# A NOTE ON INVERSE FUNCTION THEOREMS

E. B. LEACH

In this note we use the Newton-Raphson approach to inverse function theorems. We draw natural conclusions when only a left or a right inverse to the differential at a point is available. By using a strengthened version of differential, we are able to use differentiability at a single point as the smoothness condition. Although the method has been used before (cf. [2, p. 167 ff.]), analysis books have tended to use an approach that assumes finite dimensionality of the reference spaces.

DEFINITION. Let $U$ and $V$ be Banach spaces and $f\colon U \to V$ a function. A *strong differential* of $f$ at a point $x_0$ in $U$ is a bounded linear transformation $\alpha\colon U \to V$ which approximates changes of $f$ in the sense that for every $\epsilon > 0$, there is a $\delta > 0$ such that if $x'$ and $x''$ are nearer than $\delta$ to 0, then:

$$|f(x') - f(x'') - \alpha(x' - x'')| \leqq \epsilon |x' - x''|.^{1} \tag{1}$$

A relation such as (1) for any $\epsilon$ and $\delta$ implies a Lipschitz condition, and so continuity of $f$ in the sphere $\{|x - x_0| < \delta\}$. If $f$ should have a Fréchet differential $\alpha_1$ at a point $x_1$ of that sphere, it is clear that $|\alpha_1 - \alpha| \leqq \epsilon$. The following theorem is specialized by considering $x_0 = 0$ in $U$, and by the condition that $f(0) = 0$. It is evident that a more general situation reduces to this by translation of range and domain; however, the statement and the proof of the theorem are a little more complicated in the general case.

THEOREM. *Let $f\colon U \to V$ be a function such that $f(0) = 0$ and $f$ has strong differential $\alpha$ at 0. Let $\beta\colon V \to U$ be a bounded linear transformation such that $\beta\alpha\beta = \beta$. Then there is a function $g\colon V \to U$ such that $g(0) = 0$, $g$ has strong differential $\beta$ at 0, and $g$ satisfies for $y$ near 0 the identities:*

$$\beta(f(g(y))) = \beta(y), \tag{2'}$$

$$\beta(\alpha(g(y))) = g(y), \tag{2''}$$

$$g(f(\beta(y))) = \beta(y). \tag{2'''}$$

*Any two $g$'s satisfying these conditions are identical for $y$ near 0.*

Presented to the Society, February 25, 1961; received by the editors April 5, 1960 and, in revised form, November 3, 1960.

[1] We will use the notation $|x|$ or $|\alpha|$ for the norms of vectors or of linear transformations.

PROOF. For any $y$ in $V$, sequences $\{x_n\}$ and $\{y_n\}$ are defined by $x_0=0$ and the recursion formulas:

$$y_n = f(x_n), \qquad x_{n+1} = x_n + \beta(y - y_n) \qquad (n = 0, 1, 2, \cdots). \tag{3}$$

We define $g(y)$ to be the limit of $\{x_n\}$ for all $y$ such that the limit exists, and to be 0 otherwise. For all $y$ sufficiently near to 0, $g(y)$ is the limit of $\{x_n\}$.

To prove convergence of $\{x_n\}$ for $y$ near 0, consider $\epsilon>0$ such that $\epsilon|\beta|<1/2$ and $\delta>0$, chosen so that (1) is valid for $f$ at the origin of $U$. We prove inductively that $|x_n|<\delta$ and that $|x_{n+1}-x_n|<(\epsilon|\beta|)^n\delta/2$, provided $|y|<\delta/2|\beta|$, for $n=0, 1, 2, \cdots$. The bound on $|x_1-x_0|$ is an immediate consequence of (3), for $|y|<\delta/2|\beta|$. If $|x_{m+1}-x_m| \leqq (\epsilon|\beta|)^m\delta/2$ for $m\leqq n-1$, then $|x_n| \leqq \sum_{m=0}^{n-1} (\epsilon|\beta|)^m\delta/2<\delta$, since $\epsilon|\beta|<1/2$. Now we have by (3) and the relation $\beta\alpha\beta=\beta$, for $n\geqq 1$:

$$\begin{aligned} x_{n+1} - x_n &= \beta(y - y_n) = \beta(y - y_{n-1}) - \beta(y_{n-1} - y_n) \\ &= \beta\alpha\beta(y - y_{n-1}) - \beta(y_{n-1} - y_n) \\ &= \beta[\alpha(x_n - x_{n-1}) - f(x_{n-1}) + f(x_n)]. \end{aligned} \tag{4}$$

The vector in brackets is essentially the vector of the left side of (1), with $x'$ and $x''$ replaced by $x_n$ and $x_{n-1}$, and since these fall within $\delta$ of 0, (1) implies $|x_{n+1}-x_n| \leqq |\beta| \cdot \epsilon |x_n - x_{n-1}|$, from which the asserted bound on $|x_{n+1}-x_n|$ follows by induction. The sequence $\{x_n\}$ therefore converges uniformly to $g(y)$ if $|y|<\delta/2|\beta|$.

