ISRAEL JOURNAL OF MATHEMATICS, Vol. 32, No. 1, 1979

STRONG CONVERGENCE OF CONTRACTION SEMIGROUPS AND OF ITERATIVE METHODS FOR ACCRETIVE OPERATORS IN BANACH SPACES†

BY

OLAVI NEVANLINNA†† AND SIMEON REICH‡

ABSTRACT

Let A be an m-accretive operator in a Banach space E. Suppose that $A^{-1}0$ is not empty and that both E and E^* are uniformly convex. We study a general condition on A that guarantees the strong convergence of the semigroup generated by $-A$ and of related implicit and explicit iterative schemes to a zero of A. Rates of convergence are also obtained. In Hilbert space this condition has been recently introduced by A. Pazy. We also establish strong convergence under the assumption that the interior of $A^{-1}0$ is not empty. In Hilbert space this result is due to H. Brezis.

1. Introduction

In his recent study [14] of strong convergence of nonlinear contraction semigroups in Hilbert space, A. Pazy has introduced a general condition on the generator of a semigroup S which guarantees the strong convergence of $S(t)x$ as $t \to \infty$ for each x in the domain of S. One of our goals in the present paper is to show that his approach also works outside Hilbert space. For simplicity, we assume throughout most of the paper that both the Banach space and its dual are uniformly convex.

In Section 2 we define the convergence condition and present some examples. In Section 3 we establish the strong convergence of trajectories of contraction semigroups whose generators satisfy the convergence condition. In Sections 4 and 5 we consider discrete implicit and explicit iterative schemes for finding

† Sponsored by the United States Army under Contract No. DAAG29-75-C-0024.

†† *Permanent address*: Department of Mathematics, Oulu University, 90101 Oulu 10, Finland.

‡ *Current address*: Department of Mathematics, University of Southern California, Los Angeles, California 90007 USA.

Received April 27, 1978

zeros of accretive operators A that satisfy the convergence condition. The implicit scheme replaces the problem $0 \in Ax$ by a sequence of easier problems, each of which has to be solved only approximately. No boundedness assumptions on A are needed to establish convergence in this case. The result for the explicit method, on the other hand, is restricted to bounded operators (or more generally, to locally bounded operators in the presence of a priori information on the location of the zeros). In each case we also study the rate of convergence. Some of our results are new even in Hilbert space and may be applied, for example, to convex programming (cf. [18]).

In Hilbert space there are two notable cases where strong convergence occurs even though the convergence condition need not be satisfied. In the first case the semigroup is generated by certain odd operators, and in the second it has a fixed point set with a nonempty interior. The first case was partially extended to Banach spaces in [1, corollary 4.1]. In Section 6 we assume that the interior of $A^{-1}0$ is nonempty and extend the second case outside Hilbert space. In Hilbert space the result is due to H. Brezis [2] (see also [13]).

In Section 7 we indicate possible extensions and generalizations of our results to other Banach spaces.

For previous results on the implicit and explicit schemes in Hilbert space see, for example, [4, 6, 12, 14, 18] and the references mentioned there. Banach space results can be found in [8, 17]. See also [7, 11, 15, 16].

Our terminology and notation follow that of [10].

2. The convergence condition

Let E^* be the dual of a real Banach space E, and denote the norm of both E and E^* by $|\cdot|$. For simplicity we shall assume in the sequel that both E and E^* are uniformly convex. In Section 7 we show how some of our results can be extended in some sense to an arbitrary Banach space.

In our setting, the duality map $J: E \to E^*$ (defined by $(x, Jx) = |x|^2$ and $|Jx| = |x|$) is single-valued and continuous. If C is a closed convex subset of E, then also the nearest point mapping $P: E \to C$ (defined by $|x - Px| = \inf\{|x - y|: y \in C\}$) is single-valued and continuous.

We recall that a possibly multivalued operator $A \subset E \times E$ with domain $D(A)$ and range $R(A)$ is accretive if for each $x_i \in D(A)$ and each $y_i \in Ax_i$, $i = 1, 2$, $(y_1 - y_2, J(x_1 - x_2)) \geqq 0$. It is m-accretive if, in addition, $R(I + rA) = E$ for all $r > 0$. We shall always assume that $0 \in R(A)$, so that $A^{-1}0$ is nonempty, closed and convex (see [3, theorem 1.2] and notice that $A^{-1}0$ is the fixed point set of the resolvent).

Let A be m-accretive and let $P: E \to A^{-1}0$ be the nearest point mapping. We shall say that A satisfies the *convergence condition* if $[x_n, y_n] \in A$, $|x_n| \leqq C$, $|y_n| \leqq C$, and $\lim_{n\to\infty}(y_n, J(x_n - Px_n)) = 0$ imply that $\liminf_{n\to\infty}|x_n - Px_n| = 0$.

Every strongly accretive A (i.e. an operator of the form $B + \alpha I$, where B is m-accretive and $\alpha > 0$) satisfies this condition, and Pazy's perturbation theorem [14, theorem 2.6] holds in our setting too. Since both J and P are continuous, the following proposition is also true (cf. [14, proposition 3.2]).

PROPOSITION 1. *Let A be m-accretive with $A^{-1}0 \neq \emptyset$. If $(y, J(x - Px)) > 0$ for every $[x, y] \in A$ with $x \notin A^{-1}0$, and the resolvent $(I + A)^{-1}$ is compact, then A satisfies the convergence condition.*

Using this result one can extend some results of Pazy [14] on nonlinear parabolic equations in L^2-spaces to L^p-spaces for $1 < p < \infty$.

One can prove that if an operator satisfies the convergence condition, then so does its Yosida approximation.

