JOURNAL OF RESEARCH of the National Bureau of Standards — B. Mathematics and Mathematical Physics
Vol. 71B, Nos. 2 and 3, April–September 1967

A Converse to Banach's Contraction Theorem

Philip R. Meyers

Institute for Basic Standards, National Bureau of Standards, Washington, D. C. 20234

(March 3, 1967)

The class of all continuous self-mappings of a metrizable space which can become contractions (in the sense of Banach) under metrics compatible with the topology on the space is characterized. The characterization amounts to a converse to the Contraction Mapping Principle.

Key Words: Contractions, functional analysis, metric spaces, topology.

1. Introduction

Throughout this paper, X denotes a metrizable topological space and $f: X \to X$ denotes a continuous map. If ρ is a metric on X, then f is called a *ρ-contraction* in case there exists $\lambda \epsilon (0, 1)$ which is a *contraction constant* for f on (X, ρ), i.e., for all $x, y \epsilon X$,

$$\rho(f(x), f(y)) \leqslant \lambda \rho(x, y). \tag{1.1}$$

Consider the hypothesis (H) that X admits a metric ρ (yielding the correct topology) such that f is a ρ-contraction. In a previous paper [1][1] we gave conditions on X such that f, if for some metric it satisfied (1.1) *locally* at all points of X, would in fact obey (H). Note that such results have a hypothesis which is *metric* in nature. The present paper is concerned with topological conditions for (H), conditions whose statements do not refer to any particular metric or metrics on X.

In view of the Banach Contraction Theorem [2], two conditions which naturally suggest themselves are that, for some $\xi \epsilon X$,

(i) $$f(\xi) = \xi,$$

(ii) $$f^n(x) \to \xi \text{ as } n \to \infty, \text{ for all } x \epsilon X.$$

These are among the usually stated conclusions of the Contraction Theorem, and so are surely necessary if (H) is to hold for some *complete*[2] metric ρ. We shall add to them a *third* condition, an easy (though usually unstated) conclusion of the Contraction Theorem, and then show that these three conditions together are *sufficient* (as well as necessary) to guarantee (H). The third condition is that there exist an open neighborhood U of ξ such that

(iii) $$f^n(U) \to \{\xi\}.$$

This means that for any open neighborhood V of ξ, there is an $n(V) > 0$ such that $f^n(U) \subset V$ for all $n \geqslant n(V)$.

If (H) holds, take $U = \{x : \rho(x, \xi) < 1\}$. Since $\lambda < 1$ in (1.1), we can for any neighborhood V of ξ choose $n(V)$ so large that for all $n \geqslant n(V)$,

$$\{x : \rho(x, \xi) < \lambda^n\} \subset V,$$

which implies by (1.1) that $f^n(U) \subset V$. Thus (iii) is indeed a consequence of (H). It remains to show how (H) can be deduced from (i), (ii), and (iii).

2. Results and Corollaries

At this point we state all of the results of the paper. As above, f is a continuous self-mapping of the metrizable topological space X.

THEOREM 1: *If* f *satisfies conditions* (i), (ii), *and* (iii) *above, then for each* $\lambda \epsilon (0, 1)$ *there exists a metric* ρ_λ *on* X, *complete if* X *admits a complete metric, such that* f *is a* ρ_λ*-contraction with contraction constant* λ.

COROLLARY 1.1: *If* ξ *has a compact neighborhood, then* (i) *and* (ii) *are sufficient conditions in Theorem 1.*

COROLLARY 1.2:[3] *If* X *is compact, and if* $\{\xi\}$ *is the only nonempty set* D *such that*

$$\mathrm{f(D) = D}$$

then the conclusions of Theorem 1 are valid.

THEOREM 2: *If some iterate* $\mathrm{f^m}$ *of* f *satisfies conditions* (i), (ii), *and* (iii) *then there is a metric* ρ, *complete if* X *admits a complete metric, such that* f *and all its iterates are simultaneously* ρ*-contractions.*

Theorem 2 suggests some further questions. Under what conditions do two mappings which satisfy conditions (i), (ii), and (iii) become contractions under

[1] Figures in brackets indicate the literature references at the end of this paper.

