A Less Strange Version of Milnor's Proof of Brouwer's Fixed-Point Theorem

C. A. Rogers

The American Mathematical Monthly, Vol. 87, No. 7. (Aug. - Sep., 1980), pp. 525-527.

Stable URL:
http://links.jstor.org/sici?sici=0002-9890%28198008%2F09%2987%3A7%3C525%3AALSVOM%3E2.0.CO%3B2-N

The American Mathematical Monthly is currently published by Mathematical Association of America.

A LESS STRANGE VERSION OF MILNOR'S PROOF OF BROUWER'S FIXED-POINT THEOREM

C. A. ROGERS

In a recent note [1], John Milnor gives a proof of the "hairy dog theorem" and deduces the Brouwer fixed-point theorem as a consequence. Milnor describes his proof as strange but elementary. In this note we give a less strange and more direct proof of the following retraction theorem, which is well known to be equivalent, in a completely elementary and fairly simple way, to the Brouwer fixed-point theorem.

THEOREM 1. *It is not possible for a continuous function* $\mathbf{f}$ *to map the unit ball* $B^n = \{\mathbf{x} \mid \|\mathbf{x}\| \leqslant 1\}$ *of n-dimensional Euclidean space onto the unit sphere* $S^{n-1} = \{\mathbf{x} \mid \|\mathbf{x}\| = 1\}$ *and to satisfy*

$$\mathbf{f}(\mathbf{x}) = \mathbf{x}$$

for all $\mathbf{x}$ *on* S^{n-1}.

The theorem is an immediate consequence of two lemmas.

LEMMA 1. *If there were a continuous map of* B^n *onto* S^{n-1} *leaving each point of* S^{n-1} *fixed, then there would be a continuously differentiable map with these properties.*

LEMMA 2. *It is not possible for a continuously differentiable function to map* B^n *onto* S^{n-1} *and to leave each point of* S^{n-1} *fixed.*

The proof of the first lemma uses standard ideas; the proof of the second lemma uses the ideas of Milnor.

Proof of Lemma 1. Let $\mathbf{f}$ map B^n continuously onto S^{n-1} and suppose that $\mathbf{f}(\mathbf{x}) = \mathbf{x}$ for all $\mathbf{x}$ on S^{n-1}. Then $\mathbf{f}(\mathbf{x}) - \mathbf{x}$ is continuous on B^n, vanishes on S^{n-1}, and satisfies

$$\|\mathbf{f}(\mathbf{x}) - \mathbf{x}\| \leqslant 2 \tag{1}$$

on B^n. So we can choose θ with $\frac{3}{4} < \theta < 1$ so that

$$\|\mathbf{f}(\mathbf{x}) - \mathbf{x}\| < \tfrac{1}{4}, \quad \text{for } \theta \leqslant \|\mathbf{x}\| \leqslant 1. \tag{2}$$

Let $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ be the unit vectors along the coordinate axes. By the Weierstrass approximation theorem we can choose polynomials $P_i(x_1, x_2, \ldots, x_n)$, $1 \leqslant i \leqslant n$, so that

$$\left\| \sum_{i=1}^{n} P_i(x_1, x_2, \ldots, x_n)\mathbf{e}_i - (\mathbf{f}(\mathbf{x}) - \mathbf{x}) \right\| < \tfrac{1}{4}, \tag{3}$$

for all $\mathbf{x}$ with $\|\mathbf{x}\| \leqslant 1$. Write

$$\mathbf{P}(\mathbf{x}) = \sum_{i=1}^{n} P_i(x_1, x_2, \ldots, x_n)\mathbf{e}_i,$$

for convenience. Again, using the Weierstrass approximation theorem, we can choose a polynomial Q satisfying

$$\tfrac{3}{4} \leqslant Q(r^2) \leqslant 1, \qquad 0 \leqslant r \leqslant \theta; \qquad |Q(r^2)| \leqslant 1, \qquad \theta \leqslant r \leqslant 1; \qquad Q(1) = 0.$$

C. A. Rogers received his Ph.D. from the University of London in 1949, after studying with R. G. Cooke, L. S. Bosanquet, and H. Davenport. He has been at University College London on and off since 1938. He became a Fellow of the Royal Society in 1959 and has served as President of the London Mathematical Society (1970–72). He has worked on Geometry of numbers, Packing and covering, Hausdorff measures, Convexity, and Analytic sets.—*Editors*

Write

$$\mathbf{g}(\mathbf{x}) = \mathbf{x} + Q(\|\mathbf{x}\|^2)\mathbf{P}(\mathbf{x}).$$

