Acta Sci. Math., 38 (1977). 391—396

# Generalization of the implicit function theorem and of Banach's open mapping theorem

T. SZILÁGYI

In the present paper we prove the existence of implicit functions (Theorem 1) and of "right-inverse" functions (Theorem 2) under very weak assumptions. In Theorem 3 we generalize the open mapping theorem of Banach to a non-linear case and in Theorem 4 we give a new proof of a known multiplier rule (see [4]). The proof of Theorem 1 is based on Banach's open mapping theorem (see for example [2]), on Nadler's fixed point theorem for multivalued contractions (see [3]), and on the Lagrange inequality (see for example [1]). Theorem 2 is a simple consequence of Theorem 1, Theorem 3 follows easily from Theorem 2, finally Theorem 4 is based on Theorem 2 and on the Banach—Hahn theorem.

Notations. If $X$ and $Y$ are Banach spaces, then the set of all linear continuous mappings from $X$ into $Y$ will be denoted by $L(X, Y)$.

For defining equations we use the symbol $:=$ on the left side of which we write the "quantity" (number, function, set, etc.) to be defined.

If $(X, d)$ is a metric space, $r$ a positive number and $x \in X$, then

$$S(x, r) := \{y \in X \mid d(x, y) < r\} \quad and \quad B(x, r) := \{y \in X \mid d(x, y) \leqq r\}.$$

The dual of a Banach-space $X$ will be denoted by $X'$.

In $X$ and $Y$ are Banach-spaces and $A \in L(X, Y)$, then

$$p(A) := \sup_{y \in Y \setminus \{0\}} \left\{ \|y\|^{-1} \cdot \inf \{\|x\| \mid x \in X, Ax = y\} \right\}.$$

Lemma 1. *If $X$ and $Y$ are Banach-spaces, $A \in L(X, Y)$ and* $\operatorname{Im} A = Y$, *then $p(A)$ is finite.*

Proof. The conditions of Banach's open mapping theorem are fulfilled, therefore there exists a positive $r$ such that $B(O_Y, r)$ is contained in the $A$-image of

Received December 9, 1976.

$B(O_X, 1)$. Let us take an arbitrary $0 \neq y \in Y$, then

$$\inf\{\|x\| \, x \in X, Ax = y\} = \frac{\|y\|}{r} \inf\left\{\|x\| \,|\, x \in X, Ax = \frac{ry}{\|y\|}\right\} \leqq \frac{\|y\|}{r};$$

consequently $p(A) \leqq \dfrac{1}{r}$.

Lemma 2. *Let $(X, d)$ be a complete metric space, $\bar{x} \in X$, $r>0$ and $\Phi: S(\bar{x}, r) \to 2^X$ such that*

a) *for all $x \in S(\bar{x}, r)$, $\Phi(x)$ is a non-empty closed subset of $X$,*

b) *for all $x_1, x_2 \in S(\bar{x}, r)$, the Hausdorff distance*

$$h(\Phi(x_1), \Phi(x_2)) := \max\{\sup_{x \in \Phi(x_1)} d(x, \Phi(x_2)), \sup_{x \in \Phi(x_2)} d(\Phi(x_1), x)\}$$

*satisfies*

$$h(\Phi(x_1), \Phi(x_2)) \leqq \frac{1}{2} d(x_1, x_2),$$

c) $d(\bar{x}, \Phi(\bar{x})) < \dfrac{r}{2}$.

*Then there exists an $x \in S(\bar{x}, r)$ such that $x \in \Phi(x)$.*

The proof of this lemma can be found in [3], and in [4].

Lemma 3. *Let $X$ be a normed space, $L$ a linear subspace of $X$, $u_1, u_2 \in X$; $M_i := u_i + L$ $(i=1, 2)$. Then the Hausdorff distance* (see Lemma 2) *of $M_1$ and $M_2$ equals* $\inf\{\|v_1 - v_2\| : v_1 \in M_1, v_2 \in M_2\}$.