[2] Completeness is rather inessential if we are willing to operate with an equivalence class of Cauchy sequences rather than an actual point.

[3] This corollary is a restatement of a theorem due to Janos [3]. The proof of Theorem 1 uses ideas first used by Janos in his proof of this theorem; the proof has just been published.

the *same* metric? More generally, when can X be metrized so that a given *family* of self-mappings *simultaneously* become contractions? These and similar questions will be addressed in another paper under preparation.

The proof of Theorem 1 involves a somewhat lengthy construction of the desired metric ρ_λ. While this construction will be deferred to the next section, we give here the derivations of the remaining results from Theorem 1.

To derive Corollary 1.1, it suffices to show that (iii) follows from (i) and (ii) if ξ has a compact neighborhood C. Let $U = \mathrm{Int}(C)$, an open neighborhood of ξ. Consider any open neighborhood V of ξ. For each $x \epsilon C$, there exists by (ii) a smallest $n(x)$ such that $f^n(x) \epsilon V$ for all $n \geqslant n(x)$. We need only show that

$$n(V) = \sup \{n(x) : x \epsilon C\}$$

is finite. If *not*, then C contains a sequence $\{x_i\}$ such that $n(x_i) > i$, and since C is compact we may assume $x_i \rightarrow y$ for some $y \epsilon C$. The desired contradiction follows by observing that $n(y) < \infty$, and that by continuity $n(x) \leqslant n(y) + 1$ for all x in some neighborhood of y.

To derive Corollary 1.2, we show that (ii) follows from the stated hypotheses when X is compact and then use Corollary 1.1. Define successively

$$A_0 = X,\ A_1 = f(X),\ A_2 = f^2(X) = f(A_1),\ \text{etc.}$$

Since $f(X) \subset X$ and if $X \neq \{\xi\}$ the inclusion is proper, the sets A_m are a descending sequence of compact nonempty sets. Then $\bigcap^{\infty} A_m = D \neq \phi$. We show that

$$f(D) = D$$

and conclude that $D = \{\xi\}$. It is then a standard argument that diam $(A_m) \rightarrow 0$ and hence $f^m(x) \rightarrow \xi$ for $x \epsilon X$. It is clear that

$$\bigcap_0^{\infty} A_m = D \supset f(D) = \bigcap_1^{\infty} A_m.$$

In the other direction choose $x \epsilon D$ and consider the sequence of nonempty compact sets

$$S_m = f^{-1}(x) \cap A_m.$$

Since $\cap S_m = f^{-1}(x) \cap D \neq \phi$ we have $x \epsilon f(D)$ and $f(D) \supset D$ which completes the proof.

To derive Theorem 2, we show that f obeys conditions (i), (ii), and (iii) for the same point ξ as does its iterate f^m. Then by Theorem 1 there is a metric ρ and a $\lambda \epsilon (0, 1)$ such that f is a ρ-contraction with contraction constant λ. It follows that any iterate f^n is a ρ-contraction with contraction constant $\lambda < 1$.

Since $f^m(\xi) = \xi$, we have

$$f^{nm}(f(\xi)) = f(f^{nm}(\xi)) = f(\xi)$$

for all $n \geqslant 0$. But by (ii) applied to f^m, with $x = f(\xi)$,

$$(f^m)^n(f(\xi)) = f^{nm}(f(\xi)) \rightarrow \xi \qquad \text{as } n \rightarrow \infty$$

and so $f(\xi) = \xi$, condition (i) for f. For $0 \leqslant k < m$ we have, for any $x \epsilon X$,

$$f^{nm+k}(x) = (f^m)^n(f^k(x)) \rightarrow \xi \qquad \text{as } n \rightarrow \infty,$$

which implies that (ii) holds for f. And if U_m is an open neighborhood of ξ such that $f^{nm}(U_m) \rightarrow \{\xi\}$, then

$$U = \bigcap_{k=0}^{m-1} f^{-k}(U_m)$$

is an open neighborhood of ξ such that, for $0 \leqslant k < m$,

$$f^{mn+k}(U) = (f^m)^n(f^k(U)) \subset f^{nm}(U_m) \rightarrow \{\xi\}$$

as $n \rightarrow \infty$, implying that (iii) holds for f. This completes the derivation of Theorem 2.

