ON THE CONVERGENCE OF SUCCESSIVE APPROXIMATIONS FOR NONLINEAR FUNCTIONAL EQUATIONS

BY

FELIX E. BROWDER

It is the object of the present note to give a brief and transparent proof of the following generalization of the classical Picard–Banach contraction principle in its quantitative form:

Theorem 1. *Let X be a complete metric space, M a bounded subset of X, T a mapping of M into M. Suppose that there exists a monotone nondecreasing function $\psi(r)$ for $r \geqslant 0$, with ψ continuous on the right, such that $\psi(r) < r$ for all $r > 0$, while for all x and y in M,*

$$d(Tx, Ty) \leqslant \psi(d(x, y)),$$

(where d is the distance function on X).

Then: For each x_0 in M, $T^n x_0$ converges to an element ξ of X, independent of x_0, and

$$d(T^n x_0, \xi) \leqslant \psi^n(d_0)$$

where d_0 is the diameter of M, ψ^n is the n-th iterate of ψ, and

$$d_n = \psi^n(d_0) \to 0, \quad (n \to +\infty).$$

For the classical Picard–Banach theorem, $\psi(r) = \alpha r$ with $\alpha < 1$. We shall give sharper specializations of Theorem 1 below, as well as a discussion of its relation to other generalizations of the contraction principle in the literature.

We emphasize explicitly the importance of the explicit estimate given in Theorem 1 for the error term, since it is the explicit control over the error term in the Picard theorem which contributes so much to its widespread usefulness.

Proof of Theorem 1. For a fixed x_0 in M, let

$$x_j = T^j x_0, \, j \geqslant 1.$$

If d_0 is the diameter of M,

$$d(x_j, x_k) < d_0, \, (j, k \geqslant 0),$$

while by hypothesis,

$$d(x_j, x_k) = d(Tx_{j-1}, Tx_{k-1}) \leqslant \psi(d(x_{j-1}, x_{k-1}))$$

for all $j, k \geqslant 1$.

We set

$$\lambda_n = \sup_{j,k \geqslant n} |d(x_j, x_k)|.$$

Then

$$\lambda_n \leqslant \sup_{j,k \geqslant n} \psi(d(x_{j-1}, x_{k-1})) \leqslant \psi(\sup_{j,k \geqslant n} d(x_{j-1}, x_{k-1}))$$

by the monotonicity of ψ, i.e.

$$\lambda_n \leqslant \psi(\lambda_{n-1}), \ n \geqslant 1.$$

Iterating and applying the monotonicity of ψ, we see that

$$\lambda_n \leqslant \psi^n(\lambda_0) \leqslant \psi^n(d_0), \ (n \geqslant 1).$$

If $d_n = \psi^n(d_0)$, we note that

$$d_n = \psi(d_{n-1}) \leqslant d_{n-1}, \ d_n \geqslant 0.$$

Hence $d_n \to d_\infty$ for some $d_\infty \geqslant 0$. By the right continuity of ψ,

$$\psi(d_{n-1}) \to \psi(d_\infty),$$

and hence

$$d_\infty \leqslant \psi(d_\infty).$$

Since $\psi(r) < r$ for $r > 0$, it follows that $d_\infty = 0$, and hence that $\lambda_n \to 0$, i.e. $\{x_n\}$ is a Cauchy sequence in X and hence converges to an element ξ of the complete metric space X.

Finally, if y_0 is another point of M, $y_n = T^n y_0$, then by the same argument $y_n \to \xi_1$ for an element ξ_1 of X. Then

$$d(\xi, \xi_1) = \lim_n d(T^n x_0, T^n y_0) \leqslant \lim_n \psi(d(T^{n-1} x_0, T^{n-1} y_0)) \leqslant \lim_n \psi^n(d_0) = 0,$$

so that $\xi = \xi_1$. q.e.d.

Corollary to Theorem 1. *Under the hypotheses of Theorem 1, T can be extended in one and only one way to a continuous mapping of the closure of M in X into itself, and ξ is the unique fixed point of this extended mapping.*

We give some applications of Theorem 1 under variant hypotheses.

Theorem 2. *Let X be a complete metric space, M a bounded subset of X with diameter d_0. Suppose that T is a mapping of M into M and that for each $s \geqslant 0$, there exists $\Phi(s)$ with $0 \leqslant \Phi(s) < s$ for $s > 0$ such that for all x and y in M,*

$$d(Tx, Ty) \leqslant \Phi(d(x, y)).$$

Suppose further that on each compact subinterval $[\beta, d_0]$ of $[0, d_0]$ $(\beta > 0)$, the function $s^{-1}\Phi(s)$ is uniformly bounded by a constant $\theta(\beta) < 1$.