Proof. Clearly,

$$\inf_{v_i \in M_i} \|v_1 - v_2\| \leqq \inf_{v_2 \in M_2} \|u_1 - v_2\| = d(u_1, M_2) \leqq \sup_{v_1 \in M_1} d(v_1, M_2) \leqq h(M_1, M_2).$$

If $v_1 \in M_1$ and $v_2 \in M_2$, then $v_i = u_i + y_i$ $(i=1, 2, y_i \in L)$, thus $u_2 + y_2 - y_1 \in M_2$ and

$$\|v_1 - v_2\| = \|u_1 - (u_2 + y_2 - y_1)\| \geqq d(u_1, M_2),$$

consequently,

$$\inf_{v_i \in M_i} \|v_1 - v_2\| = d(u_1, M_2).$$

Similar arguments show that for all $v_1 \in M_1$ and $v_2 \in M_2$

$$d(u_1, M_2) = d(v_1, M_2) = d(M_1, v_2),$$

therefore $d(u_1, M_2) = h(M_1, M_2)$.

Theorem 1. *Let $X$ and $Y$ be Banach spaces, $T$ a topological space, $G \subset T \times X$ an open set, $(t_0, x_0) \in G$ and $F: G \to Y$ a function such that*

a) $F(t_0, x_0) = O_Y$, $\lim\limits_{t \to t_0} F(t, x_0) = O_Y$;

b) *for every* $(t, x) \in G$ *the function* $F(t, \cdot)$ *has a Fréchet-derivative at* $x$ *denoted by* $D_2F(t, x)$,

c) *the function* $D_2F\colon G \to L(X, Y)$ *is continuous at* $(t_0, x_0)$, *and*

d) $\operatorname{Im} D_2F(t_0, x_0) = Y$.

*Then for every neighborhood* $V \subset X$ *of* $x_0$ *there exist a neighborhood* $U \subset T$ *of* $t_0$ *and a function* $\varphi\colon U \to V$ *such that* $F(t, \varphi(t)) = O_Y$ *for all* $t \in U$.

Proof. Let us denote $A := D_2F(t_0, x_0)$. By Lemma 1 we have $p(A) < +\infty$. By assumption c) there exist a neighborhood $U_1$ of $t_0$ and a positive number $r$ such that $W := S(x_0, r) \subset V$ and for all $(t, x) \in U_1 \times W$

$$\|D_2F(t, x) - A\| < \frac{1}{2p(A)}.$$

Now we get from the Lagrange inequality for all $(t, z_1)$ and $(t, z_2) \in U_1 \times W$

$$(1) \qquad \|F(t, z_2) - F(t, z_1) - A(z_2 - z_1)\| \leqq$$

$$\leqq \sup_{\lambda \in [0,1]} \|D_2F(t, \lambda z_1 + (1-\lambda) z_2) - A\| \, \|z_2 - z_1\| < \frac{\|z_2 - z_1\|}{2p(A)}.$$

By assumption a) we may choose a neighborhood $U$ of $t_0$ such that

$$(2) \qquad p(A)\|F(t, x_0)\| < \frac{r}{2} \quad \text{for all} \quad t \in U.$$

Let $t \in U$ be a fixed element. We shall show that the equation $F(t, x) = 0$ has a solution $x \in W$. We apply Lemma 2 to the Banach space $X$, to the element $\bar{x} := O_X$ and to the (multivalued) function

$$x \mapsto \Phi(x) := \{z \in X \mid Ax - Az = F(t, x_0 + x)\}.$$

If $x \in S(0, r)$, assumption d) implies that $\Phi(x)$ is non-empty, the continuity of $A$ implies that $\Phi(x)$ is closed. Moreover, since $A$ is linear, $\Phi(x)$ is an affine subspace. Therefore (by Lemma 3) if $x_1, x_2 \in S(O_X, r)$, then

$$h(\Phi(x_1), \Phi(x_2)) = \inf_{v_i \in \Phi(x_i)} \|v_1 - v_2\| =$$

$$= \inf \{\|v_1 - v_2\| \mid Av_i = Ax_i - F(t, x_0 + x_i)\ i = 1, 2\}.$$

Since $\operatorname{Im} A = Y$, the latter infimum equals

$$\inf \{\|v\| \mid Av = A(x_1 - x_2) - F(t, x_0 + x_1) + F(t, x_0 + x_2)\} \leqq$$

$$p(A)\|F(t, x_0 + x_2) - F(t, x_0 + x_1) - A(x_2 - x_1)\| \leqq \frac{1}{2}\|x_1 - x_2\|$$