3. Proof of Theorem 1

Throughout this section it is assumed that $\lambda \epsilon (0, 1)$, that $f: X \rightarrow X$ is a continuous map obeying the conditions (i), (ii), and (iii), and that ρ_0 is a metric on X, giving the correct topology for X and complete if X admits a complete metric. Our aim is to construct a metric ρ_λ, topologically equivalent to ρ_0 and complete if ρ_0 is, such that for all $x, y \epsilon X$,

$$\rho_\lambda(f(x), f(y)) \leqslant \lambda \rho_\lambda(x, y). \tag{3.1}$$

For the construction, it will be convenient if the neighborhood U of (iii) satisfies

$$f(U) \subset U. \tag{3.2}$$

We show first, therefore, that there exists an open neighborhood W of ξ such that $f(W) \subset W$ and $W \subset U$, the latter implying $f^n(W) \rightarrow \{\xi\}$. Then W can replace U in the construction to follow.

Since $f^n(U) \rightarrow \{\xi\}$, there is some integer k such that $f^k(U) \subset U$. Let

$$W = \bigcap_{j=0}^{k-1} f^{-j}(U) \subset U.$$

Then for $x \epsilon W$ we have, for $1 \leqslant j \leqslant k-1$, $x \epsilon f^{-j}(U)$ and thus $f(x) \epsilon f^{-(j-1)}(U)$: moreover $x \epsilon U$, so that $f^k(x) \epsilon f^k(U) \subset U$ and thus $f(x) \epsilon f^{-(k-1)}(U)$. Hence $x \epsilon W$ implies $f(x) \epsilon W$, which was to be shown.

We now proceed with the main line of the proof. The construction has three steps. The first step yields a metric ρ_M, topologically equivalent to ρ_0 and complete if ρ_0 is, with respect to which f is *nonexpanding* in the sense of satisfying the weak version

$$\rho_M(f(x), f(y)) \leqslant \rho_M(x, y) \tag{3.3}$$

of (3.1). The second step yields a function d_λ which has all the desired properties except perhaps for satisfying the triangle inequality. This is corrected in the third step, in which $\rho_\lambda(x, y)$ is introduced as what might be called the "d_λ-geodesic" distance between x and y.

For the *first* step, we set

$$\rho_M(x, y) = \max\{\rho_0(f^n(x), f^n(y)) : n \geq 0\}. \quad (3.4)$$

That the maximum is finite and actually attained follows from (ii), and (3.3) is obvious. The positive definiteness and symmetry of ρ_M, as well as $\rho_M(x, x) = 0$, follow at once from the corresponding properties of ρ_0. The triangle inequality for ρ_M follows from the observation that, for all $n \geq 0$,

$$\rho_0(f^n(x), f^n(y)) \leq \rho_0(f^n(x), f^n(z)) + \rho_0(f^n(z), f^n(y))$$
$$\leq \rho_M(x, z) + \rho_M(z, y).$$

Thus ρ_M is indeed a metric, which must still be shown to be topologically equivalent to ρ_0, and complete if ρ_0 is.

From the inequality

$$\rho_0 \leq \rho_M \quad (3.4)$$

it follows that any ρ_M-convergent sequence is also ρ_0-convergent (with the same limit point). To prove the implication in the opposite direction, note that (iii) implies the existence for each $\delta > 0$ of an N such that

$$(\rho_0\text{-diam})[f^n(U)] < \delta \qquad \text{for } n > N. \quad (3.5)$$

For each $x \epsilon X$, it follows from (ii) that

$$\nu(x) = \min\{n \geq 0 : f^n(x) \epsilon U\}$$

is finite. Since f is continuous, there is an $\eta > 0$ so small that $\rho_0(x, y) < \eta$ implies

$$f^{\nu(x)}(y) \epsilon U,\ \rho_0(f^j(x), f^j(y)) < \delta \qquad \text{for } 0 \leq j \leq N + \nu(x). \quad (3.6)$$

By (3.2), $f^{n+N+\nu(x)}(x) \epsilon f^{n+N}(U)$ and $f^{n+N+\nu(x)}(y) \epsilon f^{n+N}(U)$ for all $n > 0$, so that (3.6) implies

$$\rho_0(f^j(x), f^j(y)) < \delta \text{ for } j > N + \nu(x).$$

Thus $\rho_0(x, y) < \eta$ implies $\rho_M(x, y) < \delta$. This shows that a sequence which is ρ_0-convergent to x is also ρ_M-convergent to x, completing the proof of topological equivalence.