For $0 \leqslant \|\mathbf{x}\| \leqslant \theta$, we have

$$\begin{aligned}\|\mathbf{g}(\mathbf{x})\| &= \|\mathbf{x} + Q(\|\mathbf{x}\|^2)\mathbf{P}(\mathbf{x})\| \\ &= \|\mathbf{f}(\mathbf{x}) + Q(\|\mathbf{x}\|^2)\{\mathbf{P}(\mathbf{x}) - \mathbf{f}(\mathbf{x}) + \mathbf{x}\} + \{Q(\|\mathbf{x}\|^2) - 1\}\{\mathbf{f}(\mathbf{x}) - \mathbf{x}\}\| \\ &\geqslant \|\mathbf{f}(\mathbf{x})\| - |Q(\|x\|^2)| \cdot \|\mathbf{P}(\mathbf{x}) - \mathbf{f}(\mathbf{x}) + \mathbf{x}\| - |1 - Q(\|\mathbf{x}\|^2)| \cdot \|\mathbf{f}(\mathbf{x}) - \mathbf{x}\| \\ &\geqslant 1 - 1 \cdot \tfrac{1}{4} - \tfrac{1}{4} \cdot 2 = \tfrac{1}{4}.\end{aligned}$$

Similarly, for $\theta \leqslant \|\mathbf{x}\| \leqslant 1$, we have

$$\begin{aligned}\|\mathbf{g}(\mathbf{x})\| &= \|\mathbf{x} + Q(\|\mathbf{x}\|^2)\{\mathbf{P}(\mathbf{x}) - \mathbf{f}(\mathbf{x}) + \mathbf{x}\} + Q(\|\mathbf{x}\|^2)\{\mathbf{f}(\mathbf{x}) - \mathbf{x}\}\| \\ &\geqslant \|\mathbf{x}\| - |Q(\|\mathbf{x}\|^2)| \big[\, \|\mathbf{P}(\mathbf{x}) - \mathbf{f}(\mathbf{x}) + \mathbf{x}\| + \|\mathbf{f}(\mathbf{x}) - \mathbf{x}\| \,\big] \\ &\geqslant \theta - 1 \cdot \big[\tfrac{1}{4} + \tfrac{1}{4}\big] \geqslant \tfrac{1}{4}.\end{aligned}$$

So

$$\|\mathbf{g}(\mathbf{x})\| \geqslant \tfrac{1}{4} \quad \text{for } \|\mathbf{x}\| \leqslant 1. \tag{4}$$

For $\|\mathbf{x}\| = 1$, we have

$$\mathbf{g}(\mathbf{x}) = \mathbf{x}.$$

Now each component of $\mathbf{g}$ is a polynomial in $x_1, x_2, \ldots, x_n$. So $\mathbf{g}$ is continuously differentiable and so is $\mathbf{h}$, defined by

$$\mathbf{h}(\mathbf{x}) = \mathbf{g}(\mathbf{x}) / \|\mathbf{g}(\mathbf{x})\|$$

for $\|\mathbf{x}\| \leqslant 1$. Clearly, $\mathbf{h}$ is a continuously differentiable map of B^n onto S^{n-1}, leaving each point of S^{n-1} fixed.

Proof of Lemma 2. Suppose that $\mathbf{f}$ is a continuously differentiable map of B^n onto S^{n-1}, leaving each point of S^{n-1} fixed. Write

$$\begin{aligned}\mathbf{g}(\mathbf{x}) &= \mathbf{f}(\mathbf{x}) - \mathbf{x}, \\ \mathbf{f}_t(\mathbf{x}) &= \mathbf{x} + t\mathbf{g}(\mathbf{x}) = (1-t)\mathbf{x} + t\mathbf{f}(\mathbf{x}),\end{aligned}$$

for $\|\mathbf{x}\| \leqslant 1$ and $0 \leqslant t \leqslant 1$.

As $\mathbf{f}$ is continuously differentiable, so is $\mathbf{g}$, and there is a constant C such that

$$\|\mathbf{g}(\mathbf{y}) - \mathbf{g}(\mathbf{x})\| \leqslant C\|\mathbf{y} - \mathbf{x}\|$$

for all $\mathbf{x}, \mathbf{y}$ in B^n. If $0 \leqslant t < 1/C$, and $\mathbf{f}_t(\mathbf{x}) = \mathbf{f}_t(\mathbf{y})$, then

$$\begin{aligned}\|\mathbf{x} - \mathbf{y}\| &= \|t\mathbf{g}(\mathbf{y}) - t\mathbf{g}(\mathbf{x})\| \\ &\leqslant tC\|\mathbf{y} - \mathbf{x}\|,\end{aligned}$$

and $\mathbf{x} = \mathbf{y}$ as $tC < 1$. Thus, the map $\mathbf{f}_t$ from B^n to B^n is injective when $0 \leqslant t < 1/C$.