(see (1)). From (2) it follows that condition c) in Lemma 2 is fulfilled, too:

$$d\big(0_X, \Phi(0_X)\big) = \inf\{\|z\| \,|\, Az = -F(t, x_0)\} \leqq p(A)\|F(t, x_0)\| < \frac{r}{2}.$$

Thus, by Lemma 2, there exists an element $g(t) \in S(O_X, r)$ such that $g\big(t(\in \Phi(g/t)\big)$, consequently

$$0_Y = A\big(g(t)\big) - A\big(g(t)\big) = F\big(t, x_0 + g(t)\big).$$

Finally, for the element $\varphi(t) := x_0 + g(t) \in V$ we have

$$F\big(t, \varphi(t)\big) = 0_Y.$$

Theorem 2. *Suppose we are given two Banach spaces: $X$ and $Y$, an open set $V \subset X$, an element $x_0 \in V$ and a Fréchet-differentiable function $f\colon V \to Y$ for which*

a) *$f'\colon V \to L(X, Y)$ is continuous at $x_0$,*

b) $\operatorname{Im} f'(x_0) = Y$.

*Then there exist a neighborhood $U$ of the point $t_0 := f(x_0)$ and a function $\varphi\colon U \to V$ such that $f \circ \varphi = id_U$ (that is $f\big(\varphi(t)\big) = t$ for all $t \in U$); consequently, $t_0$ is an interior point of the range of $f$.*

Proof. Let us define the function $F\colon Y \times V \to Y$ by letting

$$F(t, x) := f(x) - t.$$

Clearly, we can apply Theorem 1 whith $T := Y$ and $G := Y \times V$ and this gives the result to be proved.

Theorem 3. *If $X$ and $Y$ are Banach spaces, $g\colon X \to Y$ is a Fréchet-differentiable function, $g'\colon X \to L(X, Y)$ is continuous and* $\operatorname{Im} g'(x) = Y$ *in every point $x \in X$, then $g$ is an open mapping.*

Proof. Let $V \subset X$ be an open set and $f := g|_V$. We must prove that the range $R$ of $f$ is an open set in $Y$. If $t_0 \in R$, then there exists a point $x_0 \in V$, for which $f(x_0) = t_0$. From Theorem 2 it follows that $R$ is a neighborhood of $t_0$.

Theorem 4. *Let $X$ and $Z$ be Banach spaces, $W \subset X$ an open set, $g\colon W \to R$ and $G\colon W \to Z$ Fréchet-differentiable functions. If a point $x_0 \in W$ affords a local minimum to $g$ under the constraint $G(x) = O_Z$, $g'$ and $G'$ are continuous at $x_0$ and* $\operatorname{Im} G'(x_0)$ *is closed in $Z$, then there exist a real number $\lambda$ and a continuous linear functional $l \in Z'$ such that*

