JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS **88**, 319–332 (1982)

Random Products of Contractions in Metric and Banach Spaces

RONALD E. BRUCK*

Department of Mathematics, University of Southern California, Los Angeles, California 90007

Submitted by K. Fan

Suppose (X, d) is a metric space and $\{T_0,..., T_N\}$ is a family of quasi-nonexpansive self-mappings on X. We give conditions sufficient to guarantee that every possible iteration of mappings drawn from $\{T_0,..., T_N\}$ converges. As a consequence, if $C_0,..., C_N$ are closed convex subsets of a Hilbert space with nonempty intersection, one of which is compact, and the proximity mappings are iterated in any order (provided only that each is used infinitely often), then the resulting sequence converges strongly to a point of the common intersection.

INTRODUCTION

This paper was motivated by the following question: Suppose H is a Hilbert space and $C_0,..., C_N$ are closed convex subsets of H with nonempty intersection. Denoting the proximity map of H on C_i by P_i, under what circumstances can we iterate $\{P_0, P_1,..., P_N\}$ randomly and obtain a convergent sequence? By a random iteration we mean one of the form

$$x_0 \in H,$$
$$x_n = P_{r(n)} x_{n-1} \qquad (n \geqslant 1),$$

where $\{r(n)\}$ is an arbitrary sequence drawn from $\{0,..., N\}$.

Prager [1] showed that such a random iteration always converges if H is finite-dimensional and the C_i are linear subspaces of H, while Amemiya and Ando [2] proved weak convergence when H is infinite-dimensional and the C_i are closed linear subspaces. Under recurrence, selection, or periodicity hypotheses—which are, of course, nonrandom—more is known (cf. [3–6]). On the other hand, nothing is known about random iterations when H is not a Hilbert space (but see [7, 8]).

We prove the strong convergence in Hilbert space when one of the C_i is

* Partially supported by NSF Grant MCS 78-02305-2.

0022-247X/82/080319–14$02.00/0

compact (and, of course, $\bigcap_i C_i \neq \varnothing$), provided the compact proximity mapping is used infinitely often in the iteration. We state a more general version of this, involving quasi-nonexpansive mappings; the crux of the argument is that a finitely generated semigroup of quasi-nonexpansive mappings is uniformly quasi-nonexpansive if each generator is strongly quasi-nonexpansive.

Replacing compactness with symmetry about the origin, we are able to prove a similar result—with weak convergence replacing strong convergence—but only for three sets. Strong convergence remains unresolved in this case, as does weak convergence for more than three sets or without symmetry.

1. Strongly Quasi-Nonexpansive Mappings

Throughout this section, (X, d) denotes a metric space (not necessarily compact or complete). A mapping $T: X \to X$ is said to be quasi-nonexpansive if for each f in $F(T)$, the fixed-point set of T, and for each x in X,

$$d(Tx, f) \leqslant d(x, f). \tag{1.1}$$

T is strictly quasi-nonexpansive if the inequality in (1.1) holds strictly when $Tx \neq x$. T is said to be strongly quasi-nonexpansive if for each f in $F(T)$ and $\varepsilon > 0$ there exists $\delta > 0$ such that

$$d(x, f) - d(Tx, f) < \delta \Rightarrow d(x, Tx) < \varepsilon. \tag{1.2}$$

(This implies T is quasi-nonexpansive.) Equation (1.2) is analogous to strong nonexpansiveness, introduced in Bruck and Reich [6]; indeed, a mapping strongly nonexpansive in the sense of [6] is strongly quasi-nonexpansive in the sense of (1.2), uniformly in $f \in F(T)$.

Finally, a family $\mathscr{S}$ of self-mappings of X is said to be uniformly quasi-nonexpansive if each T in $\mathscr{S}$ is quasi-nonexpansive, the common fixed-point set $F(\mathscr{S}) = \bigcap \{F(T): T \in \mathscr{S}\}$ is nonempty, and for each f in $F(\mathscr{S})$ and $\varepsilon > 0$ there exists $\delta > 0$ such that (1.2) holds for each T in $\mathscr{S}$.

