ANNALES SOCIETATIS MATHEMATICAE POLONAE
Series I: COMMENTATIONES MATHEMATICAE XXIX (1989)
ROCZNIKI POLSKIEGO TOWARZYSTWA MATEMATYCZNEGO
Seria I: PRACE MATEMATYCZNE XXIX (1989)

TADEUSZ POREDA (Łódź)

On the geometrical properties of local diffeomorphisms in Banach spaces

Abstract. Let X, Y be Banach spaces. In this paper we are occupied in local diffeomorphisms of the class $\mathscr{C}^2$ which map an open set $\Omega \subset X$ into Y. It is shown that each local diffeomorphism of class $\mathscr{C}^2$ is locally convex in weakened sense.

Let X be a real Banach space with the norm $\|\cdot\|$. The ball $\{x \in X;\ \|x-x_0\| < r\}$ is denoted by $B(x_0, r)$, and we will use the abbreviations $B_r = B(0, r)$, $B = B_1$.

The class of all real continuous linear functionals on X regarded as a real linear space is denoted by X'. For each $x \in X-\{0\}$ we define the set

$$T(x) = \{x' \in X';\ \|x'\| = 1,\ x'(x) = \|x\|\};$$

the Hahn–Banach theorem guarantees that $T(x)$ is nonempty.

Let $Q_{\theta,r}$, for $0 < \theta \leqslant 1$ and $r > 0$, be the class of maps $h\colon B_r \times B_r \to X$ such that $x'(h(x, y)) > 0$ for $x' \in T(x)$ and $(x, y) \in B_r \times B_r$ such that $\|y\| < \theta\|x\|$. If $\theta = 1$ then the set $Q_{1,r}$ will be denoted by Q_r.

In the space X we introduce the semi-inner product as follows. Let us choose one nonzero element with the norm equal to 1 from each line in X containing the point $x = 0$ and denote the set of all chosen elements by X_0. Then, to each $y \in X_0$ let us assign any functional $J_0(y) \in T(y)$. We have thus defined the map $J_0\colon X_0 \to X'$. Let us extend that map onto X by putting $J(\lambda y) = \lambda J_0(y)$ for $y \in X_0$ and $\lambda \in \boldsymbol{R}$. Now, we can define the semi-inner product denoted by $\langle\cdot,\cdot\rangle$. For $x, y \in X$, we put $\langle x, y\rangle = J(y)(x)$.

It has the following properties:

(a) it maps $X \times X$ into $\boldsymbol{R}$,

(b) $\langle x+y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$, $\langle \lambda x, y\rangle = \lambda\langle x, y\rangle$, $\langle x, \lambda y\rangle = \lambda\langle x, y\rangle$ for $x, y, z \in X$ and $\lambda \in \boldsymbol{R}$,

(c) $\langle x, y\rangle = \|x\|^2$ for each $x \in X$,

(d) $|\langle x, y\rangle|^2 \leqslant \langle x, x\rangle\langle y, y\rangle$ for $x, y \in X$.

Let Ω be an open set in X and let Y be another real Banach space, and let $f\colon \Omega \to Y$.

DEFINITION 1. We shall say that the map f is *locally starlike* if for any $x_0 \in \Omega$ there exists a ball $B(x_0, r) \subset \Omega$ such that for any $0 < \varrho \leqslant r$ the set $f(B(x_0, \varrho))$ is starlike with respect to $f(x_0)$.

DEFINITION 2. We say that the map f is *locally convex* on Ω if for any $x_0 \in \Omega$ there exists $r > 0$ such that $B(x_0, r) \subset \Omega$ and for any $0 < \varrho \leqslant r$, $f(B(x_0, \varrho))$ is a convex set.

DEFINITION 3. The map f is called *locally convex in weakened sense* if for any $x_0 \in \Omega$ and $\theta \in (0, 1)$ there exists $r > 0$ such that for any $0 < \varrho \leqslant r$ and for $x, y \in B(x_0, \varrho)$ satisfying the condition $\theta \|x - x_0\| \geqslant \|y - x_0\|$,

$$tf(x) + (1-t)f(y) \in f(B(x_0, \varrho)) \quad \text{for } t \in [0, 1].$$

Remark. The following implications are true. If f is a locally convex map, then it is locally convex in weakened sense. If f is a locally convex in weakened sense map, then it is starlike.

