ALTERNATIVE APPROACHES TO PROOFS OF CONTRACTION MAPPING FIXED POINT THEOREMS

James E. Joseph and Myung H. Kwack

While preparing a lecture on Banach's Contraction Mapping Theorem for a classroom presentation to a group of mathematics students in elementary real analysis, we discovered elegant alternative proofs of the theorem and some of its generalizations and variations. The approaches to these proofs are more in the nature of ones which an undergraduate mathematics major might attempt while studying complete metric spaces if asked to produce a proof of Banach's theorem before having seen the ingenious proof given by Banach.

In [1], Banach proved the celebrated Contraction Mapping Theorem.

(1) If (X, d) is a complete metric space and $g\colon X \to X$ is a function satisfying $d(g(x), g(y)) \leq \mu d(x, y)$ for all $x, y \in X$ and some fixed $0 < \mu < 1$, then g has a unique fixed point.

Banach showed that for any $v \in X$, the sequence of iterates $\{g^n(v)\}$ is a Cauchy sequence and is therefore convergent to some $p \in X$. By the continuity of g, the subsequence $\{g^{n+1}(v)\}$ converges to $g(p)$. Thus, $g(p) = p$. Most contraction mapping type theorems have been established by variations on this technique (see [4,5,6,7,8]).

In [2], Boyd and Wong gave an alternative approach to the proof of (1), using Cantor's intersection property. Their proof, while not the first to directly utilize the Cantor theorem, was more direct than those previously given (see [7]). We now present another alternative approach. Let $I(g) = \{d(x, g(x)) \ : \ x \in X\}$. We show first that $0 = \inf I(g)$ and then that $0 \in I(g)$.

Our Proof of (1). Let $c = \inf I(g)$. If $c > 0$ we have $c/\mu > c$ and an $x \in X$ satisfying

$$d(g(x), g(g(x))) \leq \mu d(x, g(x)) < c,$$

a contradiction. Thus, $c = 0$. Let $\{x_n\}$ be a sequence in X with $d(x_n, g(x_n)) \to 0$. Then $\{x_n\}$ is a Cauchy sequence since

$$d(x_n, x_m) \leq d(x_n, g(x_n)) + d(g(x_n), g(x_m)) + d(x_m, g(x_m))$$

gives

$$(1-\mu)d(x_n,x_m) \leq d(x_n g(x_n)) + d(x_m, g(x_m)).$$

Thus, there is a $p \in X$ with $x_n \to p$, and $g(x_n) \to p$ follows from the fact that $d(x_n, g(x_n)) \to 0$. It follows that $d(g(x_n), g(p)) \leq \mu d(x_n, p)$, so $g(x_n) \to g(p)$ and $g(p) = p$. The uniqueness of p follows as usual from the contractive nature of g.

Remark 1. One of the useful features of Banach's proof of the theorem is that

$$d(g^n(v), p) \leq \frac{\mu^n}{1-\mu} d(v, g(v)).$$

We see readily that this feature is captured by our method since for any $v \in X$, we have by induction that

$$d(g^n(v), p) = d(g^n(v), g^n(p)) \leq \mu^n d(v,p) \leq \frac{\mu^n}{1-\mu} d(v, g(v)).$$

As another application of this method, we present an alternative proof of the following theorem due to Fisher [5].

(2) If g is a mapping of a complete metric space (X, d) into itself satisfying the condition

$$d(g(x), g(y)) \leq \mu[d(x, g(y)) + d(y, g(x))]$$

for all $x, y \in X$ and some $0 \leq \mu < 1/2$, then g has a unique fixed point.

Our Proof of (2). Let $c = \inf I(g)$ and suppose $0 < \mu$. If $c > 0$ then $((1-\mu)/\mu)c > c$ since $\mu < 1/2$; so there is an $x \in X$ satisfying

$$d(g(x), g^2(x)) \leq \frac{\mu}{1-\mu} d(x, g(x)) < c.$$

Hence, $c = 0$ and $d(x_n, g(x_n)) \to 0$ for some sequence $\{x_n\}$ in X. Such a sequence satisfies

$$d(x_n, x_m) \leq d(x_n, g(x_n)) + d(g(x_n), g(x_m)) + d(g(x_m), x_m)$$

$$\leq \frac{1+\mu}{1-2\mu}[d(x_n, g(x_n)) + d(x_m, g(x_m))].$$

Hence, $\{x_n\}$ is a Cauchy sequence. Choose $p \in X$ with $x_n \to p$. Then $g(x_n) \to p$ since $d(x_n, g(x_n)) \to 0$. Since

$$d(g(x_n), g(p)) \leq \mu[d(p, g(x_n)) + d(x_n, g(p))]$$

we obtain $d(p, g(p)) \leq \mu d(p, g(p))$ and $g(p) = p$.