Paralleling the proofs of Proposition 1.1 and Lemma 2.1 of [6], we easily prove:

Lemma 1.1. *If T_1 and T_2 are strongly quasi-nonexpansive and $F(T_1) \cap F(T_2) \neq \varnothing$, then $T_1 T_2$ is strongly quasi-nonexpansive and $F(T_1 T_2) = F(T_1) \cap F(T_2)$.*

If $\{T_0, \ldots, T_N\}$ are quasi-nonexpansive mappings on X, we denote by $\langle T_0, \ldots, T_N \rangle$ the multiplicative semigroup generated by $\{T_0, \ldots, T_N\}$ (including the identity I). The lemma guarantees that if each T_i is strongly quasi-nonexpansive then each element of $\langle T_0,\ldots, T_N\rangle$ is also strongly quasi-nonexpansive, provided the T_i have a common fixed-point. Establishing the uniform quasi-nonexpansiveness of $\langle T_0,\ldots, T_N\rangle$ seems to require an additional hypothesis. We say that T satisfies condition (P) provided

$$d(x_n, Tx_n) \to 0 \text{ as } n \to \infty \Rightarrow \{x_n\} \text{ has a convergent subsequence.}$$

For continuous T this is equivalent to: $F(T)$ is compact and for each $\varepsilon > 0$ there exists $\delta > 0$ such that $d(x, Tx) < \delta \Rightarrow d(x, F(T)) < \varepsilon$. When T is the proximity map of a Hilbert space onto a closed convex subset C, for example, (P) is equivalent to the compactness of C.

An idea of Amemiya and Ando [2] can be used to prove:

LEMMA 1.2. *Suppose a semigroup of quasi-nonexpansive self-mappings of X is finitely generated. If the generators have a common fixed-point, are uniformly continuous, are strongly quasi-nonexpansive, and satisfy condition* (P), *then the semigroup is uniformly quasi-nonexpansive.*

We shall prove a more precise version of Lemma 1.2, since we do not want to assume that each generator satisfies (P):

LEMMA 1.2′. *Suppose $T_0,\ldots, T_N$ are strongly quasi-nonexpansive, uniformly continuous, and have a common fixed-point. If T_0 satisfies* (P) *then $\langle T_0,\ldots, T_N\rangle$ is* "T_0-front loaded" *uniformly quasi-nonexpansive in the sense that for each $\varepsilon > 0$ and common fixed-point f, there exists $\delta > 0$ such that for any S in $\langle T_0,\ldots, T_N\rangle$,*

$$d(x, f) - d(ST_0x, f) < \delta \Rightarrow d(x, ST_0x) < \varepsilon.$$

Proof. By induction on N. The case $N = 0$ follows because

$$d(x, f) - d(T_0x, f) \leqslant d(x, f) - d(T_0^nx, f)$$

and because $d(x, T_0x)$ small implies $d(x, F(T_0))$ is small (by condition (P)), which in turn implies $d(x, T_0^nx)$ is small.

Now assume the lemma is true for a certain $N \geqslant 0$, and consider generators $T_0,\ldots, T_{N+1}$. Suppose we are given f_0 in $\bigcap_{k=0}^{N+1} F(T_k)$ and $\varepsilon > 0$. Note first that we can choose $\alpha > 0$ such that

$$d(x, T_kx) < \alpha \text{ for } 0 \leqslant k \leqslant N+1 \Rightarrow d\left(x, \bigcap_{k=0}^{N+1} F(T_k)\right) < \varepsilon/2. \qquad (1.3)$$

(If this were not so, we could find a sequence $\{x_n\}$ such that $d(x_n, T_k x_n) \to 0$ as $n \to \infty$ for each k, yet

$$d\left(x_n, \bigcap_k F(T_k)\right) \geqslant \varepsilon/2. \tag{1.4}$$

Since T_0 satisfies (P), some subsequence of $\{x_n\}$ converges; without loss of generality we may assume $\{x_n\}$ itself converges to some x^*. Since the T_k are continuous and $d(x_n, T_k x_n) \to 0$, $x^* \in F(T_k)$ for each k, surely contradicting (1.4).)

Second, by uniform continuity we can choose $\xi > 0$ such that

$$d(x, y) < \xi \Rightarrow d(T_k x, T_k y) < \alpha/3 \tag{1.5}$$

for each $0 \leqslant k \leqslant N+1$.