THEOREM 1. *Let X, Y be Banach spaces and $\Omega \subset X$ be an open set. If $f\colon \Omega \to Y$ is a local diffeomorphism of the class $\mathscr{C}^2$, then it is a locally convex in weakened sense map.*

The proof of this theorem is preceded by three lemmas.

LEMMA 1. *Let $h\colon B_r \to X$ be a map of the class $\mathscr{C}^1$ bounded with the first derivative on each ball $B_\varrho \subset B_r$ for $0 < \varrho < r$. Assume that $x'(h(x)) > 0$ for any $x' \in T(x)$, where $x \in B_r$ and $\|x\| > r_0$ for a certain $r_0 < r$. Then, for any $x_0 \in B_r$ such that $\|x_0\| > r_0$ the differential equation*

$$\frac{dv}{dt}(t) = -h(v(t)), \quad v(0) = x_0 \tag{1}$$

possesses exactly one solution defined for $t \in [0, \infty)$ and such that $v(t) \in \bar{B}_{\|x_0\|}$ for $t \in [0, \infty)$.

In our further considerations, this solution will be denoted by $v = v(x_0, t)$ for $t \in [0, \infty)$.

Proof. Let $\|x_0\| = r_1$, $r_1 > r_0$ and let $r_2 > 0$ be a number such that $r_1 + r_2 < r$. Then $B(\tilde{x}, r_2) \subset B_{r_1 + r_2}$ for any $\tilde{x} \in \bar{B}_{r_1}$. By the assumption, there exist numbers K, L such that $\|h(x)\| \leqslant K$ and $\|Dh(x)\| \leqslant L$ for $x \in B_{r_1 + r_2}$. Let τ_0 be a number satisfying the conditions $0 \leqslant \tau_0 \leqslant \min(r_0/K, 1/L)$. By Theorem IX.2′ from [2] there exists exactly one solution $v = v(x_0, t)$ of equation (1) for $t \in [-\tau_0, \tau_0]$.

We prove that $v(x_0, t) \in \bar{B}_{\|x_0\|}$ for $t \in [0, \tau_0]$.

First, let us notice that

$$\frac{\partial}{\partial t}\|v(x_0, t)\|^2 = -2\langle h(v(x_0, t)), v(x_0, t)\rangle \tag{2}$$

for almost all $t\in[0, \tau_0]$ (see [1]).

Suppose now that $\varrho_0 = \sup_{0\leqslant t\leqslant \tau_0}\|v(x_0, t)\| > \|x_0\|$. Then there exists $t_1\in[0, \tau_0]$ such that $\|v(x_0, t_1)\| = \varrho_0$ and there exists $\delta > 0$ such that $t_1-\delta > 0$ and $\|x_0\| < \|v(x_0, t)\| \leqslant \varrho_0$ for $t\in[t_1-\delta, t_1]$. Since $\|v(x_0, t)\| > r_0$ for $t\in[t_1-\delta, t_1]$ and equality (2) holds, therefore from the assumptions about the function h it follows that $\frac{\partial}{\partial t}\|v(x_0, t)\| < 0$ for almost all $t\in[t_1-\delta, t_1]$. Hence it follows at once that $\|v(x_0, t_1-\delta)\| > \|v(x_0, t_1)\|$. This contradicts the definitions of t_1 and ϱ_0. Thus $v(x_0, t)\in\bar{B}_{\|x_0\|}$ for $t\in[0, \tau_0]$.

Let us denote $\tilde{x}_0 = v(x_0, \tau_0)$ and notice that $\tilde{x}_0\in\bar{B}_{\|x_0\|}$.

Now we shall consider the equation

$$\frac{dv}{dt}(t) = -h(v(t)), \quad v(\tau_0) = \tilde{x}_0. \tag{3}$$

Proceeding analogously as in the first part of the proof, we can show that this equation possesses a solution $v = v(\tilde{x}_0, t)$ for $t\in[\tau_0, 2\tau_0]$ then equation (1) possesses a solution $v = v(x_0, t)$ defined on $[0, 2\tau_0]$. Repeating the same argumentation infinitely many times we obtain that equation (1) possesses a solution $v = v(x_0, t)$ defined for $t\in[0,\infty)$. The uniqueness of this solution follows immediately by Theorem IX.2′ from [2].

