GLOBAL INVERSE FUNCTION THEOREM

by

Felix F. Wu and Charles A. Desoer

Memorandum No. ERL-M300

23 April 1971

Electronics Research Laboratory

College of Engineering
University of California, Berkeley
94720

April 22, 1971

GLOBAL INVERSE FUNCTION THEOREM

by

Felix F. Wu and Charles A. Desoer

Department of Electrical Engineering and Computer Sciences
and the Electronics Research Laboratory,
University of California, Berkeley, California 94720

ABSTRACT

A simple proof of global inverse function theorem in $\mathbb{R}^n$ is given. A global homeomorphic version of the theorem is proved first. A global diffeomorphic version follows by an application of the classical local inverse function theorem.

Research sponsored by National Science Foundation, Grant GK-10656X.

The problem of determining that a given function from $\mathbb{R}^n$ into $\mathbb{R}^n$ has an inverse is very useful in applications. In 1959, Palais established the necessary and sufficient condition for a function to be a diffeomorphism of R^n onto itself, which appeared as an episode in a paper[1] dealing with the determination of spaces of intertwining operators on differential forms. This global version of the classical inverse function theorem has been applied widely in nonlinear network theory[2-12] and is generally referred to as Palais Theorem by circuit theorists. Palais originally stated it without proof as a corollary in [1]. We believe that a simple proof of this useful theorem will be helpful to the readers of this Journal.

This note presents a proof that is intuitively appealing and easily understood with a modest background in mathematical analysis[13,14]. We first prove the necessary and sufficient condition for a global homeomorphism[2]; the case of a global diffeomorphism follows easily by an application of the classical inverse function theorem.

THEOREM

Let f be a map from $\mathbb{R}^n$ into $\mathbb{R}^n$, then f is a homeomorphism[1] of $\mathbb{R}^n$ onto $\mathbb{R}^n$ if and only if f is

(1) a local homeomorphism[2] and

(2) a proper map[3].

Proof: ⇒ By assumption f is a (global) homeomorphism, hence it is a local homeomorphism. Because f^{-1} is continuous, it maps any compact set into a compact set. [13, p. 78; 14, Theorem 4.1, p. 207].

⇐ We prove this in three steps: (1) f is surjective (onto), (2) f is injective (one-to-one), (3) f^{-1} is continuous. To facilitate the presentation, we denote the domain of f by X and the range by Y; of course, $X = Y = \mathbb{R}^n$.

(1) Surjective: Let Y_1 be the image of f, i.e., $Y_1 = f(X)$, or more specifically, $Y_1 = \{y \in Y | f^{-1}(y)$ is a nonempty subset of $X\}$. We know that Y is connected. If Y_1 is both open and closed, knowing also the fact that Y_1 is not empty, we can conclude that $Y_1 = Y$ [13, p. 59], i.e., f is surjective.

---

1. A homeomorphism of X onto Y is, by definition, a continuous bijective map f: $X \to Y$ such that f^{-1} is also continuous.

2. A map f: $X \to Y$ is said to be a local homeomorphism if whenever $x \in X$ and $y \in Y$ are such that $f(x) = y$ then there exist open neighborhoods U of x and V of y such that f restricted to U is a homeomorphism of U onto V.

3. A continuous map is said to be proper if the inverse image of any compact set is compact.

(i) Y_1 open: Let $y_1 \in Y_1$, then there exists an $x_1 \in X$ such that $f(x_1) = y_1$. Now f is a local homeomorphism means that there exist open neighborhoods U of x_1 and V of y_1 such that f is a homeomorphism from U onto V. So $V \subset Y_1$ and Y_1 is thus open.