Finally, by the induction hypothesis we can choose $\delta > 0$ such that if Q is generated by a proper subset of $\{T_0, \ldots, T_{N+1}\}$, then

$$d(x, f_0) - d(QT_0 x, f_0) < \delta \Rightarrow d(x, QT_0 x) < \min\{\alpha/3, \xi\}, \tag{1.6}$$

moreover, since each T_k is strongly quasi-nonexpansive, such that

$$d(x, f_0) - d(T_k x, f_0) < \delta \Rightarrow d(x, T_k x) < \alpha/3 \qquad (0 \leqslant k \leqslant N+1). \tag{1.7}$$

Now let $S \in \langle T_0, \ldots, T_{N+1} \rangle$ and suppose $d(x, f_0) - d(ST_0 x, f_0) < \delta$. Without loss of generality we may assume all of the T_k are needed to generate S. For any T_k, then, we can find P_k in $\langle T_0, \ldots, T_{N+1} \rangle$ and Q_k in $\langle T_0, \ldots, \hat{T}_k, \ldots, T_{N+1} \rangle$ such that $S = P_k T_k Q_k$ (i.e., Q_k consists of the part of some factorization of S which is prior to the first occurrence of T_k; it may happen that $Q_k = I$).

Now

$$d(x, T_k x) \leqslant d(x, Q_k T_0 x) + d(Q_k T_0 x, T_k Q_k T_0 x) + d(T_k Q_k T_0 x, T_k x). \tag{1.8}$$

Since

$$d(Q_k T_0 x, f_0) - d(T_k Q_k T_0 x, f_0) \leqslant d(x, f_0) - d(ST_0 x, f_0) < \delta,$$

we have by (1.7)

$$d(Q_k T_0 x, T_k Q_k T_0 x) < \alpha/3. \tag{1.9}$$

Since we also have

$$d(x, f_0) - d(Q_k T_0 x, f_0) \leqslant d(x, f_0) - d(ST_0 x, f_0) < \delta,$$

by (1.6) we have

$$d(x, Q_k T_0 x) < \alpha/3, \tag{1.10}$$

and

$$d(x, Q_k T_0 x) < \xi.$$

By (1.5) the latter implies

$$d(T_k x, T_k Q_k T_0 x) < \alpha/3. \tag{1.11}$$

Summing (1.9), (1.10), and (1.11) and using (1.8) thus yields

$$d(x, T_k x) < \alpha \qquad (0 \leqslant k \leqslant N+1).$$

In view of (1.3), this means

$$d\left(x, \bigcap_{k=0}^{N+1} F(T_k)\right) < \varepsilon/2,$$

or finally, since

$$d(x, ST_0 x) \leqslant d(x, f) + d(ST_0 x, f) \leqslant 2d(x, f)$$

for any f in $\bigcap_{k=0}^{N+1} F(T_k)$, we have

$$d(x, ST_0 x) < \varepsilon. \qquad \text{Q.E.D.}$$

The application of Lemma 1.2 to random iterations is immediate:

THEOREM 1.1. *Suppose $T_0, \ldots, T_N$ are strongly quasi-nonexpansive, uniformly continuous, and have a common fixed-point. If T_0 satisfies condition* (P), *then any random iteration*

$$x_0 \in X,$$
$$x_n = T_{r(n)} x_{n-1} \qquad (n > 1)$$

for which $r(n) = 0$ for infinitely many n, converges to a point of $\limsup_{n \to \infty} F(T_{r(n)})$.

Proof. Let f be a common fixed-point of the T_i. Since the T_i are quasi-nonexpansive, $\{d(x_n, f)\}$ is a nonincreasing sequence, and therefore converges to a limit. It follows from the strong quasi-nonexpansiveness of the mappings T_i that $\lim_{n \to \infty} d(x_n, x_{n-1}) = 0$.

Whenever k is an integer such that $n(k) = 0$, we have for each $n > k$

$$x_n = S_n T_0 x_{k-1}$$

for some S_n in $\langle T_0, \ldots, T_N\rangle$. By Lemma 1.2′, therefore, if $\varepsilon > 0$ is given we can choose k so large that

$$d(x_n, x_{k-1}) < \varepsilon \qquad \text{for all} \quad n > k.$$

This obviously implies $\{x_n\}$ is Cauchy. We do not need to assume X is complete, because for infinitely many k we have $x_k = T_0 x_{k-1}$, while $\lim_{n\to\infty} d(x_n, x_{n-1}) = 0$; since T_0 satisfies (P), $\{x_n\}$ has a convergent subsequence, and hence $\{x_n\}$ itself converges.

Let x^* denote the limit. Since $\lim_{n\to\infty} d(x_n, x_{n-1}) = 0$, for any i for which $r(n) = i$ infinitely often we have $T_i x^* = x^*$, i.e., $x^* \in \limsup_{n\to\infty} F(T_{r(n)})$. Q.E.D.

