shall say that z is a *regular point for f relative to S* , or that f is *regular at z relative to S* if there are $\varepsilon > 0$, $r > 0$ such that

$$d_Q(x) \leq r |f(x) - f(z)| \quad (1.1)$$

whenever $x \in S$, $\|x - z\| \leq \varepsilon$.

If Y is another Banach space and F is a mapping from a neighbourhood of z into Y , then z will be called a *regular point for F relative to S* if $f(x) = \|F(x) - F(z)\|$ is regular at z relative to S . If $S = X$, we shall merely say that z is a *regular point for f (or F)*.

Our purpose is to present certain sufficient conditions for regularity in the just-defined sense. The main result (§2) gives such a condition in terms of generalized gradients of Clarke [2], [3]. Among the corollaries presented in the

Received by the editors January 25, 1978.

AMS (MOS) subject classifications (1970). Primary 58E15; Secondary 49B30.

¹*Editor's Note.* The author's original manuscript contained two more sections which the referee felt were not yet ready for publication. We asked the author's approval for this deletion, but have not received any reply. We are therefore publishing this part of the work; it should not be suppressed merely because of difficulty in communication.

rest of the paper, there is an extension of the Ljusternik theorem to the case $S \neq X$ (§3), an infinite dimensional generalization of a result of Hoffman [5] containing an estimate for the distance to the set of solutions of a system of linear inequalities (§4).

The method of proving the main theorem is quite different from that used in [7]. The proof given here is based upon the variational principle of Ekeland [4] and seems to be much more direct and natural for (1.1)-like estimates (those with nonspecified constants).

This research was mainly motivated by the theory of necessary conditions of extrema. It is closely related to another paper by the author [6] where it is shown in particular that a constrained problem is reducible to an unconstrained one if the extremum point is regular for the equality constraint mapping relative to the nonfunctional constraints.

We presume that the reader is familiar with certain elementary facts of convex analysis such as duality between closed convex sets and sublinear functions. All these facts can be found in [7], [11]. We use the same symbols $\|\cdot\|$ and $\langle \cdot, \cdot \rangle$ to denote norms and pairings in different spaces. Hopefully, this will lead to no confusion.

2. The main theorem. To begin with, we recall the notion of generalized gradient introduced by Clarke [2], [3]. Let X be a Banach space, and let $f(x)$ be a function defined and Lipschitz in a neighbourhood U of a point $z \in X$, that is

$$\|f(x) - f(u)\| \leq L\|x - u\|, \quad \forall x, u \in U. \quad (2.1)$$

Then the function

$$x \rightarrow f^0(z, x) = \limsup_{\substack{u \rightarrow z \\ t \downarrow 0}} t^{-1}(f(u + tx) - f(u))$$

is convex and continuous on X . The set

$$\partial f(z) = \{x^* | f^0(z, x) \geq \langle x^*, x \rangle, \forall x \in X\},$$

the *generalized gradient* of f at z , is thereby convex, weak* compact and nonempty.

We refer the reader to [3] for the following elementary fact.

PROPOSITION 1. *Let (2.1) be satisfied, and let $z \in S \subset X$. If f attains its local minimum on S at z , then*

$$0 \in \partial f(z) + (L + 1)\partial d_S(z).$$

Here $d_S(x)$ is the distance from x to S .

Assume now that S is a closed subset of U , $z \in S$ and f is a function defined on U and satisfying there (2.1). We shall suppose f equal to zero at z and denote $Q = \{x | x \in S, f(x) = 0\}$.

THEOREM 1. *Let there be $c > 0$ and $\infty > \varepsilon > 0$ such that $\|x^* + u^*\| > c$ whenever $x^* \in \partial|f(x)|$, $u^* \in (L + 1)\partial d_S(x)$ for some x satisfying $x \in S$, $\|x - z\| < \varepsilon$, $f(x) \neq 0$. Then*

$$d_Q(x) < (1/c)|f(x)| \quad \text{whenever } x \in S, \|x - z\| < \varepsilon / (1 + (L/c)).$$

In particular z is a regular point for f relative to S .