As the partial differential coefficients of $\mathbf{g}$ with respect to $x_1, x_2, \ldots, x_n$ are uniformly bounded, the Jacobian matrix

$$\left(\frac{\partial \mathbf{f}_t}{\partial x_1}, \frac{\partial \mathbf{f}_t}{\partial x_2}, \ldots, \frac{\partial \mathbf{f}_t}{\partial x_n}\right) = I_n + t\left(\frac{\partial \mathbf{g}}{\partial x_1}, \frac{\partial \mathbf{g}}{\partial x_2}, \ldots, \frac{\partial \mathbf{g}}{\partial x_n}\right) \tag{5}$$

is dominated by its diagonal and so is nonsingular provided $0 \leqslant t \leqslant t_0$, with t_0 a sufficiently small positive number. Now, for $0 \leqslant t \leqslant t_0$, the inverse function theorem tells us that $\mathbf{f}_t$ maps the interior of B^n into an open set, G_t say, contained in B^n. Consider any point $\mathbf{e}$ in B^n that is not in G_t for some t with $0 \leqslant t \leqslant t_0$. Join $\mathbf{e}$ to any point $\mathbf{g}$ of G_t and choose a point $\mathbf{b}$ on the line segment

$\mathbf{e}, \mathbf{g}$ on the boundary of G_t. As the image of B^n under $\mathbf{f}_t$ is compact, $\mathbf{b}=\mathbf{f}_t(\mathbf{x})$ for some $\mathbf{x}$ in B^n. As $\mathbf{b}$ is not in G_t, $\mathbf{x}$ is not in the interior of B^n and so has $\|\mathbf{x}\|=1$. Hence $\mathbf{b}=\mathbf{x}$, and $\mathbf{e}$ as well as $\mathbf{b}$ lies on the boundary of B^n. As $\mathbf{f}_t$ maps S^{n-1} onto itself, we see that, when $0\leqslant t\leqslant t_0$, $\mathbf{f}_t$ maps B^n bijectively to itself.

Now consider the integral

$$I(t)=\int_{B^n}\frac{\partial \mathbf{f}_t}{\partial \mathbf{x}}\,d\mathbf{x}=\int\cdots\int_{B^n}\det\left(\frac{\partial \mathbf{f}_t}{\partial x_1},\frac{\partial \mathbf{f}_t}{\partial x_2},\ldots,\frac{\partial \mathbf{f}_t}{\partial x_n}\right)dx_1\,dx_2\cdots dx_n,$$

for $0\leqslant t\leqslant 1$. When $0\leqslant t\leqslant t_0$, we have a formula for the volume V_n of the unit ball B_n. Thus $I(t)$ has the constant value V_n for $0\leqslant t\leqslant t_0$. But it is clear from (5) that $I(t)$ is a polynomial in t. Hence $I(t)$ is constant and has the positive value V_n for all t. But, we have

$$\mathbf{f}_1\cdot\mathbf{f}_1=1$$

identically, so that

$$\frac{\partial \mathbf{f}_1}{\partial x_i}\cdot\mathbf{f}_1=0,\qquad 1\leqslant i\leqslant n,$$

and

$$\det\left(\frac{\partial \mathbf{f}_1}{\partial x_1},\frac{\partial \mathbf{f}_1}{\partial x_2},\ldots,\frac{\partial \mathbf{f}_1}{\partial x_n}\right)=0,$$

for all $\mathbf{x}$ in B^n. Thus $I(1)=0$, and we have the required contradiction.

Note added December 1979. Dr. Roger Smart has explained to me that Brouwer's fixed-point theorem for continuously differentiable maps can be obtained directly from Lemma 2, without use of Lemma 1. The general case of Brouwer's theorem then follows by a simpler application of the Weierstrass approximation theorem. Theorem 1 then follows easily from the general Brouwer theorem.

Reference

1. J. Milnor, Analytic proofs of the "Hairy Ball Theorem" and the Brouwer Fixed Point Theorem, this Monthly, 85 (1978) 521–524.

Department of Mathematics, University College London, Gower Street, London, WC1E 6BT, England.

ON THE DEVELOPMENT OF OPTIMIZATION THEORY

ANDRÁS PRÉKOPA

1. Introduction. Farkas's famous paper of 1901 [23] became a principal reference for linear inequalities after the publication of the paper of Kuhn and Tucker, "Nonlinear Programming,"

András Prékopa received his Ph.D. in 1956 at the University of Budapest under the leadership of A. Rényi. He was Assistant Professor and later Associate Professor of Probability Theory at the same university. Since 1968 he has held a professorship of mathematics at the Technical University of Budapest. He is also head of the Applied Mathematics Department at the Computer and Automation Institute of the Hungarian Academy of Sciences. He is a corresponding member of the Hungarian Academy of Sciences, the National Academy of Engineering of Mexico, member of the I.S.I., and Fellow of the Econometric Society. His main interests are in Probability, Statistics, and Operations Research, including application in engineering design (water resources, electrical engineering), economics, and natural sciences. This paper was prepared while the author was visiting the Mathematics Research Center of the University of Wisconsin–Madison. Readers who are not familiar with the subject whose history is discussed in this article should consult Pourciau's article in this Monthly, 87 (1980) 433–451.—*Editors*