In some cases operators actually satisfy a stronger convergence condition. We shall say that an m-accretive operator satisfies the *uniform convergence condition of order* $\gamma > 0$, if for each $C > 0$ there exists a constant $\Gamma_C > 0$ such that if $[x, y] \in A$, $|x| \leqq C$ and $|y| \leqq C$, then

$$(y, J(x - Px)) \geqq \Gamma_C |x - Px|^{2\gamma}. \tag{2.1}$$

This condition will be used to obtain convergence rates.

It is clear that if B is m-accretive then $A = B + \alpha I$ satisfies the uniform convergence condition of order 1 for any $\alpha > 0$. Certain substitution operators also satisfy this condition. We also observe that if $A^{-1}0 = \{z\}$ and the origin is in the interior of Az, then A satisfies the uniform convergence condition of order $\frac{1}{2}$. In fact, let $\rho u \in Az$ for all $|u| = 1$ and some $\rho > 0$. Then $(y - \rho u, J(x - z)) \geqq 0$ for all $[x, y] \in A$. Taking $u = (x - z)/|x - z|$, we obtain $(y, J(x - z)) \geqq \rho|x - z|$.

3. The continuous case

Let A satisfy the convergence condition. Since A is m-accretive, $-A$ generates a semigroup S of nonlinear contractions on $\mathrm{cl}(D(A))$, the closure of the domain of A. If $x \in D(A)$, $S(t)x$ is a strong solution [10] of the initial value problem

$$\begin{aligned} &u'(t) + Au(t) \ni 0, \\ &u(0) = x. \end{aligned} \tag{3.1}$$

THEOREM 1. *Let* $-A$ *generate the semigroup* S *on* $\mathrm{cl}(D(A))$. *If* A *satisfies the convergence condition, then for each* $x \in \mathrm{cl}(D(A))$, $S(t)x$ *converges strongly as* $t \to \infty$ *to a zero of* A.

PROOF. We may restrict our attention to x in $D(A)$. Denote $S(t)x$ by $u(t)$, $-u'(t)$ by $v(t)$, and $J(u(t)-Pu(t))$ by $j(t)$. For almost all $t>0$ we have

$$(v(t),j(t)) = \frac{1}{h}(u(t-h)-u(t),j(t)) + (\varepsilon(h,t),j(t)),$$

where $h>0$ and $\lim_{h\to 0}\varepsilon(h,t)=0$. Since

$$u(t-h)-u(t) = Pu(t)-u(t)+u(t-h)-Pu(t-h)+Pu(t-h)-Pu(t),$$

and $(y-Pu(t),j(t)) \leqq 0$ for all $y \in A^{-1}0$, we obtain

$$\begin{aligned}
&(u(t-h)-u(t),j(t)) \\
&\quad = -|u(t)-Pu(t)|^2 + (u(t-h)-Pu(t-h),j(t)) + (Pu(t-h)-Pu(t),j(t)) \\
&\quad \leqq -|u(t)-Pu(t)|^2 + \frac{1}{2}\{|u(t-h)-Pu(t-h)|^2 + |u(t)-Pu(t)|^2\} \\
&\quad = \frac{1}{2}\{|u(t-h)-Pu(t-h)|^2 - |u(t)-Pu(t)|^2\}.
\end{aligned}$$

The mapping $t \mapsto |u(t)-Pu(t)|$ is Lipschitzian. Consequently,

$$(v(t),j(t)) \leqq -\frac{1}{2}\frac{d}{dt}|u(t)-Pu(t)|^2 \qquad \text{for almost all } t>0.$$

Since A is accretive we have $(v(t),j(t)) \geqq 0$ and we conclude that $|u(t)-Pu(t)|$ is nonincreasing and $\liminf_{t\to\infty}(v(t),j(t))=0$. By the convergence condition the latter fact implies that $\liminf_{t\to\infty}|u(t)-Pu(t)|=0$. But since $|u(t)-Pu(t)|$ is nonincreasing we also have $\lim_{t\to\infty}|u(t)-Pu(t)|=0$. Finally notice that $|u(t)-p|$ is nonincreasing for any $p \in A^{-1}0$ and therefore we can write

$$\begin{aligned}
|u(t)-u(t+h)| &\leqq |u(t)-Pu(t)| + |Pu(t)-u(t+h)| \\
&\leqq 2|u(t)-Pu(t)|,
\end{aligned}$$

from which the convergence of $u(t)$ follows, and in particular, if $z=\lim_{t\to\infty}u(t)$,

$$|u(t)-z| \leqq 2|u(t)-Pu(t)|. \tag{3.2}$$

□

It is not difficult to see that if A satisfies the uniform convergence condition of order γ, then convergence occurs in finite time if $\gamma<1$, is exponential if $\gamma=1$, and is $O(t^{-1/2(\gamma-1)})$ if $\gamma>1$. More precisely, we have ($z=\lim_{t\to\infty}u(t)$)

$$(3.3)\quad |u(t)-z| \leqq 2|x-Px|\{1+|x-Px|^{2(\gamma-1)}2(\gamma-1)\Gamma t\}^{-1/2(\gamma-1)}, \qquad \gamma \neq 1$$

and

$$(3.4)\qquad |u(t)-z| \leqq 2|x-Px|e^{-\Gamma t}, \qquad \gamma = 1.$$

Here Γ denotes the constant Γ_C appearing in (2.1). In fact, using (2.1) we see from the proof of Theorem 1 that

$$(3.5)\qquad \frac{1}{2}\frac{d}{dt}|u(t)-Pu(t)|^2+\Gamma|u(t)-Pu(t)|^{2\gamma} \leqq 0,$$

for almost all $t>0$. This gives a bound for $|u(t)-Pu(t)|$ and (3.3) and (3.4) then follow from (3.2).