PROOF. Assume that the statement is false. Then there is $u \in S$ such that

$$\|u - z\| < \varepsilon / (1 + (L/c)), \quad d_Q(u) > (1/c)|f(u)|.$$

We can obviously choose $t > 1$ to make the following inequalities valid:

$$\|u - z\| < \varepsilon / (1 + (tL/c)); \quad (2.2)$$

$$d_Q(u) > (t/c)|f(u)| = \gamma. \quad (2.3)$$

It is also obvious that

$$|f(u)| < \inf_{v \in S} |f(v)| + (c/t)\gamma.$$

We shall apply now the variational principle of Ekeland: if $g(x)$ is l.s.c. and $u \in X$, $\sigma > 0$ are such that

$$g(u) < \inf_v g(v) + \sigma,$$

then for any $\lambda > 0$ there is $u_\lambda \in X$ such that $\|u - u_\lambda\| < \lambda$, $g(v) + (\sigma/\lambda)\|v - u_\lambda\| > g(u_\lambda)$, $\forall v \in X$. The function

$$\delta(v, S) = \begin{cases} 0, & \text{if } v \in S, \\ \infty, & \text{if } v \notin S \end{cases}$$

is l.s.c. since S is closed, hence so is $g(v) = |f(v)| + \delta(v, S)$. Applying Ekeland's principle to this $g(v)$, $\sigma = (c/t)\gamma$, $\lambda = \gamma$, we find $x \in S$ such that

$$\|x - u\| < \gamma \quad (2.4)$$

and

$$\varphi(v) = |f(v)| + (c/t)\|v - x\|$$

attains its minimum on S at x . It follows from Proposition 1 that

$$0 \in \partial\varphi(x) + (L + 1)\partial d_S(x).$$

But [3, Proposition 1],

$$\partial\varphi(v) \subset \partial|f(v)| + (c/t)\partial\|v - x\|,$$

and $\partial\|v - x\|$ lies in the unit ball. Thus the two latter inclusions yield the existence of $x^* \in \partial|f(x)|$, and $u^* \in (L + 1)\partial d_S(x)$ such that

$$\|x^* + u^*\| < c/t < c. \quad (2.5)$$

According to (2.3), (2.4), $x \notin Q$ which (because $x \in S$) means that $f(x) \neq 0$. On the other hand, (2.2) and (2.4) imply that

$$\begin{aligned} \|x - z\| &< \|u - z\| + \|x - u\| < \|u - z\| + (t/c)|f(u)| \\ &< \|u - z\| + (tL/c)\|u - z\| < \varepsilon. \end{aligned}$$

This together with (2.5) contradicts the assumptions, hence proving the theorem.

The cone generated by $d_S(x)$:

$$N_S(x) = \bigcup_{\lambda > 0} \lambda \partial d_S(x)$$

is called the *normal cone* to S at x . The following result is thereby true.

COROLLARY 1.1. *Let there be $c > 0$ and $\infty > \varepsilon > 0$ such that $\|x^* + u^*\| > c$ whenever $x^* \in \partial|f(x)|$, $u^* \in N_S(x)$ for some x satisfying $x \in S$, $\|x - z\| < \varepsilon$, $f(x) \neq 0$. Then the conclusion of Theorem 1 holds.*

3. Differentiable mappings. An extension of the Ljusternik theorem. Let X and Y be Banach spaces and $F: X \rightarrow Y$. It is said that F is *strictly differentiable* at z [1] if it is Fréchet differentiable at z and

$$F(x + h) = F(x) + F'(z)h + r(x, h)\|h\|,$$

where $r(x, h) \rightarrow 0$ if $x \rightarrow z$, $h \rightarrow 0$.