(ii) Y_1 closed: Let y be an accumulation point of Y_1, then there exists a sequence $\{y_i\}_1^\infty$ with $y_i \in Y_1$ $\forall i$, and $y_i \to y$. Consider $K = \{y_i\}_1^\infty \cup \{y\}$, which is clearly closed and bounded in Y, hence compact. [13, p. 58; 14 Theorem 4.5, p. 208]. By assumption, $f^{-1}(K)$ is compact in X. Now pick $x_i \in f^{-1}(y_i)$. $\{x_i\}_1^\infty$ is a sequence in a compact set $f^{-1}(K)$ in a metric space $\mathbb{R}^n$, therefore $\{x_i\}_1^\infty$ has a convergent subsequence, say $\{x_{i_j}\}_{j=1}^\infty \to x$. [13, p. 56; 14, Th. 4-4, p. 208]. But $\{f(x_{i_j})\}_{j=1}^\infty$ is a subsequence of $\{y_i\}_1^\infty$, therefore converges to the same limit y. f is a continuous map because it is a local homeomorphism, hence $f(x) = f(\lim_{j\to\infty} x_{i_j}) = \lim_{j\to\infty} f(x_{i_j}) = y$, therefore $y \in Y_1$. Hence Y_1 is closed. [13, p. 47; 14, p. 203].

(2) Injective: Suppose that f is not injective, hence there exist two distinct points x_1, x_2 such that $f(x_1) = f(x_2)$. Without loss of generality, we can assume $f(x_1) = f(x_2) = 0$. Let α: $[0,1] \to X$ be defined by $\alpha(t) = (1-t)x_1 + tx_2$ and $\beta = f \circ \alpha$. Geometrically, α is the line segment joining x_1 to x_2 and β, its image in Y under f, is a closed curve through 0. (Fig. 1). Let B: $[0,1] \times [0,1] \to Y$ be defined by $B(t,\tau) = (1-\tau)\beta(t)$. Thus for each τ, $B(\cdot,\tau)$ is obtained by shrinking the closed curve β toward the origin. The rough idea of the proof is to shrink the curve β toward the origin, the corresponding curve α will be continuously deformed into some curve joining x_1 to x_2; the contradiction will be reached in the limit when β degenerates into a single point.

(i) Construction of the inverse image of B (Fig. 2): Let us define for each t a map $A(t,\cdot)$: $[0,1] \to X$ by the following process of piecing together the local inverses of B. First let $A(t,0) = \alpha(t)$. Since f is a local homeomorphism, there exist homeomorphic neighborhoods of $\alpha(t)$ and $\beta(t)$, U_1 and V_1 respectively. Define $A(t,\cdot)$: $[0,\tau_1] \to U_1$ to be the local inverse image of $B(t,\tau)$ for $\tau \in [0,\tau_1]$ where τ_1 is so chosen that $B(t,\tau) \in V_1$, for all $\tau \in [0,\tau_1]$. Thus we have $f(A(t,\tau_1)) = B(t,\tau_1)$ and we can define $A(t,\cdot)$ on $[\tau_1,\tau_2]$ with $\tau_2 > \tau_1$ as the local inverse of $B(t,\cdot)$ around $B(t,\tau_1)$. Repeat the same procedure; at each step, we extend τ from τ_k to τ_{k+1} with $\tau_{k+1} > \tau_k$. We are going to show by contradiction that the domain of $A(t,\cdot)$ can always be extended to include 1. Suppose that the above process fails to do so. Then the increasing sequence $\{\tau_k\}$ is bounded by 1 and has a least upper bound T, so $\{\tau_k\} \to T \leq 1$. But $B(t,\cdot)$ is continuous, $\lim_{k\to\infty} B(t,\tau_k) = B(t,T)$; and since $\{A(t,\tau_k)\}_{k=1}^{\infty}$ is a sequence in a compact set $f^{-1}(B(t,\tau) : \tau \in [0,1])$, it has a subsequence converging to a limit $A(t,T)$. Now because f is continuous, $f(A(t,T)) = B(t,T)$, hence the domain of $A(t,\cdot)$ is extended to include T; moreover, in the case when $T < 1$, it can even be extended beyond T by local homeomorphism. Thus, we can define a map A: $[0,1] \times [0,1] \to X$ with the property that $f \circ A = B$, and also $A(0,\tau) = x_1$, $A(1,\tau) = x_2$, $\forall\tau$.