Then:

(a) *For each x_0 in M, the sequence $x_n = T^n x_0$ converges in X to a point ξ independent of the choice of x_0.*

(b) *Let $\psi(r)$ for $r > 0$ be defined by*

$$\psi(r) = \lim_{t \to r+}(\sup_{s \leqslant t} \Phi(s)).$$

Then for each $n \geqslant 1$,

$$d(T^n x_0, \xi) \leqslant \psi^n(d_0) = d_n,$$

where $d_n \to 0$ as $n \to +\infty$.

Proof of Theorem 2. It suffices to show that the hypotheses of Theorem 1 are satisfied for the given function $\psi(r)$.

The function

$$\psi_0(t) = \sup_{s \leqslant t} \Phi(s)$$

is monotone non-decreasing in t, so that

$$\psi(r) = \lim_{t \to r+}\psi_0(t)$$

is obviously both monotone non-decreasing and continuous from the right. Moreover, $\Phi(r) \leqslant \psi(r)$ for each $r > 0$, so that the inequality

$$d(Tx, Ty) \leqslant \psi(d(x, y))$$

holds for all x and y in M. It suffices therefore to show that for all $r > 0$, $\psi(r) < r$.

Let $r > 0$ be given, and choose β with $0 < \beta < r$. For $s \leqslant \beta$, we know that $\Phi(s) \leqslant s \leqslant \beta$, while by hypothesis there exists a constant $\theta(\beta) < 1$ such that for all $s \geqslant \beta$,

$$\Phi(s) \leqslant \theta(\beta)s.$$

Hence for $t > r$,

$$\psi_0(t) = \sup_{s \leqslant t} \Phi(s) \leqslant \max(\beta, \theta(\beta)t).$$

Hence

$$s^{-1}\psi(s) \leqslant \max(s^{-1}\beta, \theta(\beta)) < 1.$$

q.e.d.

As a specialization of Theorem 2, we have the following:

Theorem 3. (RAKOTCH [15]) *Let X be a complete metric space, M a subset of X, T a mapping of M into M. Suppose that for each $s > 0$ there exists $\psi(s) < s$ such that for x and y in M with $x \neq y$,*

$$d(Tx, Ty) \leqslant \psi(d(x, y)).$$

Suppose further that $s^{-1}\psi(s)$ is non-increasing in s for $s>0$. Then for each x_0 in M, $T^n x_0$ converges in X.

Proof of Theorem 3. Since $s^{-1}\psi(s)$ is less than 1 for each $s>0$ and is non-increasing, it is bounded by a constant $\theta(\beta)<1$ on each interval of the form $[\beta, +\infty]$ with $\beta>0$. Hence, to apply Theorem 2, it suffices to show that each point x_0 of M is contained in a bounded subset M_0 of M invariant under T, i.e. that the orbit of each point x_0 under T is bounded. This, however, follows from the following more general result:

Theorem 4. *Let M be a metric space, T a mapping of M into M such that there exists a function ψ with $\psi(s)<s$ for each $s>0$ such that for all x and y in M with $x \neq y$,*

$$d(Tx, Ty) \leqslant \psi(d(x, y)).$$

Suppose that x_0 is a point of M, and that there exists a constant $R>0$ such that for $r>R$,

$$r - \psi(r) > 2d(x_0, Tx_0).$$

Then the orbit of x_0 under T is a bounded subset of M of diameter at most R.

Proof of Theorem 4. It suffices to show that for each $n \geqslant 1$,

$$d(T^n x_0, x_0) \leqslant R.$$

Indeed, this last inequality shows that the orbit of x_0 is bounded, and since the hypothesis of the Theorem is invariant if one replaces x_0 by x_j for any $j \geqslant 1$, it will also follow that for $0 \leqslant j < k$,

$$d(x_j, x_k) \leqslant R.$$

For each $n \geqslant 1$, we have

$$d(T^n x_0, x_0) \leqslant d(T^n x_0, T^{n+1} x_0) + d(T^{n+1} x_0, Tx_0) + d(Tx_0, x_0) \leqslant$$
$$\leqslant 2d(x_0, Tx_0) + \psi(d(x_0, T^n x_0)).$$

Hence

$$d(x_0, T^n x_0) - \psi(d(x_0, T^n x_0)) \leqslant 2d(x_0, Tx_0),$$

and it follows that $d(x_0, T^n x_0) \leqslant R$.   q.e.d.