COROLLARY 1.1. *Suppose (X, d) is a compact metric space and $\{T_0, \ldots, T_N\}$ are continuous, strictly quasi-nonexpansive self-mappings of X with a common fixed-point. Then every random iteration drawn from $\{T_0, \ldots, T_N\}$ converges.*

Proof. It is a simple exercise to show that a continuous, strictly quasi-nonexpansive mapping on a compact space satisfies condition (P) and is strongly quasi-nonexpansive and uniformly continuous. Q.E.D.

COROLLARY 1.2. *Suppose $C_0, \ldots, C_N$ are closed convex subsets of a Hilbert space H with nonempty intersection. Let P_i denote the proximity map of H on C_i. If C_0 is compact, then any random iteration of $\{P_0, \ldots, P_N\}$ which uses P_0 infinitely often, converges strongly.*

Proof. The proximity map of H onto a closed convex subset C of H is the resolvent $(I + A)^{-1}$ of the subdifferential A of the indicator function of C; and as such is strongly nonexpansive (cf. [6]). In particular, proximity mappings in Hilbert space are strongly quasi-nonexpansive. Thus the corollary follows directly from the theorem. Q.E.D.

Of course in both corollaries the limit is in the lim sup of the fixed-point sets.

2. A NON-COMPACT CASE

Compactness plays an important role in Theorem 1.1 via Condition (P). It remains an open problem whether compactness is necessary in Corollary 1.2, but in its absence we do have a curious special case.

THEOREM 2.1. *Suppose H is a Hilbert space and C_0, C_1, C_2 are symmetric closed convex subsets of H (that is, $C_i = -C_i$). If the proximity* *mappings P_0, P_1, P_2 are iterated in any order, then the resulting sequence converges weakly.*

It will be seen that the restriction to three sets is critical to our method of proof. This based on more explicit estimate possible in the case of two sets:

LEMMA 2.1. *Let P_0, P_1 be the proximity mappings of a Hilbert space H onto symmetric closed convex sets C_0, C_1. Then the iteration*

$$x_0 \in H,$$
$$x_{2n+1} = P_0 x_{2n}, \qquad x_{2n+2} = P_1 x_{2n+1} \qquad (n \geqslant 0) \tag{2.1}$$

converges strongly, with estimate

$$\|x_n - x_m\|^2 \leqslant 3\,\|x_m\|^2 - 3\,\|x_n\|^2 \qquad (n \geqslant m \geqslant 1). \tag{2.2}$$

Proof. Put $f(x) = \max\{f_0(x), f_1(x)\}$, where $f_i(x) = 1/2\, d(x, C_i)^2$. Being the maximum of two continuous convex functions, f itself is a continuous convex function. It is well-known that $\operatorname{grad} f_i(x) = x - P_i(x)$, and it is easy to see that the subdifferential ∂f of f, defined by

$$\partial f(x) = \{w \in H: f(y) \geqslant f(x) + (w, y - x) \text{ for all } y \text{ in } H\},$$

is given by

$$\partial f(x) = \operatorname{co}\{\operatorname{grad} f_i(x): f_i(x) = f_i(x)\}$$

(where co denotes the convex hull). In particular, $\partial f(x) = x - P_i(x)$ if x is in C_i and not in C_{1-i}. Thus the iteration defined by (2.1) assumes the form

$$x_{n+1} \in x_n - \partial f(x_n) \qquad (n \geqslant 0). \tag{2.3}$$

Moreover, since C_0 and C_1 are symmetric, f is an even function.

It is also well-known that P_0 and P_1 are firmly nonexpansive, i.e.,

$$\|(x-y) - (P_i x - P_i y)\|^2 \leqslant \|x - y\|^2 - \|P_i x - P_i y\|^2.$$

Taking $x = x_n$ and $y = 0$, we find

$$\|x_{n-1} - x_n\|^2 \leqslant \|x_n\|^2 - \|x_{n+1}\|^2. \tag{2.4}$$

It follows that $\{\|x_n\}$ is nonincreasing. So is $\{\|x_n - x_{n+1}\|\}$; for when n is odd we have $x_{n+1} = P_1 x_n$, and $\|x_n - x_{n+1}\| = \|x_n - P_1 x_n\| \leqslant \|x_n - x_{n-1}\|$ since $x_{n-1} = P_1 x_{n-2} \in C_1$ and $P_1 x_n$ is the point of C_1 closest to x_n. A similar argument for n even completes the proof that $\{\|x_n - x_{n+1}\|\}$ is nonincreasing.