The identities (2′) and (2″) are readily derived for all $y$ such that $|y|<\delta/2|\beta|$. For (2′), let $n\to\infty$ in (3), and taking account of the fact that $f(x_n)\to f(g(y))$ by continuity of $f$, we have $g(y)=g(y)+\beta(y-f(g(y)))$, which simplifies immediately to $\beta fg=\beta$ for $|y|<\delta/2|\beta|$. For (2″), $\beta\alpha(x_{n+1}-x_n)=\beta\alpha\beta(y-y_n)=\beta(y-y_n)=x_{n+1}-x_n$, and inductively it follows that $\beta\alpha x_n=x_n$. Making $n\to\infty$, $\beta\alpha g(y)=g(y)$ for all $y$ such that $\{x_n\}$ converges (and for all other $y$, since then $g(y)=0=\beta\alpha g(y)$).

Each of the remaining steps of proof involves estimating $x'-x''$ for some choice of $x'$, $x''$ nearer than $\delta$ to 0 in $U$. In each case a differential relation involving $\beta(f(x'))-\beta(f(x''))$ is invoked. First we prove a Lipschitz condition on $g$. If $y'$ and $y''$ are nearer than $\delta/2|\beta|$ to 0 in $V$, $g(y')$ and $g(y'')$ are nearer than $\delta$ to 0 in $U$, and by (1), with this choice of $x'$ and $x''$:

$$|f(g(y')) - f(g(y'')) - \alpha(g(y') - g(y''))| \leqq \epsilon|g(y') - g(y'')|.$$

Application of $\beta$ to the vector in brackets leads to the relation:

$$|\beta(f(g(y'))) - \beta(f(g(y''))) - \beta\alpha(g(y') - g(y''))| \leqq \epsilon|\beta| \, |g(y') - g(y'')| .$$

The vector on the left simplifies, because $\beta f g = \beta$ and $\beta\alpha g = g$, to:

$$|\beta(y' - y'') - (g(y') - g(y''))| \leqq \epsilon|\beta| \, |g(y') - g(y'')| .$$

By this relation and the triangle inequality we obtain $|\beta(y'-y'')| \geqq |g(y') - g(y'')| - \epsilon|\beta| \, |g(y') - g(y'')|$, and since $|\beta(y' - y'')| \leqq |\beta| \, |y'-y''|$ and $1-\epsilon|\beta| > 0$, we obtain the Lipschitz condition valid if $y'$ and $y''$ are nearer than $\delta/2|\beta|$ to 0:

$$(5) \qquad |g(y') - g(y'')| \leqq |\beta| \, |y' - y''|/(1 - \epsilon|\beta|).$$

Instead of proving differentiability of $g$ at 0 only, we prove differentiability of $g$ at any $y$ in $V$ such that $|y| < \delta/2|\beta|$, provided $f$ is differentiable at $g(y)$ with differential $\alpha_1$. For $y'$ and $y''$ near to $y$ and within $\delta/2|\beta|$ of 0, $g(y')$ and $g(y'')$ are within $\delta$ of 0 and near to $g(y)$. For any $\epsilon_1 > 0$, differentiability of $f$ at $g(y)$ implies the relation on $f$ valid for $y'$, $y''$ near $y$:

$$|fg(y') - f(g(y'')) - \alpha_1(g(y') - g(y''))| \leqq \epsilon_1|g(y') - g(y'')| .$$

Application of $\beta$ on the left gives:

$$|\beta(f(g(y'))) - \beta(f(g(y''))) - \beta\alpha_1(g(y') - g(y''))| \leqq \epsilon_1|\beta| \, |g(y') - g(y'')| .$$

As before, we apply $\beta f g = \beta$ to simplify the first two terms. We write $\beta\alpha_1 = \beta\alpha + \beta(\alpha_1 - \alpha)$ and since $\beta\alpha g = g$, we may replace $\beta\alpha$ by 1, leading to:

$$|\beta(y' - y'') - (1 + \beta(\alpha_1 - \alpha))(g(y') - g(y''))| \leqq \epsilon_1|\beta| \, |g(y') - g(y'')| .$$