Another consequence of Theorem 2 is the following result of a type announced recently by BOYD and WONG [1]:

Theorem 5. *Let X be a complete metric space, M a subset of X, T a mapping of M into M such that there exists a function $\psi(r)$ for $r>0$ with $\psi(s)<s$ and $\psi(s)$ upper semi-continuous in s such that for all x and y in M with $x \neq y$,*

$$d(Tx, Ty) \leqslant \psi(d(x, y)).$$

Let x_0 be a point of M, and suppose that there exists an $R > 0$ such that for all $r > R$, $r - \psi(r) > 2d(x_0, Tx_0)$.

Then: $T^n x_0$ converges in X as $n \to +\infty$.

Theorem 5 follows obviously from Theorems 2 and 4, since $s^{-1} \psi(s)$ attains its upper bound on each compact subinterval $[\beta, d_0]$ with $\beta > 0$.

Remarks. (1) Theorem 1 and its proof is a specialization on the qualitative level of a mode of argument applied in the much more general context of *pseudo-metric* spaces by Kantorovich, Schröder, and others. (See KANTOROVICH [11], SCHRÖDER [17], [18], COLLATZ [7], WOUK [19]. Pseudo-metric spaces have a "metric" taking values in a cone in a partially ordered linear space.)

(2). Forms of the iteration method which work for arbitrary non-expansive operators in Hilbert space and certain other Banach spaces have been treated in BROWDER [2], [3], [4], BROWDER–PETRYSHYN [5], [6], PETRYSHYN [14], and OPIAL [13]. (For the compact and weakly continuous mappings in this class, see also KRASNOSELSKI [12] and SCHAEFER [16]). Results about iterates for contractive mappings have been considered in a number of papers by EDELSTEIN [8], [9], [10].

In conclusion, we note that a slightly sharper form of Theorem 2 holds when X is a Banach space.

Theorem 6. *Let X be a Banach space, M a bounded convex subset of X, T a mapping of M into M. Suppose that for each $s > 0$, there exists a least constant $\psi(s) < s$ such that if $d(x, y) \leqslant s$, then $d(Tx, Ty) \leqslant \psi(s)$. Then:*

(a) *For each x_0 in M, the sequence $T^n x_0$ converges to an element ξ of X.*

(b) *For each $n \geqslant 0$,* $$\| \xi - T^n x_0 \| \leqslant \psi^n(d_0) \to 0$$

as $n \to +\infty$, where d_0 is the diameter of M.

Proof of Theorem 6. For each x and y of M with $x \neq y$, we obviously have
$$d(Tx, Ty) \leqslant \psi(d(x, y)),$$

and moreover $\psi(r)$ is monotone non-decreasing in r. Hence, to apply Theorem 1 to obtain the conclusion of Theorem 6, it suffices to prove that the function $\psi(r)$ is continuous from the right.

Let $s > 0$ be fixed, and let $t > s$. If x and y are points of M with $d(x, y) = t$, it follows since M is convex that we can choose points x_1 and y_1 of M on the segment joining x to y such that the following conditions hold:
$$d(x, x_1) = \tfrac{1}{2}(t - s), \ d(x_1, y_1) = s, \ d(y_1, y) = \tfrac{1}{2}(t - s).$$

Since T is a non-expansive mapping on M, we have
$$d(Tx, Tx_1) \leqslant \tfrac{1}{2}(t - s), \ d(Ty, Ty_1) \leqslant \tfrac{1}{2}(t - s).$$

Hence
$$d(Tx, Ty) \leqslant d(Tx_1, Ty_1) + (t-s) \leqslant \psi(s) + (t-s).$$
Therefore
$$\psi(s) \leqslant \psi(t) \leqslant \psi(s) + (t-s),$$

and ψ is continuous (and indeed satisfies a Lipschitz condition with constant 1).

<div align="right">q.e.d.</div>

An addendum: After completing the previous part of the present note and making a further examination of the literature on the general topic of successive approximation techniques for nonlinear equations, we have noted that it would be useful to extend the above results by a simple argument to cover those theorems in which contractiveness hypotheses are imposed upon iterates of the mapping T rather than upon T itself. The primary example of such a result is the theorem of Cacciopoli (Atti. Accad. Naz. Lincei (6), 11 (1930), 794–799) which asserts the covergence of successive approximants $T^n x_0$ in a complete metric space X provided that for each $j \geqslant 1$, there exists a constant c_j such that

$$d(T^j x, T^j y) \leqslant c_j \, d(x, y)$$

for all x and y in M, where

$$\sum_{j=1}^{\infty} c_j < +\infty.$$

(This theorem was republished two decades later by J. Weissinger, Math. Nachr., 8 (1952), 193–212).