Thus (2.1) is a descent method for f: $f(x_{n+1}) \leqslant f(x_n)$ for all n. Let $n \geqslant i \geqslant 1$. Then

$$f(x_i) \geqslant f(x_n) = f(-x_n),$$

since f is even, while by the subdifferential inequality,

$$f(-x_n) \geqslant f(x_i) + (\partial f(x_i), -x_n - x_i).$$

Combining these and noting that $x_i - x_{i+1} \in \partial f(x_i)$ by (2.3), we conclude that

$$(x_i - x_{i+1}, x_n + x_i) \geqslant 0 \qquad \text{whenever} \quad n \geqslant i \geqslant 1. \tag{2.5}$$

But surely for $n > m$ we have the identity

$$\begin{aligned}\|x_m - x_n\|^2 &= \|x_m\|^2 - \|x_n\|^2 + 2(x_n - x_m, x_n)\\ &= \|x_m\|^2 - \|x_n\|^2 + 2\sum_{i=m}^{n-1} (x_{i+1} - x_i, x_n).\end{aligned}$$

Thus by (2.5) we have

$$\begin{aligned}\|x_n - x_m\|^2 &\leqslant \|x_m\|^2 - \|x_n\|^2 + 2\sum_{i=m}^{n-1} (x_i - x_{i+1}, x_i)\\ &= 2(\|x_m\|^2 - \|x_n\|^2) + \sum_{i=m}^{n-1} \|x_i - x_{i+1}\|^2. \end{aligned} \tag{2.6}$$

We are finally led to (2.2) by combining (2.4) with (2.6). Q.E.D.

The identity used in deriving (2.6) was used in a similar context by McCormick and Rodrigue [9].

Proof of Theorem 2.1. Consider the iteration

$$x_0 \in H,$$

$$x_n = P_{r(n)} x_{n-1}, \qquad n \geqslant 1.$$

If only two of the indices 0, 1, 2 appear infinitely often in $\{r(n)\}$ then the iteration is essentially of the form $P_0 y$, $P_1 P_0 y$, $P_0 P_1 P_0 y, \ldots$, since $P_i^2 = P_i$. It is known (cf. [6]) that in this case $\{x_n\}$ converges strongly to a point of $C_0 \cap C_1$.

Thus we may assume all three indices 0, 1, 2 appear infinitely often in $\{r(n)\}$. Consider a positive integer m, and find the smallest integer $n \geqslant m$ with the property that $x_{n+1} = P_2 x_n$. This means that in the iteration, the terms between x_m and x_n were obtained by iterating only P_0 and P_1, and thus by Lemma 2.1,

$$\|x_n - x_m\|^2 \leqslant 3\,\|x_m\|^2 - 3\,\|x_n\|^2. \tag{2.7}$$

Thus

$$\begin{aligned}\|x_m - P_2 x_m\| &\leqslant \|x_m - x_n\| + \|x_n - P_2 x_n\| + \|P_2 x_n - P_2 x_m\| \\ &\leqslant 2\,\|x_m - x_n\| + \|x_n - x_{n+1}\|\end{aligned}$$

(since $x_{n+1} = P_2 x_n$ and P_2 is nonexpansive). Therefore, by the Cauchy–Schwarz inequality and (2.4), (2.7) we have

$$\begin{aligned}\|x_m - P_2 x_m\|^2 &\leqslant 6\,\|x_n - x_m\|^2 + 3\,\|x_n - x_{n+1}\|^2 \\ &\leqslant 18\,\|x_m\|^2 - 18\,\|x_{n+1}\|^2.\end{aligned}$$

(The constant 18 can be improved.) Since $\{\|x_k\|\}$ converges, therefore $\lim_{m\to\infty} \|x_m - P_2 x_m\| = 0$. By symmetry we now have

$$\lim_{m\to\infty} \|x_m - P_i x_m\| = 0 \qquad (i = 0, 1, 2).$$

It follows that all weak subsequential limits of $\{x_m\}$ lie in $C_0 \cap C_1 \cap C_2$.

By a theorem of Reich [10], however, $\lim_{m\to\infty}(x_m, v_1 - v_2)$ exists for every v_1, v_3 in $\bigcap C_i$. It readily follows that $\{x_m\}$ can have at most one weak subsequential limit in $\bigcap C_i$. Thus $\{x_m\}$ does indeed converge weakly, as claimed.

Q.E.D.