Now it was observed in the remarks following the definition of strong differential that $|\alpha_1 - \alpha| \leqq \epsilon$, so $|\beta(\alpha_1 - \alpha)| < \epsilon|\beta| < 1/2$. So $1 + \beta(\alpha_1 - \alpha)$ has a bounded inverse of norm $\leqq 1/(1-\epsilon|\beta|)$. Applying $-(1+\beta(\alpha_1-\alpha))^{-1}$ to the vector on the left above and changing the order of the terms gives:

$$(6) \qquad \begin{aligned} &|g(y') - g(y'') - [1 + \beta(\alpha_1 - \alpha)]^{-1}(\beta(y' - y''))| \\ &\qquad \leqq \epsilon_1|\beta| \, |g(y') - g(y'')|/(1 - \epsilon|\beta|). \end{aligned}$$

But by (5) the right side may be replaced by $\epsilon_1|\beta|^2|y'-y''|/(1-\epsilon|\beta|)^2$, if $y'$ and $y''$ are near to $y$. Since the coefficient of $|y'-y''|$ approaches 0 as $\epsilon_1 \to 0$, $g$ is differentiable at $y$ with differential $[1+\beta(\alpha_1-\alpha)]^{-1}\beta$.[2] In particular, this reduces to $\beta$ at $y=0$.

[2] By letting $y''=y$ the same proof gives the existence of a Fréchet differential provided $f$ has a Fréchet differential at $g(y)$.

We now prove the third identity ($2'''$). If $y$ is near to 0, both $y$ and $f(\beta(y))$ are within $\delta/2|\beta|$ of 0. We wish to compare $g(f(\beta(y)))$ and $\beta(y)$, which therefore serve as the $x'$ and $x''$ in this case. The relation about $\beta f$ is:

$$|\beta(f(g(f(\beta(y))))) - \beta(f(\beta(y))) - \beta\alpha(g(f(\beta(y))) - \beta(y))| \leqq \epsilon|\beta|\ |g(f(\beta(y))) - \beta(y)|.$$

Now $\beta fg = \beta$ cancels the first two terms on the left, while $\beta\alpha g = g$ and $\beta\alpha\beta = \beta$ causes the relation to simplify to $|g(f(\beta(y))) - \beta(y)| \leqq \epsilon|\beta|\ |g(f(\beta(y))) - \beta(y)|$, a contradiction unless $g(f(\beta(y))) = \beta(y)$, since $\epsilon|\beta| < 1/2$.

There remains only to prove uniqueness of $g$; if $g'$ has similar properties, then $g'(y) = g(y)$ in a neighborhood of 0. We consider $y$ so near to 0 that $g$ and $g'$ satisfy ($2'$) and ($2''$), and so that $|g(y)|$ and $|g'(y)|$ are nearer than $\delta$ to 0 in $U$. Then, using the same method as before, we obtain the relation:

$$|\beta(f(g(y))) - \beta(f(g'(y))) - \beta\alpha(g(y) - g'(y))| \leqq \epsilon|\beta|\ |g(y) - g'(y)|$$

The first two terms cancel, for $\beta(f(g(y))) = \beta(y) = \beta(f(g'(y)))$. Then $\beta\alpha g = g$ and $\beta\alpha g' = g'$, so the expression simplifies to $|g(y) - g'(y)| \leqq \epsilon|\beta|\ |g(y) - g'(y)|$, and so $g(y) = g'(y)$.

REMARK. The equation $\beta\alpha\beta = \beta$ in the hypothesis and the equations ($2'$), ($2''$), ($2'''$) are cases where $\alpha\beta$, $\beta\alpha$, $fg$ or $gf$ can be cancelled locally from a triple composite mapping $V$ into $U$. The list of identities of this sort is completed with $g\alpha\beta = g$ and $gfg = g$. These may be proved in the same way that ($2'''$) is proved. In contrast, if we propose such an identity for a triple composition from $U$ to $V$, we always have $\beta = 0$ as a possible choice of $\beta$, and the resulting $g$ is then 0. The triple composition then gives 0, while the cancellation gives $\alpha$ or $f$, which may be different from 0.

COROLLARY. *If in particular $\beta\alpha$ is the identity of $U$ or $\alpha\beta$ is the identity of $V$, then correspondingly $g(f(x)) = x$ for $x$ near* 0, *or $f(g(y)) = y$, for $y$ near* 0.

PROOF. In the first case $g(f(x)) = g(f(\beta(\alpha(x)))) = \beta(\alpha(x)) = x$, using ($2'''$). In the second case, $f(g(y)) = \alpha(\beta(f(g(y)))) = \alpha(\beta(y)) = y$, using ($2'$).

## REFERENCES

1. T. M. Apostol, *Mathematical analysis*, Reading, Massachusetts, Addison-Wesley, 1957.

2. L. V. Kantorovich, *Functional analysis and applied mathematics* (translated from the Russian by Curtis D. Benster), National Bureau of Standards Report, 1952.

CASE INSTITUTE OF TECHNOLOGY