4. The implicit scheme

In this section we consider the implicit scheme defined by

$$(4.1)\qquad x_{n+1}+\lambda_{n+1}Ax_{n+1} \ni x_n + e_{n+1}, \qquad n \geqq 0$$

where $x_0 \in E$, and $\{\lambda_n\}$ is a positive sequence.

THEOREM 2. *Let A be m-accretive and let $\{x_n\}$ be defined by* (4.1). *If A satisfies the convergence condition, $\Sigma_{n=1}^{\infty}\lambda_n=\infty$, and $\Sigma_{n=1}^{\infty}|e_n|<\infty$, then $\{x_n\}$ converges strongly to a zero of A.*

PROOF. If the conclusion is true with $e_n \equiv 0$, then it is true if $\{e_n\}$ has a compact support. Approximating any $\{e_n\} \in l^1$ by a sequence with a compact support and using the facts that the resolvents J_{λ_n} are contractions and $A^{-1}0$ is closed, we see that we may assume in the remainder of the proof that $e_n \equiv 0$.

Let $j_n = J(x_n - Px_n)$, and denote $(x_n - x_{n+1})/\lambda_{n+1}$ by $y_{n+1} \in Ax_{n+1}$. We have

$$\begin{aligned}|x_{n+1}-Px_{n+1}|^2+\lambda_{n+1}(y_{n+1},j_{n+1}) &= (x_n - Px_{n+1}, j_{n+1})\\ &= (x_n - Px_n, j_{n+1}) + (Px_n - Px_{n+1}, j_{n+1})\\ &\leqq |x_n - Px_n|\,|x_{n+1}-Px_{n+1}|\\ &\leqq \frac{1}{2}\{|x_n-Px_n|^2+|x_{n+1}-Px_{n+1}|^2\}.\end{aligned}$$

Hence

$$(4.2)\qquad |x_{n+1}-Px_{n+1}|^2+2\lambda_{n+1}(y_{n+1},j_{n+1}) \leqq |x_n-Px_n|^2,$$

and $\Sigma_{n=1}^{\infty} \lambda_{n+1}(y_{n+1}, j_{n+1}) < \infty$. Since $\{\lambda_n\} \notin l^1$ and $(y_{n+1}, j_{n+1}) \geqq 0$, it follows that $\liminf_{n\to\infty}(y_n, j_n) = 0$. The sequence $\{x_n\}$ is bounded because $A^{-1}0 \neq \emptyset$, and $\{|y_n|\}$ is decreasing. Therefore we can invoke our convergence condition and conclude that $\liminf_{n\to\infty}|x_n - Px_n| = 0$. But $\{|x_n - Px_n|\}$ is decreasing, so that $x_n - Px_n \to 0$, and $\{x_n\}$ converges strongly because

$$|x_{n+m} - x_n| \leqq |x_{n+m} - Px_n| + |Px_n - x_n| \leqq 2|x_n - Px_n|.$$

Here we used the fact that $|x_n - p|$ is nonincreasing for all $p \in A^{-1}0$. □

To obtain convergence rates we assume now that A satisfies the uniform convergence condition, and let $e_n = 0$. We denote $(x_n - x_{n+1})/\lambda_{n+1}$ by $y_{n+1} \in Ax_{n+1}$ and $\lim_{n\to\infty} x_n$ by z.

When $\gamma = 1$ we obtain immediately from (4.2) that

$$|x_n - z| \leqq 2|x_0 - Px_0| \prod_{j=1}^{n} (1 + 2\Gamma_C\lambda_j)^{-1/2} \tag{4.3}$$

where $C = \max\{|x_0| + 2|Px_0|, |y_1|\}$. Note that this is essentially the rate obtained for the continuous case (3.4).

For $\gamma \neq 1$ we shall assume for simplicity that for some $\beta \geqq -1$,

$$\lambda_n = \lambda_0 n^\beta. \tag{4.4}$$

For $\gamma < 1$ we obtain from (4.2) that

$$|x_{n+1} - Px_{n+1}| \leqq (2\Gamma_C\lambda_0 n^\beta)^{-1/2\gamma}|x_n - Px_n|^{1/\gamma}. \tag{4.5}$$

Since $|x_{n+1} - Px_{n+1}| \leqq |x_{n+1} - z| \leqq 2|x_{n+1} - Px_{n+1}|$, we get the following rates:

$$\begin{aligned} |x_{n+1} - z| &= o(|x_n - z|^{1/\gamma}) && \text{if } \beta > 0, \\ &= O(|x_n - z|^{1/\gamma}) && \text{if } \beta \geqq 0, \\ &= o(|x_n - z|^{1/\gamma - \varepsilon}) && \text{for } \varepsilon > 0 \text{ if } \beta \geqq -1. \end{aligned} \tag{4.6}$$

While the continuous semigroup converges for $\gamma < 1$ in finite time (3.3), the implicit scheme does not generally converge in a finite number of steps. This, in fact, is impossible if $x_0 \neq Px_0$ and A is single-valued at $A^{-1}0$. However, the following is true:

PROPOSITION 2. *Assume that* $A^{-1}0 = \{z\}$ *and* $0 \in \operatorname{int}(Az)$. *If* $e_n \equiv 0$ *and* $\{x_n\}$ *satisfies* (4.1) *with* $(x_n - x_{n+1})/\lambda_{n+1} \to 0$, *then there exists* N *such that* $x_n = z$ *for* $n \geqq N$.