3. Interleaving Iterations

Throughout this section (X, d) again denotes a metric space. A special role is played by mappings which satisfy, for some $f_0 \in X$ and increasing function λ: $[0, \infty) \to [0, \infty)$,

$$d(x, Tx) \leqslant \lambda(d(x, f_0)) - \lambda(d(Tx, f_0)) \tag{3.1}$$

for all $x \in X$. Their importance arises from an observation made by Golub *et al.* [4]: if X is a Hilbert space, C is a closed convex subset of X, and the closed ball of radius $\delta > 0$, centered at f_0, is entirely contained in C, then there exists $\lambda = \lambda_\delta$ such that (3.1) holds for $T =$ the proximity map of X on C. Indeed, if T is a nonexpansive mapping in Hilbert space with $B_\delta(f) \subset F(T)$, then

$$2\delta\,\|x - Tx\| \leqslant \|x - f\|^2 - \|Tx - f\|^2.$$

The importance of (3.1) in iterations arises from:

LEMMA 3.1. *Suppose two mappings T_1 and T_2 satisfy* (3.1) *for the same λ and f_0. Then T_1T_2 also satisfies* (3.1), *with the same λ and f_0.*

Proof. This is trivial, because

$$\begin{aligned} d(x, T_1T_2x) &\leqslant d(x, T_2x) + d(T_2x, T_1T_2x) \\ &\leqslant \lambda(d(x,f_0)) - \lambda(d(T_2x,f_0)) + \lambda(d(T_2x,f_0)) - \lambda(d(T_1, T_2x,f_0)) \\ &= \lambda(d(x,f_0)) - \lambda(d(T_1T_2x,f_0)). \qquad \text{Q.E.D.} \end{aligned}$$

We shall call a sequence $\{T_n\}$ of self-mappings of X iteration-normal if for all $x_0 \in X$ the iterates

$$x_n = T_nx_{n-1} \qquad (n \geqslant 1)$$

converge and the same is true of every shift $\{T_k, T_{k+1},...\}$ of the original sequence.

THEOREM 3.1. *Let (X, d) be complete. If $\{T_n\}$ is iteration-normal, each T_n is nonexpansive, and $\{P_n\}$ is a sequence of self-mappings of X which satisfy* (3.1) *for some increasing $\lambda: [0, \infty) \to [0, \infty)$ and some $f_0 \in \bigcap_{n=1}^{\infty} F(T_n)$, then the sequence $\{T_1, P_1, T_2, P_2,...\}$ obtained by* "shuffling" *$\{T_n\}$ with $\{P_n\}$ is also iteration-normal.*

Proof. Since the hypotheses are invariant under a shift, it suffices to show that

$$\{x_0, P_1x_0, T_1P_1x_0, P_2T_1P_1x_0,...\} \tag{3.2}$$

converges for any $x_0 \in X$. To this end, define $\{x_n\}$ by

$$x_n = T_nP_nx_{n-1} \qquad (n \geqslant 1).$$

Thus $\{x_n\}$ is the sequence of even-numbered terms of (3.2) (taking x_0 as the zeroth). Temporarily fix m and define $y_m, y_{m+1},...$ by

$$\begin{aligned} y_m &= x_m, \\ y_n &= T_ny_{n-1} \qquad (n > m). \end{aligned}$$

For any $i > m$,

$$\begin{aligned} d(x_i, y_i) &= d(T_iP_ix_{i-1}, T_iy_{i-1}) \\ &\leqslant d(P_ix_{i-1}, y_{i-1}) \\ &\leqslant d(x_{i-1}, y_{i-1}) + d(x_{i-1}, P_ix_{i-1}) \\ &\leqslant d(x_{i-1}, y_{i-1}) + \lambda(d(x_{i-1},f_0)) - \lambda(d(P_ix_{i-1},f_0)), \end{aligned}$$

since T_i is nonexpansive. Because f_0 is a fixed-point of T_i this leads to

$$d(x_i, y_i) \leqslant d(x_{i-1}, y_{i-1}) + \lambda(d(x_{i-1}, f_0)) - \lambda(d(x_i, f_0)) \qquad (i > m).$$

Summing for $i = m + 1, \ldots, n$ and telescoping, we get

$$d(x_n, y_n) \leqslant \lambda(d(x_m, f_0)) - \lambda(d(x_n, f_0)) \qquad (n > m). \tag{3.3}$$