LEMMA 2. *Let $f\colon B_r\to Y$ be a local diffeomorphism of the class $\mathscr{C}^2$ and let for some $y_0\in\bar{B}_{r_0}$ (where $0<r_0<r$) the map*

$$h(x, y_0) = \big(D(f(x))^{-1}(f(x)-f(y_0))\big)$$

defined for $x\in B_r$, satisfy the condition $x'(h(x, y_0)) > 0$, where $x'\in T(x)$, while $x\in B_r$ and $\|x\| > r_0$. Furthermore, assume that $h(x, y_0)$, as a function of the variable x, is bounded on every ball B_ϱ for $0<\varrho<r$.

Then, for any $x\in B_r$, $\|x\|\geqslant r_0$ and for $t\in[0, 1]$

$$tf(x)+(1-t)f(y_0)\in f(\bar{B}_{\|x\|}).$$

Proof. Under the above assumption, from Lemma 1 we get that the equation

$$\frac{\partial v}{\partial t}(x, t) = -h(v(x, t), y_0), \quad v(x, 0) = x,$$

where $x\in B_r$ and $\|x\|\geqslant r_0$, possesses exactly one solution $v = v(x, t)$ determined for $t\in[0, \infty)$ such that $v(x, t)\in\bar{B}_{\|x\|}$ for $t\in[0, \infty)$.

Since $h(x, y_0) = (Df(x))^{-1}(f(x)-f(y_0))$ for $x \in B_r$, therefore

$$\frac{\partial v}{\partial t}(x, t) = -\big(Df(v(x, t))\big)^{-1}\big(f(v(x, t))-f(y_0)\big), \quad v(x, 0) = x$$

for $t \in [0, \infty)$. Transforming this equality, we obtain

$$\frac{\partial}{\partial t} f(v(x, t)) = -f(v(x, t))+f(y_0), \quad v(x, 0) = x$$

for $t \in [0, \infty)$. Put

$$w_x(t) = f(v(x, t)) \quad \text{for } t \in [0, \infty).$$

Then the last differential equation takes form

$$\frac{d}{dt} w_x(t) = -w_x(t)+f(y_0), \quad w_x(0) = f(x),$$

for $t \in [0, \infty)$. Solving this equation (cf. [2]), we get that

$$w_x(t) = e^{-t} f(x)+(1-e^{-t}) f(y_0) \quad \text{for } t \in [0, \infty).$$

Hence

$$f(v(x, t)) = e^{-t} f(x)+(1-e^{-t}) f(y_0) \quad \text{for } t \in [0, \infty).$$

This implies at once that $e^{-t} f(x)+(1-e^{-t}) f(y_0) \in f(\bar{B}_{\|x\|})$ for $t \in [0, \infty)$, which completes the proof.

In our further investigations we shall consider the set $B_r \times B_r$ with a norm defined by the equality $\|(x, y)\| = \max(\|x\|, \|y\|)$ for $x, y \in B_r$.

LEMMA 3. *Let $h\colon B_r \times B_r \to X$ be a map of the class $\mathscr{C}^1$, $h(0, 0) = 0$, $Dh(0, 0)(x, y) = x-y$. Then, for any $\theta \in (0, 1)$ there exists $0 < \varrho < r$ such that $h \in Q_{\theta,\varrho}$.*

Proof. From the assumptions about the function h it follows that

$$h(x, y) = x-y+o(\|(x, y)\|),$$

where

$$\lim_{(x,y)\to(0,0)} \frac{\|o(\|(x, y)\|)\|}{\|(x, y)\|} = 0.$$

Hence, for any $\theta \in (0, 1)$ there exists $0 < \varrho < r$ such that $\|o(\|(x, y)\|)\| \leqslant \frac{1}{2}(1-\theta)\|(x, y)\|$ for $x, y \in B_\varrho$. Next, let us notice that

$$x'(h(x,y)) \geqslant \|x\| - \|y\| - \tfrac{1}{2}(1-\theta)\|(x,y)\|$$

for $x, y \in B_\varrho$ and $x' \in T(x)$.

Since $\|(x, y)\| = \|x\|$ for $x, y \in B_\varrho$ satisfying the condition $\theta\|x\| \geqslant \|y\|$ (where $0 < \theta < 1$), therefore $x'(h(x, y)) \geqslant \frac{1}{2}\|x\|(1-\theta)$. As a consequence, for $x, y \in B_\varrho$ such that $\theta\|x\| > \|y\|$, $x'(h(x, y)) > 0$, where $x' \in T(x)$. Hence $h \in Q_{\theta,\varrho}$ for any $0 < \theta < 1$ and for some $\varrho > 0$.