The proof is obvious. A case where $(x_n - x_{n+1})/\lambda_{n+1}$ always tends to zero is the following: if the modulus of convexity of E satisfies $\delta(\varepsilon) \geqq k\varepsilon^s$ for some $s \geqq 2$ and $k > 0$, and $\{\lambda_n\} \not\in l^s$, then by [8, theorem 2.6]

$$|x_n - x_{n+1}|/\lambda_{n+1} \leqq M\left(\sum_{j=1}^{n+1} \lambda_j^s\right)^{-1/s}.$$

If $\gamma > 1$, (4.6) is of no interest and we proceed differently. For $\beta > -1$ we set $|x_n - Px_n|^2 = k_n n^{-\alpha}$ and determine the largest α for which $\{k_n\}$ remains bounded. From (4.2) we then have

$$k_{n+1}(n+1)^{-\alpha} + 2\Gamma_C \lambda_0 n^\beta k_{n+1}^\gamma (n+1)^{-\alpha\gamma} \leqq k_n n^{-\alpha}. \tag{4.7}$$

By an induction argument we see that $\{k_n\}$ is bounded if $\alpha \leqq (1+\beta)/(\gamma - 1)$, and we obtain the rate

$$|x_n - z| = O(n^{-(1+\beta)/2(\gamma-1)}), \qquad \gamma > 1, \quad \beta > -1. \tag{4.8}$$

For $\beta = -1$, we obtain similarly

$$|x_n - z| = O((\log n)^{-1/2(\gamma-1)}), \qquad \gamma > 1, \quad \beta = -1. \tag{4.9}$$

Note that the rates (4.8) and (4.9) coincide with the rate for the continuous problem.

By Theorem 2, at each step the equation $x_{n+1} + \lambda_{n+1} A x_{n+1} \ni x_n$ has to be solved only approximately. If λ_n is sufficiently small this can often be done by using some locally convergent iteration method (e.g. Newton's). The initial guess for x_{n+1} may be taken to be x_n.

When A is continuous and defined on E (or more generally when $I - A$ maps a closed convex subset of E back into itself), one can actually use the explicit method to produce both the step size sequence $\{\lambda_n\}$ and the sequence $\{x_n\}$ itself in such a way that (4.1) is satisfied with summable errors and $\{\lambda_n\} \not\in l^1$. This can be done by using the method of [11].

5. The explicit scheme

In this section we consider the explicit scheme defined by

$$x_{n+1} \in x_n - \lambda_n A x_n, \qquad n \geqq 0 \tag{5.1}$$

where $x_0 \in E$, and $\{\lambda_n\}$ is a positive sequence. We shall study the convergence of $\{x_n\}$ under the assumptions that $\{x_n\} \subset D(A)$ and that $\{(x_n - x_{n+1})/\lambda_n\}$ is bounded. This is always the case if $D(A) = E$ and $R(A)$ is bounded. It can also be guaranteed for suitable $\{\lambda_n\}$ if $A^{-1}0$ is contained in the interior of $D(A)$ and x_0 is close enough to $A^{-1}0$.

Since E^* is uniformly convex, there is [15, p. 89] a continuous nondecreasing function $b: [0,\infty)\to[0,\infty)$ such that $b(0)=0$, $b(ct)\leqq cb(t)$ for $c\geqq 1$, and

$$|x+y|^2\leqq |x|^2+2(y,Jx)+\max\{|x|,1\}|y|b(|y|) \tag{5.2}$$

for all x and y in E.

We shall assume that $\Sigma_{n=0}^{\infty}\lambda_n=\infty$ and $\Sigma_{n=0}^{\infty}\lambda_n b(\lambda_n)<\infty$. For any given continuous nondecreasing function $b(t)$ with $b(0)=0$ such sequences $\{\lambda_n\}$ always exist. In particular, if the modulus of convexity of E^* satisfies $\delta_{E^*}(\varepsilon)\geqq k\varepsilon^r$ for some $k>0$ and $r\geqq 2$, then $b(t)\leqq ct^{s-1}$ with $s=r/(r-1)$. Therefore if $E=L^p$, $1<p<\infty$, we can take any $\{\lambda_n\}\in l^s\backslash l^1$ with $s=p$ if $1<p\leqq 2$ and $s=2$ if $p\geqq 2$.

THEOREM 3. *Let A be m-accretive and let $\{\lambda_n\}$ be a positive sequence such that $\Sigma_{n=0}^{\infty}\lambda_n=\infty$ and $\Sigma_{n=0}^{\infty}\lambda_n b(\lambda_n)<\infty$. Assume that $\{x_n\}$ satisfies* (5.1) *and that $\{(x_n-x_{n+1})/\lambda_n\}$ is bounded. If A satisfies the convergence condition, then $\{x_n\}$ converges strongly to a zero of A.*

PROOF. Let $j_n=J(x_n-Px_n)$, and denote $(x_n-x_{n+1})/\lambda_n$ by $y_n\in Ax_n$. Using (5.2) we have

$$|x_{n+1}-Px_{n+1}|^2\leqq |x_{n+1}-Px_n|^2=|x_n-Px_n-\lambda_n y_n|^2$$
$$\leqq |x_n-Px_n|^2-2\lambda_n(y_n,j_n)+\max\{|x_n-Px_n|,1\}|y_n|\lambda_n b(\lambda_n|y_n|).$$

Since for $c\geqq 1$, $b(ct)\leqq cb(t)$, this yields, for some M,

$$|x_{n+1}-Px_{n+1}|^2\leqq |x_n-Px_n|^2-2\lambda_n(y_n,j_n)+M\max\{|x_n-Px_n|,1\}\lambda_n b(\lambda_n). \tag{5.3}$$