We deduce from (3.3) that $\{\lambda(d(x_n, f_0))\}$ is nonincreasing, hence convergent. Put $L = \lim \lambda(d(x_n, f_0))$. Then (3.3) implies

$$\limsup_{n \to \infty} d(x_n, y_n) \leqslant \lambda(d(x_m, f_0)) - L.$$

Since $\{T_n\}$ is iteration-normal, however, $\{y_m\}$ converges to a point f_m (which depends on m, in general); and we therefore have

$$\limsup_{n \to \infty} d(x_n, f_m) \leqslant \lambda(d(x_m, f_0)) - L.$$

Since L is independent of m, this shows $\{f_m\}$ is Cauchy and hence convergent. For $f = \lim f_m$ we therefore have

$$\begin{aligned}\limsup_{n \to \infty} d(x_n, f) &\leqslant \limsup_{n \to \infty} d(x_n, f_m) + d(f_m, f) \\ &\leqslant \lambda(d(x_m, f_0)) - L + d(f_m, f)\end{aligned}$$

for any $m > 0$—which proves $\{x_n\}$ converges to f.

The odd-numbered terms of (3.2) are given by $\{P_n x_n\}$; and since

$$\begin{aligned}d(x_n, P_n x_n) &\leqslant \lambda(d(x_n, f_0)) - \lambda(d(P_n, x_n, f_0)) \\ &\leqslant \lambda(d(x_n, f_0)) - \lambda(d(x_{n+1}, f_0))\end{aligned}$$

we have $\lim d(x_n, P_n x_n) = 0$, finally proving that (3.2) converges. Q.E.D.

Note that if we take each $T_n = I$, Theorem 3.1 asserts that $\{P_n\}$ can be iterated in any order, with the resulting sequence converging. Note also that we do not identify the limit in terms of the fixed-point sets of the T_n, P_n.

When X is a subset of a Banach space we shall call a sequence $\{T_n\}$ of self-mappings of X weakly iteration-normal if for each x_0 in X the iteration $x_n = T_n x_{n-1}$ $(n \geqslant 1)$ converges weakly, and the same is true of every shifted sequence $\{T_k, T_{k+1}, \ldots\}$. Using the weak lower-semicontinuity of the norm, we can prove the following by an argument similar to that of Theorem 3.1.

THEOREM 3.2. *Suppose X is a weakly compact subset of a Banach space and $\{T_n\}$ is weakly iteration-normal on X. Suppose $\{P_n\}$ is a sequence of self-mappings of X which satisfy* (3.1) *for some increasing λ and some $f_0 \in \bigcap_{n=1}^{\infty} F(T_n)$. Then $\{T_1, P_1, T_2, P_2,...\}$, obtained by* "shuffling" *$\{T_n\}$ with $\{P_n\}$, is also weakly iteration-normal.*

We omit the proof. We point cut that in both theorems a more general kind of "shuffling" can be permitted: it suffices that every term of $\{T_n\}$ is used in the shuffled sequence, and in the same order, whereas any number of P's can be used, including none, and without preserving order. This is because the identity map satisfies (3.1), and any product of maps satisfying (3.1) also satisfies (3.1).

COROLLARY 3.1. *Suppose $\{C_\alpha: \alpha \in A\}$ is a family of closed convex sets in Hilbert space, while K_0, K_1, K_2 are symmetric closed convex sets with $K_0 \cap K_1 \cap K_2 \cap \operatorname{int} \bigcap \{C_\alpha: \alpha \in A\} \neq \varnothing$. Then when the proximity mappings of these sets are iterated in any order, the resulting sequence converges weakly.*

Proof. Combine Theorem 3.2 with Theorem 2.1.

4. REMARKS

The essential restriction in our results on random iterations, that the pool from which the mappings are drawn be finite, can be relaxed but not eliminated. Consider, for example, rays emanating from the origin in R^2, making angles $0 = \xi_0 < \xi_1 < \cdots < \xi_n < \cdots$ with the positive x-axis, where $\xi_n \to \infty$ and $\Delta\xi_n = \xi_n - \xi_{n-1} < 1$ for all n. If we begin with a point on the positive x-axis and successively project on these rays, the resulting vector x_n makes an angle ξ_n with the positive x-axis, and has length

$$\|x_n\| = \|x_0\| \prod_{i=1}^{n} \cos \Delta\xi_i \geqslant \|x_0\| \prod_{i=1}^{n} (1 - (\Delta\xi_i)^2/2).$$

So if we choose $\{\xi_n\}$ so $\sum (\Delta\xi_n)^2 < \infty$, the set of accumulation points of $\{x_n\}$ consists of a circle of positive radius; in any event, $\{x_n\}$ does not converge.