For a linear continuous mapping $A: X \rightarrow Y$ and a convex cone $K \subset X$, we set

$$C(A, K) = \sup_{\|y\| < 1} \inf\{\|x\| \mid Ax = y, x \in K\}.$$

As usual, we take $\inf \emptyset = \infty$ so that $C(A, K) < \infty$ only if A maps K onto Y . If K is closed, a slight modification of the Banach open map theorem shows that the inverse conclusion is also true.

Let $S \subset X$ and $z \in S$ be given. For $x \in S$, we denote by $T_S(x)$ the *tangent cone* to S at x which is by definition the polar to $N_S(x)$:

$$T_S(x) = \{h \in X \mid \langle x^*, h \rangle \leq 0, \forall x^* \in N_S(x)\}.$$

THEOREM 2. *Assume that F is continuous in a neighbourhood of z , strictly differentiable at z , $F(z) = 0$ and*

$$C(F'(z), T_S(x)) < k < \infty \quad (3.1)$$

for all $x \in S$ lying in a neighbourhood of z . Then z is a regular point for F relative to S .

PROOF. Denote for brevity $A = F'(z)$, $f(x) = \|F(x)\|$. First we shall show that for any $\varepsilon > 0$, there is a neighbourhood $U(\varepsilon)$ of z such that

$$\partial f(x) \subset F'^*(z)n(F(x)) + B_\varepsilon \quad (3.2)$$

whenever $x \in U(\varepsilon)$; here

$$n(y) = \{y^* \in Y^* \mid \|y^*\| \leq 1, \langle y^*, y \rangle = \|y\|\} \quad (3.3)$$

is the subdifferential of the Y -norm and

$$B_\varepsilon = \{y^* \in Y^* \mid \|y^*\| \leq \varepsilon\}.$$

Let $\varepsilon > 0$ be given. Take $\delta > 0$ such that $\|r(x, h)\| < \varepsilon$ if $\|x - z\| < \delta$, $\|h\| < \delta$ and choose $y^*(x) \in n(F(x))$ for any such x . We have

$$\begin{aligned} f^0(x, h) &= \limsup_{\substack{u \rightarrow x \\ t \downarrow 0}} t^{-1} (\|F(u + th)\| - \|F(u)\|) \\ &< \limsup_{\substack{u \rightarrow x \\ t \downarrow 0}} t^{-1} \langle y^*(u + th), F(u + th) - F(u) \rangle \\ &= \limsup_{\substack{u \rightarrow x \\ t \downarrow 0}} \langle y^*(u + th), Ah + r(u + th, th)\|h\| \rangle. \end{aligned}$$

Any weak* limit point of $y^*(u + th)$ belongs to $n(F(x))$ since the set-valued mapping $y \rightarrow n(y)$ is u.s.c. from the norm topology of Y into weak* topology of Y^* and F is continuous.

If $\|x - z\| < \delta$, then according to the choice of δ ,

$$f^0(x, h) < \max\{\langle y^*, Ah \rangle \mid y^* \in n(F(x))\} + \varepsilon \|h\|$$

which is exactly the same as (3.2).

Take $\varepsilon < 1/(8k)$ and choose a neighbourhood $U \subset U(\varepsilon)$ so that (3.1) holds for all $x \in U \cap S$. Let $x \in U \cap S$ be such that $F(x) \neq 0$, and choose $x^* \in \partial f(x)$ and $u^* \in N_S(x)$. We have

$$\begin{aligned} \|x^* + u^*\| &= \sup\{\langle x^* + u^*, h \rangle \mid \|h\| \leq 1\} \\ &> \sup\{\langle x^*, h \rangle \mid \|h\| \leq 1, h \in -T_S(x)\}. \end{aligned}$$

On the other hand, by (3.2)

$$x^* = F'^*(z)y^* + h^*,$$

where $\|y^*\| = 1$ (since $F(x) \neq 0$) and $\|h^*\| < \varepsilon$. Therefore

$$\|x^* + u^*\| > \sup\{\langle y^*, F'(z)h \rangle - \varepsilon \|h\| \mid \|h\| \leq 1, h \in -T_S(x)\}. \quad (3.4)$$