Proof of Theorem 1. Let us first consider a function

$$h(x, y) = (Df(x))^{-1}(f(x)-f(y))$$

defined for $x, y \in \Omega$. This function has continuous partial derivatives with respect to variables x and y then it is of the class $\mathscr{C}^1$ on $\Omega \times \Omega$.

Let $x_0 \in \Omega$ and $r > 0$ be a number such that $B(x_0, r) \subset \Omega$ and the function h and its derivative $\partial h/\partial x$ are bounded on $B(x_0, r) \times B(x_0, r)$.

We shall show that $Dh(x_0, x_0)(x, y) = x-y$ for $x, y \in X$ and $x_0 \in \Omega$.

Let us notice that for any $x, y \in B(x_0, t)$ and for $|t| < 1$

$$x_0 + t(x-x_0), \qquad x_0 + t(y-x_0) \in B(x_0, r)$$

and consider an auxiliary function

$$F(t) = h\big(x_0+t(x-x_0),\, x_0+t(y-x_0)\big) \quad \text{for } |t| < 1,$$

where $x, y \in B(x_0, r)$.

From the definition of the function F it follows that

$$F'(0) = Dh(x_0, x_0)(x-x_0, y-x_0).$$

On the other hand, taking account of the definition of the function h, we have

$$Df\big(x_0+t(x-x_0)\big)(F(t)) = f\big(x_0+t(x-x_0)\big) - f\big(x_0+t(y-x_0)\big)$$

for $x, y \in B(x_0, r)$ and $|t| < 1$. Differentiating the above equality with respect to the parameter t and taking $F(0) = 0$ into account, we get that $F'(0) = x-y$, which implies that

$$Dh(x_0, x_0)(x-x_0, y-x_0) = x-y$$

for $x, y \in B(x_0, r)$.

Put $x-x_0 = \tilde{x}$, $y-x_0 = \tilde{y}$, then $Dh(x_0, x_0)(\tilde{x}, \tilde{y}) = \tilde{x}-\tilde{y}$ for any $\tilde{x}, \tilde{y} \in B_r$. By linearity of the operator $Dh(x_0, x_0)$, we obtain that

$$Dh(x_0, x_0)(\tilde{x}, \tilde{y}) = \tilde{x}-\tilde{y} \quad \text{for } \tilde{x}, \tilde{y} \in X.$$

Next, let us observe that the function $\tilde{h}(x, y) = h(x+x_0, y+x_0)$, for $x, y \in B_r$, fulfils the assumptions of Lemma 3. Hence, by this lemma, for each $\theta \in (0, 1)$ there exists $\varrho_0 > 0$ such that $\tilde{h} \in Q_{\theta,\varrho_0}$.

With the above notation, we have

$$\tilde{h}(x, y) = \big(Df(x_0+x)\big)^{-1}\big(f(x_0+x)-f(x_0+y)\big)$$

for $x,\ y \in B_r$.

From Lemma 2 it follows that if $x, y \in B_\varrho$, where $0 < \varrho < \varrho_0$, and $\|y\| \leqslant \theta\|x\|$, then

$$\tau f(x_0+x)+(1-\tau)f(x_0+y) \in f\big(B(x_0, \varrho)\big)$$

for $\tau \in [0, 1]$. This proves that f is locally convex in weakened sense.

On account of the remark, the following corollary is obvious.

COROLLARY. *Let X, Y be Banach spaces and $\Omega \subset X$ be an open set. If $f\colon \Omega \to Y$ is a local diffeomorphism of the class $\mathscr{C}^2$, then f is locally starlike.*

The natural question arises whether Theorem 1 is true in a stronger form, namely, if local convexity in weakened sense can be replaced by local convexity. An answer to this question is given by the following example.

EXAMPLE. Let $X = Y = \boldsymbol{R}^2$, $\|x\| = \max(|x_1|, |x_2|)$ for $x = (x_1, x_2) \in \boldsymbol{R}^2$. Consider a map $f\colon B \to \boldsymbol{R}^2$ defined as follows

$$f(x) = (x_1+\tfrac{1}{2}x_2^2,\ x_2) \qquad \text{for } x = (x_1, x_2) \in B.$$

It is not difficult to prove that this map is diffeomorphism of the class $\mathscr{C}^2$, but for each $r \in (0, 1)$, $f(B_r)$ is not a convex set. Hence, the answer to the imposed question is negative.