However, the number of mappings may be infinite provided the number of distinct fixed-point sets is finite. More precisely, the following version of Lemma 1.2′ remains valid:

LEMMA 4.1. *Suppose (X, d) is a metric space and $\mathscr{F}$ is an equicon-*

tinuous, uniformly quasi-nonexpansive family with a common fixed-point. If $\{F(T)\colon T \in \mathcal{S}\}$ *is finite, each* $F(T)$ *is compact, and* $\mathcal{S}$ *satisfies condition* (P) *uniformly*:

$$\text{for each } \varepsilon > 0 \text{ there exists } \delta > 0 \text{ such that}$$

$$T \in \mathcal{S},\ d(x, Tx) < \delta \Rightarrow d(x, F(T)) < \varepsilon,$$

then the generated semigroup $\langle \mathcal{S} \rangle$ *is uniformly quasi-nonexpansive.*

We omit the proof, remarking only that it is by induction on the number of distinct fixed-point sets; otherwise it is very similar to the proof of Lemma 1.2′. (Note that the hypothesis of equicontinuity, which was equivalent to uniform continuity in Lemma 1.2′, is needed here.)

One application of Lemma 4.1 is the following: suppose X is a uniformly convex Banach space and $A_0, A_1,..., A_N$ are m-accretive mappings. Define the resolvents by $J_\lambda^i = (I + \lambda A_i)^{-1}$, and for $\delta > 0$ put $\mathcal{S}_\delta = \{J_\lambda^i : \lambda \geqslant \delta, 0 \leqslant i \leqslant N\}$. If each A_i has a compact resolvent, and $\bigcap A_i^{-1}(0) \neq \varnothing$, then $\mathcal{S}_\delta$ satisfies the hypotheses of Lemma 4.1 and, accordingly, mappings drawn from $\mathcal{S}_\delta$ can be iterated in any order, with the resulting sequence converging strongly. The uniform quasi-nonexpansiveness can be proved as in [6], whereas the uniform (P) condition follows from the resolvent identity

$$J_\lambda x = J_\mu \left(\left(1 - \frac{\mu}{\lambda}\right) J_\lambda x + \frac{\mu}{\lambda} x \right).$$

(It even suffices to assume only that A_0 has compact resolvents, provided infinitely many are used in the iteration.)

Note added in proof. The argument of Lemma 1.2′ that $d(x, T_k x) < \alpha$ for $0 \leqslant k \leqslant N + 1$ is valid only for $0 < k \leqslant N + 1$ (because the induction hypothesis (1.6) should be that Q is generated by a proper subset of $\{T_0,..., T_{N+1}\}$, *including* T_0). The case $k = 0$ has a trivial proof, however: since

$$d(x, f) - d(T_0 x, f) \leqslant d(x, f) - d(ST_0 x, f) < \delta,$$

by (1.7) we actually have $d(x, T_0 x) < \alpha/3$.

REFERENCES

1. M. PRAGER, On a principle of convergence in a Hilbert space, *Czechoslovak Math. J.* **10** (1960), 271–282.
2. I. AMEMIYA AND T. ANDO, Convergence of random products of contractions in Hilbert space, *Acta Sci. Math.* (*Szeged*) **26** (1965), 239–244.
3. L. M. BREGMAN, The method of successive projection for finding a common point of convex sets, *Dokl. Akad. Nauk. SSSR* **162** (1965), 487–490.

4. L. G. GUBIN, B. T. POLYAK, AND E. V. RAIK, The method of projections for finding the common point of convex sets, *U.S.S.R. Computational Math. Math. Phys.* **7** (1967), 1–24.
5. F. E. BROWDER AND W. V. PETRYSHYN, Construction of fixed points of nonlinear mappings in Hilbert space, *J. Math. Anal. Appl.* **20** (1967), 197–228.
6. R. E. BRUCK AND S. REICH, Nonexpansive projections and resolvents of accretive operators in Banach spaces, *Houston J. Math.* **3** (1977), 459–470.
7. B. ATLESTAM AND F. SULLIVAN, Iteration with best approximation operators, *Rev. Roumaine Math. Pures Appl.* **21** (1976), 125–131.
8. F. DEUTSCH, The alternating method of Von Neumann, *in* "Multivariate Approximation Theory," pp. 83–96, Birkhäuser, Basel, 1979.
9. S. F. MCCORMICK AND G. H. RODRIGUE, A uniform approach to gradient methods for linear operator equations, *J. Math. Anal. Appl.* **49** (1975), 275–285.
10. S. REICH, Weak convergence theorems for nonexpansive mappings in Banach spaces, *J. Math. Anal. Appl.* **67** (1979), 274–276.