Take $v \in Y$ such that $\|v\| = 1$, $\langle y^*, v \rangle > 1/2$. By virtue of (3.1), there is $u \in -T_S(x)$ such that $\|u\| \leq 2k$, $F'(z)u = v$. Let $h = u/\|u\|$. Then $\|h\| = 1$, $h \in -T_S(x)$ and

$$\langle y^*, F'(z)h \rangle = \langle y^*, v \rangle \|u\|^{-1} > 1/(4k).$$

This shows together with (3.4) that $\|x^* + u^*\| > 1/(8k)$. It remains to apply Corollary 1.1.

The Ljusternik theorem in the form stated in the Introduction follows immediately from what we have just proved.

COROLLARY 2.1. *Assume that F is continuous in a neighbourhood of z , strictly differentiable at z and that $F'(z)$ maps X onto Y . Then z is a regular point for F .*

PROOF. We have $C(F'(z), X) < \infty$ because $F'(z)$ maps X onto Y (see [7, §0.2]).

Here is a more general condition ensuring $C(A, K) < \infty$.

PROPOSITION 2. *Assume that*

$$(\text{int } K) \cap (\text{Ker } A) \neq \emptyset, \quad R(A) = Y,$$

where $\text{int } K$ is the interior of K , $\text{Ker } A$ and $R(A)$ are the null space and the range of A . Then $C(A, K) < \infty$.

PROOF. Since $R(A) = Y$, we have that $C(A, X) = c < \infty$. Let $u \in (\text{int } K) \cap (\text{Ker } A)$. This means that $Au = 0$ and there is $\varepsilon > 0$ such that $u + h \in K$ whenever $\|h\| < \varepsilon$. Take $y \in Y$ with $\|y\| < 1$ and choose $x \in X$ such that $Ax = y$, $\|x\| < 2c$. Then $w = x + (2c/\varepsilon)u \in K$, $Aw = Ax = y$ and $\|w\| < 2c(1 + \|u\|/\varepsilon)$. Hence $C(A, K) < 2c(1 + \|u\|/\varepsilon)$.

It follows that (3.1) will be fulfilled if there is $\varepsilon > 0$ such that for any x in a neighbourhood of z , we can find $u(x) \in X$ such that $F'(z)u(x) = 0$ and $u(x) + h \in T_S(x)$ whenever $\|h\| < \varepsilon\|u(x)\|$. This is obviously the case when $(\text{int } T_S(z)) \cap (\text{Ker } F'(z)) \neq \emptyset$ and the cone-valued mapping $x \rightarrow T_S(x)$ is l.s.c. at z . Hence

COROLLARY 2.2. *Let F be continuous in a neighbourhood of z and strictly differentiable at z . Assume that $(\text{int } T_S(z)) \cap (\text{Ker } F'(z)) \neq \emptyset$ and the cone-valued mapping $x \rightarrow T_S(x)$ is l.s.c. at z (from the norm topology into the norm topology). Then z is a regular point for F relative to S .*

4. An extension of a Hoffman estimate for the distance to the set of solutions to a system of linear inequalities.

THEOREM 3. *Let $A: X \rightarrow Y$ be a linear continuous mapping such that $R(A)$, the range of A , is closed. Let $x_i^* \in X^*$, $i = 1, \dots, n$, be given, and consider the set*

$$Q = \{x \in X \mid Ax = 0, \langle x_i^*, x \rangle \leq 0, i = 1, \dots, n\}.$$

Then there is $k > 0$ such that for any $x \in X$,

$$d_Q(x) \leq k \left(\|Ax\| + \sum_{i=1}^n \langle x_i^*, x \rangle^+ \right).$$

Here $t^+ = \max\{0, t\}$.

For the case of finite-dimensional X , Y , this result was established by Hoffman [5]. Theorem 1 is likely to be helpful for proving other inequalities of such sort, for instance like those considered in [10].