Now suppose ρ_0 is complete. By (3.4), any ρ_M-Cauchy sequence is also a ρ_0-Cauchy sequence, hence is ρ_0-convergent, and so (by the topological equivalence of the two metrics) is ρ_M-convergent. Thus the completeness of ρ_M follows from that of ρ_0.

For the *second* step in the construction, we begin by defining K_n to be the closure of $f^n(U)$ for $n \geq 0$, and $K_{(-n)} = f^{-n}(K_0)$, so that (iii) implies

$$K_n \to \{\xi\} \qquad \text{as } n \to \infty. \quad (3.7)$$

For $x \epsilon K_0 - \{\xi\}$, set

$$n(x) = \max\{n : x \epsilon K_n\} \geq 0;$$

finiteness is assured by (3.7). Let $n(\xi) = \infty$, and for $x \epsilon X - K_0$ set

$$n(x) = -\min\{m : f^m(x) \epsilon K_0\} = \max\{n : x \epsilon K_n\} < 0,$$

which must exist by (ii). Then d_λ is defined in terms of

$$c(x, y) = \min\{n(x), n(y)\}$$

by the formula

$$d_\lambda(x, y) = \lambda^{c(x, y)} \rho_M(x, y),$$

which has the correct limiting form $d_\lambda(\xi, \xi) = 0$. Then d_λ satisfies the metric requirements except perhaps for the triangle inequality, and from (3.3) and the fact that $n(f(x)) \geq n(x) + 1$, we see that d_λ has the property

$$d_\lambda(f(x), f(y)) \leq \lambda d_\lambda(x, y) \quad (3.8)$$

desired for ρ_λ. We turn now to the *third* and last step of the construction.

Denote by Σ_{xy} the set of chains $\sigma_{xy} = [x = x_0, x_1, \ldots, x_m = y]$ from x to y, with associated *lengths*

$$L_\lambda(\sigma_{xy}) = \sum_1^m d_\lambda(x_i, x_{i-1}),$$

and put

$$\rho_\lambda(x, y) = \inf\{L_\lambda(\sigma_{xy}) : \sigma_{xy} \epsilon \Sigma_{xy}\}. \quad (3.9)$$

We shall show that ρ_λ is the desired metric.

That f is a ρ_λ-contraction follows by applying (3.8) to the links $[x_{i-1}, x_i]$ of any chain σ_{xy}. Clearly ρ_λ is symmetric and $\rho_\lambda(x, x) = 0$; the triangle law holds since following a σ_{xy} with a σ_{yz} yields a σ_{xz}.

It remains to show that ρ_λ is positive definite. Consider any $x \neq \xi$ and any $y \neq x$: assume $n(x) \leq n(y)$ without loss of generality. If $y \neq \xi$, any chain σ_{xy} either lies in $X - K_{n(y)+1}$, or has a last link which leaves $K_{n(y)+1}$ (and possibly is followed by other links), so that

$$\rho_\lambda(x, y) \geq \lambda^{n(y)} \min\{\rho_M(x, y), \rho_M(y, K_{n(y)+1})\} > 0. \quad (3.10)$$

The remaining case, $y = \xi$, is covered by

$$\rho_\lambda(x, \xi) \geq \lambda^{n(x)} \rho_M(x, K_{n(x)+1}) > 0; \quad (3.11)$$

thus ρ_λ is positive definite and indeed a metric, which must still be proved equivalent to ρ_M.