The following two theorems give much stronger results (which include the weakening of the Cacciopoli hypothesis to the simpler condition that for some m, $c_m < 1$):

Theorem 7. *Let M be a bounded subset of the complete metric space X, T an uniformly continuous mapping of M into M. Suppose that there exists a positive integer m and a monotone function $\psi(r)$ for $r \geqslant 0$, with ψ continuous on the right, such that*
$$\psi(r) < r, \text{ for all } r > 0,$$

while for all x and y in M,

$$d(T^m x, T^m y) \leqslant \psi(d(x, y)).$$

Then:

(a) *For each x_0 in M, $T^n x_0$ converges in X to a limit point ξ which is independent of the choice of the initial approximant x_0 in M.*

(b) *If T is extended continuously to a continuous mapping of the closure of M into itself, then ξ is the unique fixed point of the extended mapping T in the closure of M.*

(c) *For each x_0 in M,*
$$d(T^n x_0, \xi) \leqslant \beta(\psi^{[n/m]}(d_0)),$$
where:

$$d_0 = the\ diameter\ of\ M;$$

$$[n/m] = the\ integer\ part\ of\ (n/m),$$

$$\beta(r) = \max_{0 \leqslant j \leqslant m-1}\ \sup_{x, y \in M;\ d(x, y) \leqslant r} |d(T^j x, T^j y)|.$$

Proof of Theorem 7. We begin by applying Theorem 1 to the iterated mapping T^m of M into M. It follows from Theorem 1 that there exists an unique element ξ in X such that as $k \to +\infty$, $T^{mk} x_0 \to \xi$ for each x_0 in M. Moreover,
$$d(T^{mk} x_0, \xi) \leqslant \psi^k(d_0)$$

where d_0 is the diameter of M. We may assume without loss of generality that T is already extended by continuity to a continuous mapping of $cl(M)$, the closure of M in X, into $cl(M)$. Then, if we continue to denote this extended mapping as T, T^m is the continuous extension of T^m/M to a continuous mapping of $cl(M)$ into $cl(M)$ and ξ is the unique fixed point of T^m in $cl(M)$.

We note the elementary fact that if for a point p in $cl(M)$, $T^m(p) = p$, then
$$T^m(Tp) = T(T^m p) = Tp,$$

i.e. T maps the fixed point set of T^m in $cl(M)$ into itself. Since ξ is the unique fixed point of T^m in $cl(M)$, it follows that $T\xi = \xi$. Since, on the other hand, every fixed point of T is also a fixed point of T^m, we know that ξ is the unique fixed point of T in $cl(M)$.

By hypothesis, T is uniformly continuous as a mapping of M into M. It follows by induction that for each positive integer j, T^j is uniformly continuous as a mapping of M into M. Hence the function $\beta(r)$ of the conclusion (c) of Theorem 7, which is the maximum of the moduli of continuity of T^j for $0 \leqslant r \leqslant m-1$, satisfies the condition that

$$\beta(r) \to 0,\ as\ r \to 0.$$

Let n be a positive integer. We may write n in the form

$$n = mk + j,\ k = [n/m],\ 0 \leqslant j \leqslant m-1.$$

By a preceding remark, for each x_0 in M

$$d(T^{mk} x_0, \xi) \leqslant \psi^k(d_0).$$

Since $\beta(r)$ dominates the modulus of continuity of each T^j with $0 \leqslant j \leqslant m-1$, it follows that
$$d(T^n x_0, \xi) = d(T^j(T^{mk} x_0), T^j \xi) \leqslant \beta(\psi^k(d_0)).$$

By Theorem 1, $\psi^k(d_0) \to 0$ as $k \to +\infty$, i.e. as $n \to +\infty$. Since $\beta(r) \to 0$ as $r \to 0$, it follows that

$$\beta(\psi^k(d_0)) \to 0, \ (n \to +\infty).$$

Hence $T^n x_0$ converges to ξ, the unique fixed point of T in $cl(M)$, and the estimate of conclusion (c) holds.