First we shall prove the following auxiliary fact.

PROPOSITION 3. *Let A be a linear continuous mapping from X onto Y such that $C(A, X) < \infty$. Then*

$$\|A^*y^*\| \geq C(A, X)^{-1}\|y^*\|, \quad \forall y^* \in Y^*.$$

PROOF. Let $y^* \in Y^*$ be given. Take $\varepsilon > 0$, and choose $y \in Y$ such that $\|y\| = 1$, $\langle y^*, y \rangle > \|y^*\| - \varepsilon$. By definition, there is $u \in X$ such that $Au = y$, $\|u\| \leq C(A, X) + \varepsilon$. Let $x = u/\|u\|$. Then

$$\begin{aligned} \|A^*y^*\| &> \langle A^*y^*, x \rangle = \|u\|^{-1} \langle y^*, y \rangle \\ &> (C(A, X) + \varepsilon)^{-1} (\|y^*\| - \varepsilon). \end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, the result follows.

PROOF OF THEOREM 3. For $n = 0$, the result follows from Corollary 2.1 because $R(A)$ is a Banach space. Suppose the theorem is true whenever we have not more than $n - 1$ inequality, and consider the case of exactly n inequalities.

Assume first that x_1^*, \dots, x_n^* are such that

$$A^*y^* + \lambda_1 x_1^* + \dots + \lambda_n x_n^* \neq 0 \quad (4.1)$$

if the λ_i are nonnegative and not all equal to zero.

Since $R(A)$ is a closed subspace of Y , we may consider $R(A)$ the space into which A is acting or, in other words, identify $R(A)$ with Y . Then A^* is a linear homeomorphism.

Consider the following quantity:

$$a = \inf \left\{ \left\| A^*y^* + \sum_{i=1}^n \lambda_i x_i^* \right\| \mid \max\{\|y^*\|, \lambda_1, \dots, \lambda_n\} = 1, \right. \\ \left. \lambda_i \geq 0, i = 1, \dots, n \right\}.$$

Let us verify that $a > 0$. Indeed, if $a = 0$, then there are sequences $\{y_k^*\}$, $\{\lambda_{ik}\}$, $i = 1, \dots, n$, such that $\max\{\|y_k^*\|, \lambda_{1k}, \dots, \lambda_{nk}\} = 1$ for all k , $\lambda_{ik} > 0$ and

$$u_k^* = A^*y_k^* + \sum_{i=1}^n \lambda_{ik} x_i^*$$

norm-converges to zero. We can obviously suppose that λ_{ik} norm-converges to certain λ_i as $k \rightarrow \infty$ for every i , so that $A^*y_k^*$ also norm-converges in X^* . It follows, due to the fact that A^* is a homeomorphism, that y_k^* norm-converges to some y^* . We have

$$A^*y^* + \sum_{i=1}^n \lambda_i x_i^* = 0 \quad (4.2)$$

and

$$\max\{\|y^*\|, \lambda_1, \dots, \lambda_n\} = 1, \quad \lambda_i \geq 0, i = 1, \dots, n.$$

If $\lambda_i = 0$, $i = 1, \dots, n$, then $\|y^*\| = 1$ and (4.2) contradicts Proposition 3. On the other hand, if not all of λ_i are equal to zero, then (4.2) contradicts (4.1). Hence $a > 0$.

Consider now the function

$$f(x) = \|Ax\| + \sum_{i=1}^n \langle x_i^*, x \rangle^+.$$

We see that

$$Q = \{x \in X | f(x) = 0\}.$$

On the other hand, f is sublinear, Q is a convex cone and hence the inequality $d_Q(x) \leq kf(x)$ will hold everywhere if it is true in a neighbourhood of the origin. This inequality, however, is just what we want to prove. According to Theorem 1, it suffices to verify that there is $c > 0$ such that $\|x^*\| > c$ whenever $x^* \in \partial f(x)$ and $f(x) \neq 0$.