Remark 2. An analysis of the proof of (2) shows that any sequence $\{x_n\}$ in X with $d(x_n, g(x_n)) \to 0$ must converge to the fixed point. We observe that if $v \in X$ and $x_n = g^{n-1}(v)$ it follows by induction that

$$d(x_n, g(x_n)) \leq \left(\frac{\mu}{1-\mu}\right)^{n-1} d(v, g(v)),$$

so the sequence of iterates of any point converges to the fixed point.

Remark 3. If p is the fixed point assured in (2) and $\{x_n\}$ is any sequence in X, then

$$\begin{aligned} d(x_n, p)) &\leq d(x_n, g(x_n)) + d(g(x_n), g(p)) \\ &\leq \frac{1+\mu}{1-2\mu} d(x_n, g(x_n)). \end{aligned}$$

Therefore, from Remark 2, if $v \in X$ and $x_n = g^{n-1}(v)$ it follows that

$$d(x_n, p) \leq \frac{1+\mu}{1-2\mu} \left(\frac{\mu}{1-\mu}\right)^{n-1} d(v, g(v)),$$

which gives an estimate for the rate of convergence of the iterates of the point v to the fixed point.

Remark 4. We observe that a consequence of (2) is that if (X, d) is a complete metric space and $g: X \to X$ is a function satisfying $d(g(x), g(y)) \leq \mu d(x, y)$ for all

$x, y \in X$ and some fixed $0 < \mu < 1/3$, then g has a unique fixed point. This follows from

$$d(g(x), g(y)) \le \mu[d(x, g(y)) + d(g(x), g(y)) + d(y, g(x))]$$

$$\le \frac{\mu}{1-\mu}[d(x, g(y)) + d(y, g(x))],$$

and $(\mu/(1-\mu)) < 1/2$.

Our next illustration of this method is a proof of a common generalization of (1) and (2) which is due to Hardy and Rogers [6].

(3) Let (X, d) be a complete metric space and let g be a self mapping of X satisfying the following condition for all $x, y \in X$ and fixed nonnegative a, b, c, e, f with $a + b + c + e + f < 1$:

$$d(g(x), g(y)) \le ad(x, g(x)) + bd(y, g(y)) + cd(x, g(y)) + ed(y, g(x)) + fd(x, y).$$

Then g has a unique fixed point p and, for each $v \in X$, $g^n(v) \to p$.

Our Proof of (3). Let $q = \inf I(g)$. By the lemma in [6] there is a $\mu < 1$ satisfying

$$d(g(x), g^2(x)) \le \mu d(x, g(x))$$

for all $x \in X$. Thus, $q = 0$. There is a sequence $\{x_n\}$ in X with $d(x_n, g(x_n)) \to 0$. We see that $\{x_n\}$ is a Cauchy sequence since

$$[1 - (c + e + f)]d(x_n, x_m) \le (1 + a + e)d(x_n, g(x_n)) + (1 + b + c)d(x_m, g(x_m)).$$

Let $p \in X$ satisfy $x_n \to p$. Then $g(x_n) \to p$ and $g(p) = p$ since

$$\begin{aligned} d(g(x_n), g(p)) \le\ & ad(x_n, g(x_n)) + bd(p, g(p)) + cd(x_n, p) \\ & + cd(p, g(p)) + ed(x_n, p) + ed(p, g(p)) + fd(x_n, p), \end{aligned}$$

which gives

$$d(p, g(p)) \le (b + c)d(p, g(p)).$$

The uniqueness of p is readily established. It is seen from the argument above that if $\{x_n\}$ is any sequence in X with $d(x_n, g(x_n)) \to 0$, we have $x_n \to p$; since