(i) *at least one of them is different from* 0,

(ii) *for all $x \in X$, $\lambda g'(x_0)x + l\big(G'(x_0)x\big) = 0$.*

Proof. Let us choose an open set $V \subset X$ containing $x_0$ such that $x_0$ minimizes the function $g|_V$ under the constraint $G|_V = O_Z$ and let us denote $Y := R \times Z$; for all $x \in V$ $f(x) := (g(x), G(x))$. From our assumptions it follows that the function $f: V \to Y$ is Fréchet-differentiable, $f'$ is continuous at $x_0$, and for all $x \in X$

$$f'(x_0)x = (g'(x_0)x, G'(x_0)x).$$

First we observe that $\mathrm{Im}\, f'(x_0) \neq Y$. Indeed, if $\mathrm{Im}\, f'(x_0)$ were the whole space $Y$, then we could apply Theorem 2: there would be points $x \in V$ with $G(x)=0$ and $g(x) < g(x_0)$, since $(g(x_0), G(x_0))$ would be an interior point to the range of $f$. But this is impossible, because $x_0$ is a solution of the minimum problem on $V$. Therefore $\mathrm{Im}\, f'(x_0)$ is a proper linear subspace of $Y$. If it is a closed subspace, then we can apply a known corollary of the Banach—Hahn theorem: there exists a $0 \neq \bar{l} \in Y'$ such that $\bar{l} \circ f'(x_0) = 0$; and since the continuous linear functional $\bar{l}$, defined on the product space $R \times Z$ is of the form

$$\bar{l}(t, z) = \lambda t + l(z)$$

(where $\lambda \in R$ and $l \in Z'$), in this case the proof is complete. If the subspace $\mathrm{Im}\, f'(x_0)$ is not closed, then there exist a sequence $(x_n) \subset X$ and an element $(r, z) \in Y \setminus \mathrm{Im}\, f'(x_0)$ such that

$$\lim g'(x_0)x_n = r \quad \text{and} \quad \lim G'(x_0)x_n = z.$$

Since $\mathrm{Im}\, G'(x_0)$ is closed, there is an element $u \in X$ such that $G'(x_0)u = z$. Now we observe that if $x \in \mathrm{Ker}\, G'(x_0)$, then $g'(x_0)x = 0$ (and consequently if $G'(x_0)u_1 = G'(x_0)u_2$, then $g'(x_0)u_1 = g'(x_0)u_2$). Indeed, if $g'(x_0)x$ were different from 0, then for the real number

$$t := \frac{r - g'(x_0)u}{g'(x_0)x}$$

we would get

$$g'(x_0)(u + tx) = r \quad \text{and} \quad G'(x_0)(u + tx) = z,$$

that is, $(r, z) \in \mathrm{Im}\, f'(x_0)$. Therefore we can define a functional $l_1$ on $\mathrm{Im}\, G'(x_0)$ in the following way: if $z \in \mathrm{Im}\, G'(x_0)$ and $u \in X$ such that $G'(x_0)u = z$, then

$$l_1(z) := g'(x_0)u.$$

Obviously, $l_1$ is linear and $g'(x_0) = l_1 \circ G'(x_0)$. Moreover, $l_1$ is continuous: if $U \subset R$ is any open set, then $(g'(x_0))^{-1}(U)$ is open since $g'(x_0)$ is continuous, and

$$l_1^{-1}(U) = G'(x_0)[(g'(x_0)^{-1}(U)]$$

is open in the Banach space $\mathrm{Im}\, G'(x_0)$, because $G'(x_0): X \to \mathrm{Im}\, G'(x_0)$ is an open mapping. By the Banach—Hahn theorem there is an extension $l \in Z'$ of $l_1$, it satisfies (ii) with $\lambda = -1$, as $l \circ G'(x_0) = l \circ G'(x_0) = g'(x_0)$.

Remark. It is known (see [4]) that Theorem 4 implies various transversality conditions and Euler—Lagrange equations concerning the classical problems im the calculus of variations.

## References

[1] J. Dieudonné, *Foundations of Modern Analysis*, Academic Press (New York—London, 1960).
[2] W. Rudin, *Functional Analysis*, McGraw-Hill Book Company (New York, 1973).
[3] S. B. Nadler, Jr., Multivalued contraction mappings, *Pacific J. Math.*, 30 (1969), 475—488.
[4] А. Д. Иоффе—В. М. Тихомиров, *Теория экстремальных задач*, Изд. Наука (Москва, 1974).

ELTE TTK ANALÍZIS II
1445 BUDAPEST 8.
PF. 323.