(ii) Continuity of $A(\cdot,\tau)$: We will show that for each τ, $A(\cdot,\tau)$: $[0,1] \to X$ is continuous by open-set arguments [13, p. 70; 14, pp. 201-202]. Let $\mathcal{O}$ be any open set in X and let its inverse image under $A(\cdot,\tau)$ be denoted by $\mathcal{T}$; equivalently, $\mathcal{T}$ is the inverse image under $A(\cdot,\tau)$ of the intersection of $\mathcal{O}$ with the image of $A(\cdot,\tau)$. Now for each t, let the homeomorphic neighborhoods of $A(t,\tau)$ and $B(t,\tau)$ be U_t and V_t, respectively. Note that

$\bigcup_{t\in[0,1]} (U_t \cap \mathcal{O})$ has $\mathcal{T}$ as its inverse image under $A(\cdot,\tau)$ and $\bigcup_{t\in[0,1]} f(U_t \cap \mathcal{O})$ also has $\mathcal{T}$ as its inverse image under $B(\cdot,\tau)$ since $B = f\circ A$. But $f(U_t \cap \mathcal{O})$ are open in Y, so does $\bigcup_{t\in[0,1]} f(U_t \cap \mathcal{O})$. Because $B(\cdot,\tau)$ is continuous, $\mathcal{T}$ is open in [0,1]. This completes our proof that $A(\cdot,\tau)$ is continuous.

Now for $\tau = 1$, $B(t,1) = 0$, $\forall t$. Geometrically, this is done by shrinking β to the origin. The corresponding $A(t,1)$ is still a continuous curve joining two distinct points x_1 and x_2. But the inverse image of a single point under a local homeomorphism f can not be a continuous curve. To demonstrate this, suppose it were true, every neighborhood of x_1 would contain points of $f^{-1}(0)$ other than x_1 itself, then it would be impossible for homeomorphic neighborhoods of x_1 and 0 to exist. Thus we have proved that f is injective.

Remark: Here we have in fact tacitly constructed a covering homotopy A of B (15, Th. 3, p. 59).

(3) <u>Continuity of f^{-1}</u>: Recall that continuity is a local property. [13, p. 68; 14, pp. 201-202]. The fact that f^{-1} exists globally (by (1) and (2)) together with the local homeomorphism assumption asserts that f^{-1} is continuous. Q.E.D.

Lemma 1

Let f be a continuous map from $\mathbb{R}^n$ into $\mathbb{R}^n$, then f is a proper map if and only if $\lim_{\|x\|\to\infty} \|f(x)\| = \infty$.

Proof: $\Rightarrow$ By contradiction. Suppose that there is a sequence $\{x_k\}$ with $\|x_k\| \to \infty$, yet $\|f(x_k)\| \leq M < \infty$. Consider the closed and bounded ball

$B_M = \{y \mid \|y\| \leq M\}$, because f is proper, $f^{-1}(B_M)$ is compact. However, $\{x_k\}$ is contained in $f^{-1}(B_M)$, but $\|x_k\| \to \infty$ contradicts the compactness of $f^{-1}(B_M)$.

$\Leftarrow$ f is continuous implies that for each closed set K, $f^{-1}(K)$ is closed. Suppose K is bounded yet $f^{-1}(K)$ is not, then there exists a sequence $\{x_k\}$ in $f^{-1}(K)$ with $\|x_k\| \to \infty$. Clearly $\{f(x_k)\} \subset K$. But by assumption $\|f(x_k)\| \to \infty$, which contradicts boundedness of K. Q.E.D.

Lemma 2

Let f be a C^k map $(k \geq 1)$ from $\mathbb{R}^n$ into $\mathbb{R}^n$, then f is a local C^k-diffeomorphism[4] if and only if $\det(\frac{\partial f}{\partial x}) \neq 0$.

Proof: This is the well-known classical local inverse function theorem. [13, p. 211 and Ex. 17, p. 217; 14, p. 167].

Corollary:

Let f be a C^k map from $\mathbb{R}^n$ into $\mathbb{R}^n$, then f is a C^k-diffeomorphism if and only if

(1) $\det(\frac{\partial f}{\partial x}) \neq 0 \quad \forall x$

(2) $\lim_{\|x\| \to \infty} \|f(x)\| = \infty$

Proof: It follows from the Theorem, Lemma 1 and 2, as well as the fact that differentiability is a local property. [13, p. 198; 14, p. 142].