$$d(g^n(v), g^{n+1}(v)) \le \mu^n d(v, g(v))$$

for any $v \in X$, it follows that $g^n(v) \to p$. Also, we observe that if p is the fixed point and $d(x_n, g(x_n)) \to 0$, we have

$$[1 - (c + e + f)]d(x_n, p) \le (1 + a + e)d(x_n, g(x_n)),$$

so

$$d(g(x_n), p) \le \frac{1 + a + e}{1 - (c + e + f)} d(g^n(v), g^{n+1}(v))$$

$$\le \frac{1 + a + e}{1 - (c + e + f)} \mu^n d(v, g(v)).$$

This furnishes an estimate for the rate of convergence of the iterates of a point in the space.

Remark 5. An alternative proof of a theorem of Edelstein [4] concerning contraction mappings on ϵ–chainable spaces may be obtained by a slight modification of the method employed above.

For our final demonstration of alternative proofs we produce two proofs of the following theorem on common fixed points, a result due to Kannan [7]. The first uses the ideas employed above in the proofs of (1), (2), (3), and the second uses the Cantor Intersection Property.

(4) If (X, d) is a complete metric space and $g, h\colon X \to X$ are functions satisfying

$$d(g(x), h(y)) \le \mu[d(x, g(x)) + d(y, h(y))]$$

for all $x, y \in X$ and some $0 < \mu < 1/2$, then g and h have a unique common fixed point.

Our First Proof of (4). Let $c = \inf(I(g) \cup I(h))$. Then $c = \inf I(g)$ or $c = \inf I(h)$. We exhaust the possibilities.

Case 1. $c = \inf I(g)$. If $c \neq 0$ then $((1-\mu)/\mu)c > c$ and there is an $x \in X$ such that $d(x, g(x)) < ((1-\mu)/\mu)c$. For such an x we have

$$d(g(x), h(g(x)) \leq \frac{\mu}{1-\mu} d(x, g(x)) < c,$$

a contradiction. Let $\{x_n\}$ be a sequence in X such that $d(x_n, g(x_n)) \to 0$. Then

$$\begin{aligned} d(x_n, h(x_n)) &\leq d(x_n, g(x_n)) + d(g(x_n), h(x_n)) \\ &\leq (1+\mu)d(x_n, g(x_n)) + \mu d(x_n, h(x_n)) \\ &\leq \frac{1+\mu}{1-\mu} d(x_n, g(x_n)). \end{aligned}$$

Hence, $d(x_n, h(x_n)) \to 0$. We see that $\{x_n\}$ is a Cauchy sequence from the fact that

$$d(x_n, x_m) \leq (1+\mu)[d(x_n, g(x_n)) + d(x_m, h(x_m))].$$

Let $x_n \to p$. Then

$$(1-\mu)d(g(p), p) \leq (1+\mu)d(h(x_n), x_n) + d(x_n, p),$$

so $g(p) = p$. To see that $h(p) = p$ note that

$$(1-\mu)d(h(p), p) \leq (1+\mu)d(g(x_n), x_n) + d(x_n, p).$$

Case 2. $c = \inf I(h)$. Entirely similar to Case 1.

To effect our second proof of (4) we utilize the following well–known notions and properties of metric spaces (see [3]).

(a) In a metric space, the diameter of a nonempty subset A ($\sup\{d(x,y) : x, y \in A\}$) is the same as the diameter of the closure $\overline{A}$. We denote the diameter of A by $\delta(A)$.

(b) A metric space is complete if and only if $\cap F_n$ is a singleton set for every decreasing sequence $\{F_n\}$ of nonempty closed subsets of the space with $\delta(F_n) \to 0$.

For convenience we adopt the following additional notations. If (X, d) is a metric space and $g : X \to X$ is a function we will denote the graph of g in

$X \times X$ by $G(g)$, and will denote by e the distance function induced on $X \times X$ by d $(e((x,y),(u,v)) = d(x,u) + d(y,v))$. With this metric, $X \times X$ is complete when X is complete.

Our Second Proof of (4). Let $c = \inf(I(g) \cup I(h))$. Then $c \geq 0$, and if $c > 0$ then $((1-\mu)/\mu)c > c$ since $\mu < 1/2$. Hence, there is an $x \in X$ satisfying either

(i) $d(x, g(x)) < ((1-\mu)/\mu)c > c$

or

(ii) $d(x, h(x)) < ((1-\mu)/\mu)c$.