Next, we formulate a theorem being the criterion of local convexity.

THEOREM 2. *Let X, Y be Banach spaces and let $\Omega \subset X$ be an open set. Let $f\colon \Omega \to Y$ be a local diffeomorphism of the class $\mathscr{C}^2$. If for any $x_0 \in \Omega$ there exists $r > 0$ such that $B(x_0, r) \subset \Omega$ and*

$$w_{x_0}(x, y) = \big(Df(x_0+x)\big)^{-1}\big(f(x_0+x)-f(x_0+y)\big),$$

for $x,\ y \in B_r$, belongs to Q_r, then the map f is locally convex.

We precede the proof of this theorem by the following lemma.

LEMMA 4. *Let $f\colon B_r \to Y$ be a local diffeomorphism of the class $\mathscr{C}^2$ and let a function*

$$h(x, y) = \big(Df(x)\big)^{-1}\big(f(x)-f(y)\big),$$

defined for $(x, y) \in B_r \times B_r$, belong to Q_r. Moreover, assume that the function h and its partial derivative with respect to the variable x are bounded on $B_\varrho \times B_\varrho$ for any $\varrho \in (0, r)$. Then $f(B_\varrho)$ is a convex set for any $\varrho \in (0, r)$.

Proof. Let x_0, y_0 be fixed points from B_r such that $\max(\|x_0\|, \|y_0\|) = \|x_0\| = r_0$. From the assumption that $h \in Q_r$ we have that $x'(h(x, y_0)) > 0$ for $x' \in T(x)$, $x \in B_r$, and $\|x\| > r_0$. By Lemma 1 it follows that the equation

$$\frac{dv}{\partial t}(t) = -h(v(t), y_0), \qquad v(0) = x_0,$$

possesses exactly one solution defined for $t \in [0, \infty)$. (Further this solution will be denoted by $v = v(x_0, y_0, t)$ for $t \in [0, \infty)$.) From Lemma 1 we get also that $v(x_0, y_0, t) \in \bar{B}_{\|x_0\|}$ for $t \in [0, \infty)$. By regard to Lemma 2 we obtain that $tf(x_0) + (1-t)f(y_0) \in f(\bar{B}_{\|x_0\|})$ for $t \in [0, 1]$. From the above and by free choice of x_0 and y_0 it follows that $f(B_\varrho)$ is a convex set for any $\varrho \in (0, r)$.

Proof of Theorem 2. In accordance with the assumptions for any $x_0 \in \Omega$ there exists $r > 0$ such that $B(x_0, r) \subset \Omega$ and $w_{x_0} \in Q_r$. So defined map w_{x_0} is of the class $\mathscr{C}^1$ on $B_r \times B_r$. Hence there exists $\varrho_0 \in (0, r)$ such that this map and its partial derivative with respect to the variable x are bounded on $B_{\varrho_0} \times B_{\varrho_0}$.

Let $x, y \in B(x_0, \varrho_0)$; then they can be represented in the form $x = x_0 + \tilde{x}$, $y = x_0 + \tilde{y}$, where $\tilde{x}, \tilde{y} \in B_{\varrho_0}$. Let us consider an auxiliary function

$$\tilde{f}(\tilde{x}) = f(x_0, \tilde{x}) - f(x_0) \quad \text{for } \tilde{x} \in B_{\varrho_0}.$$

With the above notation the map

$$(D\tilde{f}(\tilde{x}))^{-1}(\tilde{f}(\tilde{x}) - \tilde{f}(\tilde{y})), \quad \text{where } \tilde{x}, \tilde{y} \in B_{\varrho_0},$$

belongs to Q_{ϱ_0}. Hence by Lemma 4, $\tilde{f}(B_\varrho)$ is a convex set for $\varrho \in (0, \varrho_0)$, therefore we get at once that $f(B(x_0, \varrho))$ is a convex set for any $\varrho \in (0, \varrho_0)$. By free choice of $x_0 \in \Omega$ it follows local convexity of the map f on Ω.

References

[1] T. Kato, *Nonlinear semigroups and evolution equations*, J. Math. Soc. Japan 19 (1967), 508–520.

[2] K. Maurin, *Analiza*, part I, PWN, Warszawa 1971.