<div align="right">q.e.d.</div>

Theorem 8. *Let M be a bounded subset of the complete metric space X, T a uniformly continuous mapping of M into M. Suppose that there exists a positive integer m and a function $\varphi(r)$ for $r \geqslant 0$ such that on each interval of the form $[\beta, d_0]$ with $\beta > 0$ $s^{-1}\varphi(s)$ is bounded from above by a constant $\theta_\beta < 1$. Suppose that for each x and y in M,*

$$d(T^m x, T^m y) \leqslant \varphi(d(x, y)).$$

Then:

(a) *For each x_0 in M, $T^n x_0$ converges to a point ξ in X, where ξ is independent of the choice of x_0 in M and is the unique fixed point of the mapping T extended continuously to $cl(M)$.*

(b) *For each x_0 in M,*

$$d(T^n x_0, \xi) < \beta(\psi^{[n/m]}(d_0)),$$

where:

$$d_0 = \text{the diameter of } M,$$

$$[n/m] = \text{the integer part of } (n/m),$$

$$\beta(r) = \max_{0 \leqslant j \leqslant m-1} \sup_{x, y \in M; \ d(x, y) \leqslant r} d(T^j x, T^j y),$$

$$\psi(r) = \lim_{t \to r+} \sup_{s \leqslant t} \varphi(s).$$

Proof of Theorem 8. As in the proof of Theorem 2 from Theorem 1, we show that under the hypotheses of Theorem 8,

$$d(T^m x, T^m y) < \psi(d(x, y)), \ (x, y \in M),$$

and that the function ψ satisfies the restrictions imposed in Theorem 7. The proof is identical in the latter respect with the proof of Theorem 2.

BIBLIOGRAPHY

1. BOYD, D. W. and J. S. W. WONG, On nonlinear contractions, Abstract 67T-453, Notices of the American Mathematical Society, June 1967.

2. BROWDER, F. E., Existence and approximation of solutions of nonlinear variational inequalities, Proc. Nat. Acad. Sci. U.S.A., **56**, 1080–1086 (1966).

3. ———, Convergence of approximants to fixed points of nonexpansive nonlinear mappings in Banach spaces, Archive Rat. Mech. and Anal., **24**, 82–90 (1967).

4. ————, Convergence theorems for sequences of nonlinear operators in Banach spaces, Math. Zeitschrift, 100, 201–225 (1967).

5. ———— and W. V. PETRYSHYN, The solution by iteration of nonlinear functional equations in Banach spaces, Bull. Amer. Math. Soc., 72, 517–575 (1966).

6. ———— and ————, Construction of fixed points of nonlinear mappings in Hilbert space, Journ. Math. Analysis and Appl. (to appear).

7. COLLATZ, L., Funktionalanalysis und Numerische Mathematik, Grundlehren der Math. Wiss. Band 120, Springer, Berlin, 1964.

8. EDELSTEIN, M., An extension of Banach's contraction principle, Proc. Amer. Math. Soc., 12, 7–10 (1961).

9. ————, On fixed and periodic points under contractive mappings, Jour. London Math. Soc., 37, 74–79 (1962).

10. ————, On nonexpansive mappings in Banach spaces, Proc. Cambr. Phil. Soc., 60, 439–447 (1964).

11. KANTOROVICH, L. V., The method of successive approximations for functional equations, Acta Math., 71, 63–97 (1939).

12. KRASNOSELSKI, M. A., Two remarks about the method of successive approximations, Uspekhi Mat. Nauk, 19, 123–127 (1955).

13. OPIAL, Z., Weak convergence of successive approximations of nonexpansive mappings in Banach spaces. Bull. Amer. Math. Soc. (to appear).

14. PETRYSHYN, W. V., Construction of fixed points of demicompact mappings in Hilbert space, Jour. Math. Analysis and Appl., 14, 276–284 (1966).

15. RAKOTCH, E., A note on contractive mappings, Proc. Amer. Math. Soc., 13, 459–465 (1962).

16. SCHAEFER, H., Über die Methode sukzessiver Approximationen, Jahresb. Deutsch. Math. Verein., 59, 131–140 (1957).

17. SCHRÖDER, J., Das Iterationsverfahren beim allgemeineren Abstandsbegriff, Math. Zeits., 66, 111–116 (1956).

18. ————, Nichtlineare Majoranten beim Verfahren der schrittweisen Näherung, Arch. Math., 7, 471–484 (1957).

19. WOUK, A., Direct iteration, existence, and uniqueness, Nonlinear integral equations, Madison Wisc., 3–34 (1964).