Let $B_\nu = X - f^{-\nu}(U)$ for $\nu \geq 0$, so that the definition of $\nu(x)$ implies $\rho_M(x, B_{\nu(x)}) > 0$ and $n(x) \geq -\nu(x)$. For any $x \neq \xi$, if y obeys

$$\rho_M(x, y) < \delta(x) = \min\{\rho_M(x, K_{n(x)+1}), \rho_M(x, B_{\nu(x)})\}, \quad (3.12)$$

then $n(x) \geq -\nu(x)$, so that (3.9) and (3.10) . . . the latter with x and y interchanged . . . imply

$$\lambda^{n(x)}\rho_M(x, y) \leq \rho_\lambda(x, y) \leq d_\lambda(x, y) \leq \lambda^{-\nu(x)}\rho_M(x, y). \tag{3.13}$$

Choose $k(x) > \max\{0, n(x)\}$ such that $z \epsilon K_{k(x)}$ implies $\rho_M(\xi, z) < \rho_\lambda(x, \xi)/2$. Then $\rho_\lambda(x, K_{k(x)}) \geq \rho_\lambda(x, \xi)/2$, so that if y obeys

$$\rho_\lambda(x, y) < \rho_\lambda(x, \xi)/2 \tag{3.14}$$

then only chains disjoint from $K_{k(x)}$ need enter (3.9), implying

$$\rho_\lambda(x, y) \geq \lambda^{k(x)}\rho_M(x, y). \tag{3.15}$$

In particular, if

$$\rho_\lambda(x, y) < \min\{\rho_\lambda(x, \xi)/2, \lambda^{k(x)}\delta(x)\}$$

then with (3.14) and (3.15) this implies (3.12) and hence (3.13) applies. Thus $\rho_\lambda(x_n, x) \to 0$ iff $\rho_M(x_n, x) \to 0$.

As for $x = \xi$, note first that if $\rho_M(\xi, y) < \rho_M(\xi, B_0)$, then

$$\rho_\lambda(\xi, y) \leq d_\lambda(\xi, y) \leq \rho_M(\xi, y). \tag{3.16}$$

Second, for any $\eta > 0$, (iii) guarantees an $N(\eta) > 0$ such that $\rho_M(\xi, z) < \eta/2$ for all $z \epsilon K_{N(\eta)}$. Then

$$\rho_M(\xi, y) > \eta \tag{3.17}$$

implies that $\rho_M(y, K_{N(\eta)}) \geq \eta/2$ and thus that

$$\rho_\lambda(\xi, y) \geq \rho_\lambda(K_{N(\eta)}, y) \geq \lambda^{N(\eta)}\eta/2. \tag{3.18}$$

Hence $\rho_\lambda(\xi, x_n) \to 0$ iff $\rho_M(\xi, x_n) \to 0$.

To show that ρ_M-completeness is preserved, assume that $\{x_n\}$ is a ρ_λ-Cauchy sequence and that (X, ρ_M) is complete. If $\{x_n\}$ does not converge to ξ then since ρ_λ and ρ_M are equivalent, for some N and all sufficiently large n,

$$n(x_n) < N.$$

Now exactly as above choose $k(\{x_n\}) = P > \max\{0, N\}$ such that $z \epsilon K_{k(\{x_n\})}$ implies $\rho_M(\xi, z) < \inf\{\rho_\lambda(x, \xi)/2, x \epsilon \{x_n\}\} = R/2$. Then since $\{x_n\}$ is a ρ_λ-Cauchy sequence there is an $i > 0$ such that

$$\rho_\lambda(x_p, x_{p+j}) < R/2$$

for all $p > i$, and using (3.15) with $k(x) = P$, we have

$$\lambda^{-P}\rho_\lambda(x_p, x_{p+j}) \geq \rho_M(x_p, x_{p+j})$$

so that $\{x_n\}$ is a ρ_M-Cauchy sequence and the proof is complete.

The author thanks A. J. Goldman for his advice and encouragement.

4. References

[1] Meyers, P. R., Some Extensions of Banach's Contraction Theorem, J. Res. NBS **69B** (Math. and Math. Phys.) No. 3, 179–184 (1965).

[2] Kolmogorov and Fomin, Elements of the Theory of Functions and Functional Analysis 1, 43–45 (Rochester, Graylock, 1957).

[3] Janos, L., Two operations on distance functions, Notices A.M.S. **11**, No. 3, 614–7 (Aug. 1964); A converse of Banach's contraction theorem, Proc. A.M.S. **18**, No. 2, 287–289 (1967).

(Paper 71B2 & 3–202)