Q.E.D.

4. A C^k-map is, by definition, a map with continuous derivatives up to order k. A C^k-diffeomorphism is, by definition, a bijective C^k map such that the inverse is also C^k.

The authors would like to thank Prof. A. Weinstein for his helpful discussion.

REFERENCES

1. R. S. Palais, "Natural operations on differential forms," Trans. Amer. Math. Soc., vol. 92, no. 1, 1959, pp. 125-141.

2. C. A. Holzmann and R. Liu, "On the dynamical equations of nonlinear networks with n-coupled elements," Proc. 3rd Ann. Allerton Conf. Circuit and System Theory, 1965, pp. 536-545.

3. T. Ohtsuki and H. Watanabe, "State-variable analysis of RLC networks containing nonlinear coupling elements," IEEE Trans. Circuit Theory, vol. CT-16, no. 1, Feb. 1969, pp. 26-38.

4. I. W. Sandberg and A. N. Willson, Jr., "Some theorems on properties of DC equations of nonlinear networks." Bell Syst. Tech. J., vol. 48, no. 1, Jan. 1969, pp. 1-34.

5. I. W. Sandberg, "Theorems on the analysis of nonlinear transistor networks," Bell Syst. Tech. J., vol. 49, no. 1, Jan. 1970, pp. 95-114.

6. I. W. Sandberg, "Theorems on the computation of the transient responses of nonlinear networks containing transistors and diodes," Bell Syst. Tech. J., vol. 49, no. 8, Oct. 1970, pp. 1739-1776.

7. I. W. Sandberg, "Necessary and sufficient conditions for the global invertibility of certain nonlinear operators that arise in the analysis of networks," IEEE Trans. Circuit Theory, vol. CT-18, no. 2, Mar. 1971, pp. 260-263.

8. E. S. Kuh and I. N. Hajj, "Nonlinear circuit theory: resistive networks," Proc. IEEE, vol. 59, no. 3, Mar. 1971, pp. 340-355.

9. T. Fujisawa and E. S. Kuh, "Some results on existence and uniqueness of solution of nonlinear networks," IEEE Trans. Circuit Theory, to appear.

10. T. Fujisawa and E. S. Kuh, "Piecewise-linear theory of nonlinear networks," to appear.

11. L. O. Chua and Y. F. Lam, "Foundations of nonlinear network theory." Tech. Report #TR-EE 70-22, Purdue Univ., June 1970.

12. L. O. Chua, "The linear transformation converter and its application to the synthesis of nonlinear networks," IEEE Trans Circuit Theory, vol. CT-17, no. 4, Nov. 1970, pp. 584-594.

13. M. Rosenlicht, "Introduction to Analysis," Scott, Foresman and Co., Glenview, Ill., 1968.

14. L. H. Loomis and S. Sternberg, "Advanced Calculus," Addison-Wesley, Reading, Mass., 1968.

15. I. M. Singer and J. A. Thorpe, "Lecture Notes on Elementary Topology and Geometry," Scott, Foresman and Co., Glenview, Ill., 1967.

FOOTNOTES

1. A homeomorphism of X onto Y is, by definition, a continuous bijective map f: $X \to Y$ such that f^{-1} is also continuous.

2. A map f: $X \to Y$ is said to be a local homeomorphism if whenever $x \in X$ and $y \in Y$ are such that $f(x) = y$ then there exist open neighborhoods U of x and V of y such that f restricted to U is a homeomorphism of U onto V.