This inequality implies that $\{|x_n-Px_n|\}$ is bounded. In fact, if we set $d_n=\max(|x_n-Px_n|,1)$ then

$$d_{n+1}^2\leqq d_n^2+Md_n\lambda_n b(\lambda_n),$$

and consequently, since $d_{n+1}^2-d_n^2=(d_{n+1}-d_n)(d_{n+1}+d_n)$,

$$d_{n+1}\leqq d_n+M\lambda_n b(\lambda_n),$$

which implies the boundedness of d_n and hence of $|x_n-Px_n|$. But then (5.3) actually implies that

$$|x_n-Px_n|^2\leqq |x_k-Px_k|^2+\varepsilon_k,\qquad n>k \tag{5.4}$$

where $\varepsilon_k=C\,\Sigma_{j=k}^{\infty}\lambda_j b(\lambda_j)\to 0$ as $k\to\infty$. Hence $\lim_{n\to\infty}|x_n-Px_n|$ exists. Furthermore

$$2\sum_{i=0}^{\infty}\lambda_i(y_i,j_i)\leqq |x_0-Px_0|^2+C\sum_{i=0}^{\infty}\lambda_i b(\lambda_i). \tag{5.5}$$

Since $\Sigma_{n=0}^{\infty}\lambda_n = \infty$ and $(y_i, j_i) \geqq 0$ this implies that $\liminf_{n\to\infty}(y_n, j_n) = 0$. Now we can apply the convergence condition since, by assumption, $\{y_n\}$ is bounded and repeating the boundedness argument for $x_n - Px_n$ with a fixed $p \in A^{-1}0$ in place of Px_n we also have $\{x_n\}$ bounded. Thus $\liminf_{n\to\infty}|x_n - Px_n| = 0$, and, since $\lim_{n\to\infty}|x_n - Px_n|$ exists, $\lim_{n\to\infty}|x_n - Px_n| = 0$. Finally, we have for all $n > k$,

$$|x_n - Px_k|^2 \leqq |x_k - Px_k|^2 + C_1 \sum_{i=k}^{\infty} \lambda_i b(\lambda_i). \tag{5.6}$$

Consequently, $|x_n - x_k| \leqq |x_n - Px_k| + |Px_k - x_k| \leqq 2|x_k - Px_k| + \varepsilon_k$, is small when k is large and the proof is complete. □

In order to obtain convergence rates, we again assume that A satisfies the uniform convergence condition of order γ. We also assume that $b(t) \leqq ct^{s-1}$ for some $1 < s \leqq 2$, and that $\lambda_n = \lambda_0 n^{\beta}$ with

$$-1 \leqq \beta < -1/s, \tag{5.7}$$

so that $\Sigma_{n=0}^{\infty}\lambda_n = \infty$ and $\Sigma_{n=0}^{\infty}\lambda_n^s < \infty$. For the explicit scheme the rates depend on γ and β in a more complicated way than in the case of the implicit scheme. If the order γ is very small, then the rate of convergence is independent of γ and is determined by β and s. The rate is better if β is smaller. On the other hand, if γ is large the rate is independent of s and equals the rate for the implicit scheme. In this case, the larger β is, the better the rate is. In general, for any fixed γ and s, there exists a unique optimal choice $\beta = \beta^*$, defined by

$$\beta^* = \begin{cases} -1, & \text{if } \gamma \leqq 1, \\ \gamma\{(1-\gamma)s - 1\}^{-1}, & \text{if } \gamma \geqq 1. \end{cases} \tag{5.8}$$

We shall obtain the following convergence rates ($z = \lim_{n\to\infty} x_n$):

$$|x_n - z| = \begin{cases} O(n^{(1+\beta s)/2}), & \beta \in [\beta^*, -1/s), \\ O(n^{-(1+\beta)/2(\gamma-1)}), & \gamma > 1, \quad \beta \in (-1, \beta^*], \\ O((\log n)^{-1/2(\gamma-1)}), & \gamma > 1, \quad \beta = -1. \end{cases} \tag{5.9}$$

(For $\gamma = 1$, $\beta = \beta^* = -1$, choose $\lambda_0 \geqq 1/\Gamma_C$, where Γ_C appears in the definition of the uniform convergence condition.)

Since for $\gamma > 1$ and $\beta \leqq \beta^*$ we obtain the same rates that were obtained for the implicit scheme and the continuous problem, we do not expect to be able to improve the rate by using other sequences. We shall also show below that in Hilbert space with $\gamma = 1$, the rate $O(n^{-1/2})$ obtained with $\lambda_n = \lambda_0/n$ is in some sense optimal over all possible choices $\{\lambda_n\} \in l^2 \backslash l^1$. This shows that the result of Bruck [6] is, in some sense, the best possible. On the other hand, for $\gamma < 1$ the rate (5.9) may be pessimistic. In fact, if $A^{-1}0$ is a singleton, then the convergence is faster.