We have

$$\partial f(x) = A^*(n(Ax)) + \sum_{i=1}^n m(\langle x_i^*, x \rangle) x_i^* \quad (4.3)$$

where

$$m(t) = \begin{cases} 1, & \text{if } t > 0, \\ 0, & \text{if } t < 0, \\ [0, 1], & \text{if } t = 0, \end{cases}$$

is the subdifferential of the function $t \rightarrow t^+$. In other words, any $x^* \in \partial f(x)$ can be represented as follows:

$$x^* = A^*y^* + \sum_{i=1}^n \lambda_i x_i^*, \quad (4.4)$$

where $y^* \in n(Ax)$, $\lambda_i \in m(\langle x_i^*, x \rangle)$, $i = 1, \dots, n$.

Let $f(x) \neq 0$. If $\lambda_i < 1$ for all $i = 1, \dots, n$, then $\langle x_i^*, x \rangle^+ = 0$, hence $Ax \neq 0$ and $\|y^*\| = 1$ whenever $y^* \in n(Ax)$. Therefore $\max\{\|y^*\|, \lambda_1, \dots, \lambda_n\} = 1$ in (4.4) and it follows from what we have already proved that $\|x^*\| > a > 0$.

Thus the theorem is true under the additional assumption (4.1). Assume now that this does not hold, that is to say, one of x_i^* , say x_n^* , can be represented in the form

$$x_n^* = - \sum_{i=1}^{n-1} \beta_i x_i^* - A^*z^* \quad (4.5)$$

for certain $\beta_i > 0$, $i = 1, \dots, n-1$, $z^* \in Y^*$. It follows that $\langle x_n^*, x \rangle = 0$ for all $x \in Q$.

Consider a mapping $A': X \rightarrow Y \times R$ defined by

$$A'x = (Ax, \langle x_n^*, x \rangle).$$

The range of A' is clearly closed and

$$Q = \{x | A'x = 0, \langle x_i^*, x \rangle \leq 0, 1 \leq i \leq n-1\}.$$

For A' , x_1^*, \dots, x_{n-1}^* the theorem is true by induction and the inequality

$$|\langle x_n^*, x \rangle| \leq \langle x_n^*, x \rangle^+ + \sum_{i=1}^{n-1} \beta_i \langle x_i^*, x \rangle^+ + \|Ax\| \|z^*\|$$

following immediately from (4.5) completes the proof.

REFERENCES

1. N. Bourbaki, *Variétés différentielles et analytiques*, Hermann, Paris, 1971.
2. F. H. Clarke, *Necessary conditions for nonsmooth problems in optimal control*, Ph. D. Thesis, Univ. of Washington, 1973.
3. ———, *A new approach to Lagrange multipliers*, Math. Oper. Res. 1 (1976), 165–174.
4. I. Ekeland, *On the variational principle*, J. Math. Anal. Appl. 47 (1974), 324–353.
5. A. Hoffman, *On approximate solutions of systems of linear inequalities*, J. Res. Nat. Bur. Standards Sect. B 49 (1952), 263–265.
6. A. D. Ioffe, *Necessary and sufficient conditions for a local minimum* (to appear).
7. A. D. Ioffe and V. M. Tikhomirov, *Theory of extremal problems*, "Nauka", Moscow, 1974. (Russian)
8. G. Lebourg, *Valeur moyenne pour gradient généralisé*, C. R. Acad. Sci. Paris Sér. A-B 281 (1975), 795–797.
9. L. A. Ljusternik and V. I. Sobolev, *Elements of functional analysis*, "Nauka", Moscow, 1965.
10. S. M. Robinson, *An application of error bounds for convex programming in a linear space*, SIAM J. Control 13 (1975), 271–273.
11. R. T. Rockafellar, *Convex analysis*, Princeton Univ. Press, Princeton, N. J., 1970.

PROFSOUZNAJA 97-1-203, MOSCOW B-279, RUSSIA