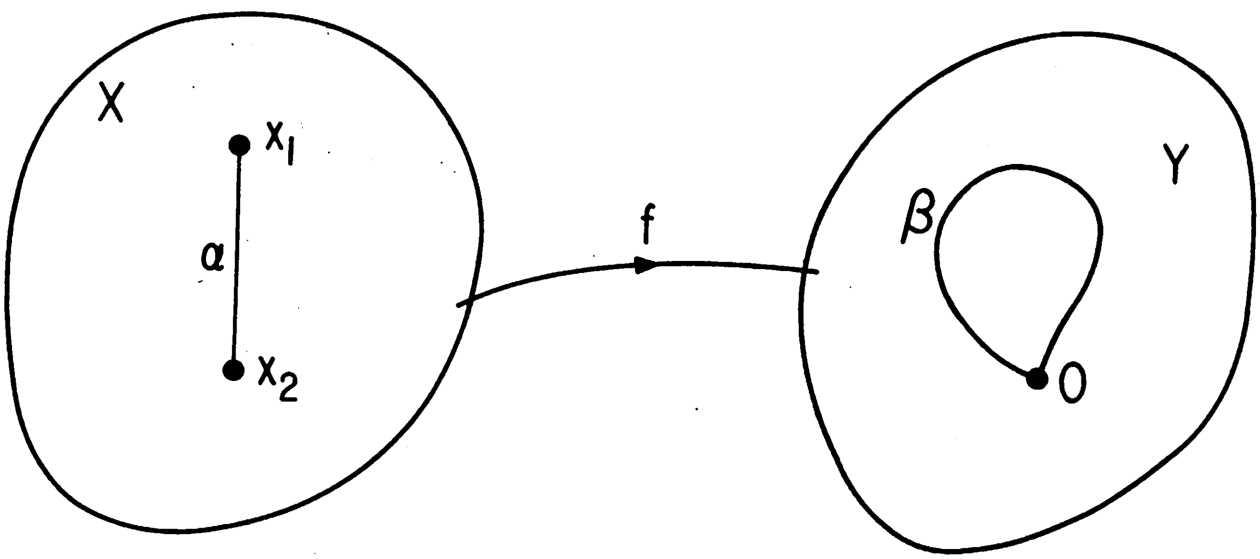
3. A continuous map is said to be proper if the inverse image of any compact set is compact.

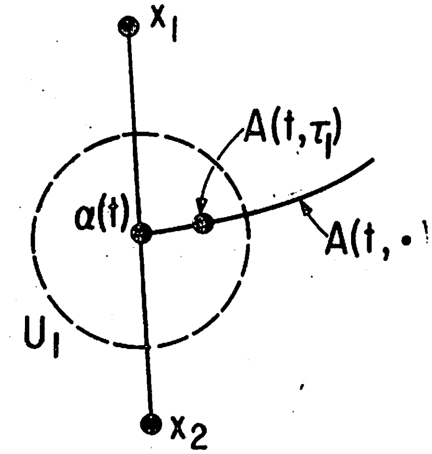
4. A C^k-map is, by definition, a map with continuous derivatives up to order k. A C^k-diffeomorphism is, by definition, a bijective C^k map such that the inverse is also C^k.

FIGURE CAPTIONS

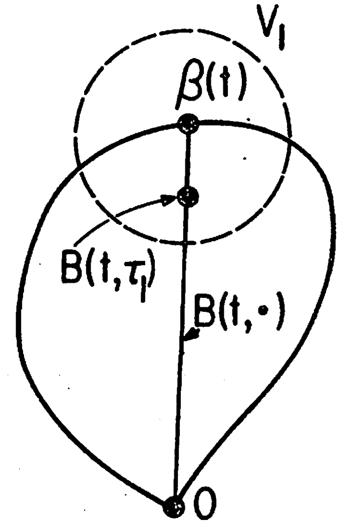
Fig. 1. For the proof by contradiction, it is assumed that $x_1 \neq x_2$ and that $f(x_1) = f(x_2) = 0$. The line segment α which joins x_1 to x_2 is mapped by f onto the closed curve β.

Fig. 2. As τ goes from 0 to 1, $B(t,\tau) = (1-\tau)\ \beta(t)$ travels in a straight line from $\beta(t)$ to 0. For the same fixed t, the corresponding curve $A(t,\tau)$ is constructed from $B(t,\tau)$ by successive local homeomorphisms.


X
x_1
α
x_2
f
Y
β
0



x1
A(t, τ1)
α(t)
A(t, •)
U1
x2



V1
β(t)
B(t, τ1)
B(t, •)
O