We shall now establish (5.9). From (5.3) we obtain

$$|x_{n+1} - Px_{n+1}|^2 \leqq |x_n - Px_n|^2 - 2\lambda_n(y_n, j_n) + C\lambda_n^s. \tag{5.10}$$

Let $\lambda_n = \lambda_0 n^\beta$ and $|x_n - Px_n|^2 = k_n n^{-\alpha}$. By (5.10),

$$k_{n+1}(n+1)^{-\alpha} \leqq k_n n^{-\alpha} - 2\Gamma\lambda_0 n^\beta k_n^\gamma n^{-\alpha\gamma} + Cn^{\beta s}. \tag{5.11}$$

Let $\alpha = -(s-1)\beta/\gamma$. Then

$$k_{n+1} \leqq \left(1 + \frac{1}{n}\right)^{-(s-1)\beta/\gamma} \{k_n - [2\Gamma\lambda_0 k_n^\gamma - C]n^\delta\} \tag{5.12}$$

with $\delta = \beta(1 - s + s\gamma)/\gamma$. If $\gamma > 1$ and $\beta \geqq \beta^*$ $(\delta \geqq -1)$, then $\{k_n\}$ must be bounded. The same is true if $\gamma \leqq 1$ and $\delta > -1$ which holds for all β if $\gamma < 1$, and for $\beta > -1$ if $\gamma = 1$. If $\gamma = 1$ and $\beta = -1$, then (5.12) yields $(s \leqq 2)$

$$k_{n+1} \leqq k_n + [k_n(1 - 2\Gamma\lambda_0) + C]n^{-1} + [C - 2\Gamma\lambda_0 k_n]n^{-2}. \tag{5.13}$$

Since $\Gamma\lambda_0 \geqq 1$, $\{k_n\}$ must again be bounded. For $\gamma > 1$ and $\beta < \beta^*$ one proceeds as in the implicit case. Hence we have

$$|x_n - Px_n| = \begin{cases} O(n^{(s-1)\beta/2\gamma}), & \beta \in [\beta^*, -1/s), \\ O(n^{-(1+\beta)/2(\gamma-1)}), & \gamma > 1, \quad \beta \in (-1, \beta^*], \\ O((\log n)^{-1/2(\gamma-1)}), & \gamma > 1, \quad \beta = -1. \end{cases} \tag{5.14}$$

To obtain rates for $|x_n - z|$, note that $|x_n - z| \leqq |x_n - Px_n| + |Px_n - z|$, and by (5.6),

$$|z - Px_n|^2 \leqq |x_n - Px_n|^2 + C\sum_{i=n}^{\infty} \lambda_i^s \leqq |x_n - Px_n|^2 + C_1 n^{1+\beta s}.$$

The rates (5.9) now follow from (5.14).

Let B be a maximal monotone operator in Hilbert space such that $0 \in B0$ and $|Bx| = 1$ for $x \neq 0$. Consider the explicit scheme for $A = I + B$. Clearly $A^{-1}0 = \{0\}$ and we have

$$|x_{n+1}|^2 \leqq (1 - \lambda_n)^2 |x_n|^2 + \lambda_n^2.$$

It is therefore natural to ask how fast the solutions of the equation

$$(5.15)\qquad |x_{n+1}|^2 = (1-\lambda_n)^2|x_n|^2 + \lambda_n^2$$

can decay when $\{\lambda_n\} \in l^2 \backslash l^1$.

Proposition 3. *If* $\{\lambda_n\} \in l^2\backslash l^1$ *and* $\{x_n\}$ *satisfies* (5.15) *with* $|x_n| \leqq Kn^{-q}$ *for all* n, *then* $q \leqq 1/2$.

Proof. If $\{\lambda_n\} \in l^2\backslash l^1$, then for every $\varepsilon > 0$ there are infinitely many k such that $\sum_{j=k}^{\infty} \lambda_j^2 \geqq 1/k^{1+\varepsilon}$ [9, lemma 1]. By (5.15)

$$|x_k|^2 \leqq 2\sum_{j=k}^{\infty} \lambda_j |x_j|^2 - \sum_{j=k}^{\infty} \lambda_j^2$$

$$\leqq \left(\sum_{j=k}^{\infty} \lambda_j^2\right)\left\{\frac{2\left(\sum_{j=k}^{\infty} |x_j|^4\right)^{1/2}}{\left(\sum_{j=k}^{\infty} \lambda_j^2\right)^{1/2}} - 1\right\}.$$

The term in the brackets must of course be positive. Using $|x_j| \leqq Kj^{-q}$ and $\sum_{j=k}^{\infty} \lambda_j^2 \geqq 1/k^{1+\varepsilon}$, we obtain $q \leqq 1/2 + \varepsilon/4$ and the result follows. □

6. Another sufficient condition for convergence

In this section we study another sufficient condition for strong convergence.

Let S be a contraction semigroup in E, and let F be its fixed point set. In [2], Brezis proved that in Hilbert space, if int (F), the interior of F, is nonempty, then for each x in the domain of S, $S(t)x$ converges strongly as $t \to \infty$ to a point in F. Using the idea of Pazy's proof [13] of this result we show here that the result can be extended outside Hilbert space. We also show that it is not true in all Banach spaces.

Theorem 4. *Let* E *be a Banach space and assume that both* E *and* E^* *are uniformly convex. Let* S *be a contraction semigroup in* E *with a fixed point set* F. *If* int $(F) \neq \emptyset$, *then for each* x *in the domain of* S, $S(t)x$ *converges strongly as* $t \to \infty$ *to a point in* F.