If x satisfies (i) then

$$d(g(x), h(g(x))) \leq \frac{\mu}{1-\mu} d(x, g(x)) < c.$$

If x satisfies (ii) then

$$d(h(x), g(h(x))) \leq \frac{\mu}{1-\mu} d(x, h(x)) < c.$$

Consequently, $c \neq \inf(I(g) \cup I(h))$. Hence, $c = 0$. For each positive integer n let

$$F_n = d^{-1}([0, 1/n]) \cap (G(g) \cup G(h)).$$

Then $\overline{F}_n \neq \emptyset$ and $\overline{F}_{n+1} \subset \overline{F}_n$. We will complete the proof by distinguishing two cases.

Case 1. $(x, g(x)), (y, h(y)) \in F_n$. Then

$$\begin{aligned} e((x,g(x)),(y,h(y))) &= d(x,y) + d(g(x),h(y)) \\ &\leq d(x,g(x)) + d(g(x),h(y)) + d(h(y),y) + d(g(x),h(y)) \\ &\leq d(x,g(x)) + 2\mu[d(x,g(x)) + d(h(y),y)] + d(h(y),y) \\ &\leq (1+2\mu)[d(x,g(x)) + d(h(y),y)] \leq (1+2\mu)\frac{2}{n}. \end{aligned}$$

Case 2. $(x, g(x)), (y, g(y)) \in F_n$. Then $(g(x), h(g(x)) \in F_n$, so

$$e((x, g(x)), (y, g(y))) \le e((x, g(x)), (g(x), h(g(x))) + e((g(x), h(g(x)))), (y, g(y)))$$

$$\le (1 + 2\mu)\frac{4}{n}$$

from Case 1. Hence, $\delta(\overline{F}_n) = \delta(F_n) \to 0$, and $\cap \overline{F}_n = \{(x, y)\}$; $d(x, y) \in [0, 1/n]$ for each n so $x = y$. If $(x_k, g(x_k)) \to (x, x)$, then

$$d(g(x_k), h(x)) \le \mu[d(x_k, g(x_k)) + d(x, h(x))],$$

so

$$d(x, h(x)) \le \mu d(x, h(x))$$

and $h(x) = x$. Also

$$d(x, g(x)) = d(h(x), g(x)) \le \mu d(x, g(x))$$

and $g(x) = x$. A similar argument gives $g(x) = h(x) = x$ if $(x_k, h(x_k)) \to (x, x)$.

We conclude with two exercises for the reader.

Exercise 1. Give a proof of Banach's Contraction Mapping Theorem using the Cantor Intersection Theorem.

Exercise 2. Give a proof of (2) using the Cantor Intersection Theorem.

References

1. S. Banach, "Sur les Operations dan les Ensembles Abstraite et leur Application aux Equations Integrales," *Fund. Math.*, 3 (1922), 133–181.
2. D. W. Boyd and J. S. W. Wong, "Another Proof of the Contraction Mapping Theorem," *Canad. Math. Bull.*, 11 (1968), 605–606.
3. J. Dugundji, *Topology*, Allyn & Bacon, Inc., Boston, Mass., 1966.

4. M. Edelstein, "An Extension of Banach's Contraction Principle," *Proc. Amer. Math. Soc.*, 12 (1961), 7–10.

5. B. Fisher, "A Fixed Point Theorem," *Math. Mag.*, 48 (1975), 223–225.

6. G. E. Hardy and T. D. Rogers, "A Generalization of a Fixed Point Theorem of Reich," *Canad. Math. Bull.*, 16 (1973), 201–206.

7. R. Kannan, "Some Results on Fixed Points," *Bull. Cal. Math.*, 60 (1968), 71–76.

8. I. I. Kolodner, "On the Proof of the Contraction Mapping Theorem," *Amer. Math. Monthly*, 74 (1967), 1212–1213.

James E. Joseph
Department of Mathematics
Howard University
Washington, D.C. 20059
email: jjoseph@scs.howard.edu

Myung H. Kwack
Department of Mathematics
Howard University
Washington, D.C. 20059