Proof. Let a ball with center x_0 and radius $r > 0$ be contained in F. Since, by (5.2), $2(y, Jx) + |x|^2 \leqq |x + y|^2 \leqq |x|^2 + 2(y, Jx) + \max(|x|, 1)|y|b(|y|)$ for all x and y in E, we obtain for $t \geqq s$, $0 < \rho \leqq r$, and $|u| = 1$,

$$\begin{aligned}
&-2\rho(u, J(S(t)x - x_0)) + |S(t)x - x_0|^2 \\
&\leqq |S(t)x - x_0 - \rho u|^2 \\
&\leqq |S(s)x - x_0 - \rho u|^2 \\
&\leqq |S(s)x - x_0|^2 - 2\rho(u, J(S(t)x - x_0)) + M\rho b(\rho).
\end{aligned}$$

Hence

$$2\rho(u, J(S(s)x - x_0) - J(S(t)x - x_0)) \leqq |S(s)x - x_0|^2 - |S(t)x - x_0|^2 + M\rho b(\rho).$$

Denote $J(S(t)x - x_0)$ by $y(t)$, and let $u = J^{-1}(y(s) - y(t))/|y(s) - y(t)|$. We obtain

$$|y(s) - y(t)| \leqq \{|S(s)x - x_0|^2 - |S(t)x - x_0|^2\}/2\rho + \frac{1}{2}Mb(\rho).$$

Given $\varepsilon > 0$, choose ρ such that $Mb(\rho) < \varepsilon$, and then choose s_0 such that

$$\{|S(s)x - x_0|^2 - |S(t)x - x_0|^2\}/\rho < \varepsilon \qquad \text{for all } t \geqq s \geqq s_0.$$

Then $|y(s) - y(t)| \leqq \varepsilon$, so that $\lim_{t\to\infty} y(t)$ exists. Since J^{-1} is continuous, $\lim_{t\to\infty} S(t)x$ also exists and it clearly belongs to F. □

If A is m-accretive and $\operatorname{int}(A^{-1}0) \neq \emptyset$, then a similar argument shows that the implicit scheme (4.1) will always converge. It is not difficult to see that if $\Sigma_{n=1}^{\infty} \lambda_n = \infty$, then the limit will belong to $A^{-1}0$. The same is true for the explicit scheme (5.1) if, in addition, $\Sigma_{n=0}^{\infty} \lambda_n b(\lambda_n) < \infty$ and the operator A is, for example, bounded.

Theorem 4 does not hold in all Banach spaces. To see this, consider the space C[0, 1] with the max norm. Let A be the operator

$$Au(x) = \max\{a(x)u(x), 0\}, \qquad u \in C[0,1], \quad 0 \leqq x \leqq 1,$$

where $a(x)$ is a continuous function satisfying $a(0) = 0$, $a(x) > 0$ for $x > 0$. A is m-accretive and $A^{-1}0 = \{f \in C[0,1]: f(x) \leqq 0 \text{ for } 0 \leqq x \leqq 1\}$. Let S be the semigroup generated by $-A$. For $u_0 \equiv 1$, $S(t)u_0(x) = e^{-a(x)t}$ which does not converge in $C[0,1]$.

7. Extensions and refinements

In this section we show that under certain assumptions that always hold if E and E^* are uniformly convex, versions of Theorem 1 and 2 hold in any Banach space. We also observe that the hypotheses of Theorem 3 and 4 can, in fact, be weakened.

Let E be a real Banach space with dual E^*. The duality mapping J from E into the family of nonempty weak star compact convex subsets of E^* is defined by

$$J(x) = \{x^* \in E^*: (x, x^*) = |x|^2 \quad \text{and} \quad |x^*| = |x|\}$$

for each x in E. Denote the distance between a point $x \in E$ and a set $V \subset E$ by

$d(x, V)$. A point $z \in V$ is said to be a best approximation to $x \in E$ if $|x - z| = d(x, V)$.

Recall that a set V in E is called a *sun* [19] if whenever $z \in V$ is a best approximation to $x \in E$, then z is also a best approximation to $z + t(x - z)$ for all $t \geqq 0$. Every convex set is a sun. If V is a sun and $z \in V$ is a best approximation to $x \in E$, then there exists $j \in J(x - z)$ such that $(y - z, j) \leqq 0$ for all $y \in V$. The set V is said to be *proximinal* if each x in E has at least one best approximation in V.

An operator $A \subset E \times E$ is said to be accretive in the sense of Browder [5] if for each $x_i \in D(A)$ and each $y_i \in Ax_i$, $i = 1, 2$, $(y_1 - y_2, j) \geqq 0$ for all $j \in J(x_1 - x_2)$. This stronger notion of accretivity is needed in the generalization of Theorem 2 but in Theorem 1 it is superfluous since we have to assume that the semigroup is differentiable.

Let $A \subset E \times E$ be an accretive operator in the sense of Browder (in certain cases only the fact that A is accretive is really needed) with $0 \in R(A)$, and assume that $A^{-1}0$ is a proximinal sun. If P is a selection of the nearest point mapping onto $A^{-1}0$, then for each $x \in E$, there is at least one $j \in J(x - Px)$ such that $(y - Px, j) \leqq 0$ for all y in $A^{-1}0$. We will denote this j by $J_p(x - Px)$.

In this general setting, A will be said to satisfy the *convergence condition* if there is a selection P of the nearest point mapping onto $A^{-1}0$ such that if $[x_n, y_n] \in A$, $|x_n| \leqq C$, $|y_n| \leqq C$, and $\lim_{n\to\infty}(y_n, J_P(x_n - Px_n)) = 0$, then $\liminf_{n\to\infty} |x_n - Px_n| = 0$.

With this extension of the convergence condition, Theorem 1 is true as stated if the semigroup S is differentiable. Theorem 2 is true as stated. In Theorem 3 one only needs to assume that $A^{-1}0$ is proximinal. Theorem 4 is true if E^* is uniformly convex with a Fréchet differentiable norm. A somewhat different argument shows that Theorem 4 is also true if S is differentiable and E is uniformly convex and smooth.

The following example may be of some interest in connection with the convergence condition in general Banach spaces. Let $E = C[0, 1]$, and let $a(x)$ be a nonnegative continuous function. Let $(Au)(x) = a(x)u(x)$ for $u \in E$ and let S be the semigroup generated by $-A$. It is clear that $S(t)u_0$ converges for all $u_0 \in E$ if and only if $a(x) \equiv 0$ or $a(x) > 0$ for $x \in [0, 1]$. We show that for these operators the convergence condition is not only a sufficient condition but also necessary. Assume therefore that $a(x) \geqq \alpha > 0$ for $x \in [0, 1]$. Then $A^{-1}0 = \{0\}$ and P is trivial so that in the convergence condition we can use any $j_n \in J(u_n)$. Let $x_n \in [0, 1]$ be a point such that $|u_n(x_n)| = |u_n|_\infty$, where $\{u_n\}$ is any bounded sequence in E. Assume then that

$$\lim_{n\to\infty} (Au_n, j_n) = 0.$$

But $(Au_n, j_n) = a(x_n)|u_n|_\infty^2$ and since $a(x) \geqq \alpha > 0$ we conclude $|u_n|_\infty^2 \to 0$, i.e. A satisfies the convergence condition. If on the other hand $a(x) \equiv 0$ then A satisfies trivially the condition, and hence in this case the strong convergence of the trajectories of the semigroup is equivalent to the convergence condition on its generator. This is in contrast with the situation for similar operators in Hilbert space (cf. [14], or consider A in l^2 given by $\xi_n \mapsto \alpha_n \xi_n$ where $\alpha_n \searrow 0$).

Finally we remark that if A does not satisfy the convergence condition, the problem $0 \in Ax$ can still sometimes be solved iteratively by approximating A by $A + p_n I$ and letting $p_n \searrow 0$ slowly during the iteration. See, e.g. [7, 15, 16].

Acknowledgment

We wish to thank Professors M. G. Crandall and A. Pazy for many helpful conversations.

Note added in proof

(1) An equivalent definition of the convergence condition is obtained when lim inf is replaced by lim.

(2) In connection with Section 6, see also the papers by J. J. Moreau, *Un cas de convergence des itérées d'une contraction d'un espace hilbertien*, C. R. Acad. Sci. Paris **286** (1978), 143–144, and by B. Beauzamy, *Un cas de convergence forte des itérés d'une contraction dans un espace uniformément convexe*, to appear.

(3) In Section 7, E may be a general Banach space, but V and $A^{-1}0$ must be convex. Theorems 1 and 2 hold if E is, for example, reflexive, strictly convex, and smooth.

References

1. J. B. Baillon, R. E. Bruck and S. Reich, *On the asymptotic behavior of nonexpansive mappings and semigroups in Banach spaces*, Houston J. Math., to appear.

2. H. Brezis, *Monotonicity methods in Hilbert space and some applications to nonlinear partial differential equations*, in *Contributions to Nonlinear Functional Analysis* (E. H. Zarantonello, ed.), Academic Press, New York, 1971, pp. 101–156.

3. H. Brezis, *Opérateurs Maximaux Monotones et Semigroupes de Contractions dans les Espaces de Hilbert*, North-Holland Publishing Co., Amsterdam, 1973.

4. H. Brezis and P. L. Lions, *Produits infinis de résolvantes*, Israel J. Math. **29** (1978), 329–345.

5. F. E. Browder, *Nonlinear accretive operators in Banach spaces*, Bull. Amer. Math. Soc. **73** (1967), 470–476.

6. R. E. Bruck, *The iterative solution of the equation $y \in x + Tx$ for a monotone operator T in Hilbert space*, Bull. Amer. Math. Soc. **79** (1973), 1258–1261.

7. R. E. Bruck, *A strongly convergent iterative solution of the equation* $0 \in U(x)$ *for a maximal monotone operator U in Hilbert space*, J. Math. Anal. Appl. **48** (1974), 114–126.

8. R. E. Bruck and S. Reich, *Nonexpansive projections and resolvents of accretive operators in Banach spaces*, Houston J. Math. **3** (1977), 459–470.

9. M. D. Canon and C. D. Cullum, *A tight upper bound on the rate of convergence of the Frank-Wolfe algorithm*, SIAM J. Control **6** (1968), 509–516.

10. M. G. Crandall and T. M. Liggett, *Generation of semi-groups of nonlinear transformations on general Banach spaces*, Amer. J. Math. **93** (1971), 265–298.

11. M. G. Crandall and A. Pazy, *On the range of accretive operators*, Israel J. Math. **27** (1977), 235–246.

12. J. C. Dunn, *Iterative construction of fixed points for multivalued operators of the monotone type*, J. Functional Analysis **27** (1978), 38–50.

13. A. Pazy, *On the asymptotic behavior of semigroups of nonlinear contractions in Hilbert space*, J. Functional Analysis **27** (1978), 292–307.

14. A. Pazy, *Strong convergence of semigroups of nonlinear contractions in Hilbert space*, MRC Report #1828, 1978.

15. S. Reich, *An iterative procedure for constructing zeros of accretive sets in Banach spaces*, Nonlinear Analysis **2** (1978), 85–92.

16. S. Reich, *Iterative methods for accretive sets*, Proc. Conf. on Nonlinear Equations, Academic Press, to appear.

17. S. Reich, *Weak convergence theorems for nonexpansive mappings in Banach spaces*, J. Math. Anal. Appl., to appear.

18. R. T. Rockafellar, *Monotone operators and the proximal point algorithm*, SIAM J. Control and Optimization **14** (1976), 877–898.

19. I. Singer, *Best Approximation in Normed Linear Spaces*, Springer, Berlin, 1970.

Mathematics Research Center
University of Wisconsin
Madison, Wisconsin 53706 USA