New results concerning monotone operators

and nonlinear semigroups

Haïm BREZIS

Our purpose is to describe here some recent developments in three different directions.

In §I we discuss a property of the range $R(A+B)$ of the sum of two monotone operators. Surprisingly, it turns out that in "many" cases $R(A+B)$ is "almost" equal to $R(A)+R(B)$. A number of applications to nonlinear partial differential equations are given.

In §II we prove some estimates showing that $(I+tA)^{-1}$ and $S(t)$ have the same modulus of continuity at $t = 0$ ($S(t)$ denotes the semigroup generated by $-A$). Next we present some consequences.

In §III we give a very general form of the convergence theorem of Trotter - Kato - Neveu type for nonlinear semigroups.

§I "$R(A+B) \simeq R(A)+R(B)$" and applications

Let $H$ be a real Hilbert space and let $A$ and $B$ be maximal monotone operators such that $A+B$ is again maximal monotone.

We say that two subsets $K_1$ and $K_2$ of $H$ are almost equal ($K_1 \simeq K_2$) if $K_1$ and $K_2$ have the same closure and the same interior. We prove here, under various assumptions, that
\( R(A+B) \simeq R(A) + R(B) \); we discuss here only the simplest forms (for more elaborate results see [7]).

**Theorem 1** Suppose \( A \) and \( B \) are subdifferentials of convex functions. Then \( R(A+B) \simeq R(A) + R(B) \).

**Proof** First we prove that \( \overline{R(A+B)} = \overline{R(A) + R(B)} \); it is sufficient to verify that \( R(A) + R(B) \subset \overline{R(A+B)} \). Given \( f \in \overline{R(A) + R(B)} \), there exist \( \xi \in D(A) \) and \( \eta \in D(B) \) such that \( f \in A\xi + B\eta \). The equation

\[
(1) \quad \varepsilon u_\varepsilon + Au_\varepsilon + Bu_\varepsilon \ni f
\]

has a unique solution \( u_\varepsilon \). The conclusion follows provided we show that \( \varepsilon u_\varepsilon \to 0 \) as \( \varepsilon \to 0 \). Let \( x \in D(A) \cap D(B) \) be fixed. Since \( A \) and \( B \) are cyclically monotone (see [21]) we have

\[
(2) \quad (Au_\varepsilon, u_\varepsilon - x) + (Ax, x - \xi) + (A\xi, \xi - u_\varepsilon) \geq 0
\]

\[
(3) \quad (Bu_\varepsilon, u_\varepsilon - x) + (Bx, x - \eta) + (B\eta, \eta - u_\varepsilon) \geq 0
\]

and therefore by adding (2) and (3) we obtain

\[
(f - \varepsilon u_\varepsilon, u_\varepsilon - x) + C - (f, u_\varepsilon) \geq 0,
\]

where \( C \) is independent of \( \varepsilon \). Hence

\[
\varepsilon |u_\varepsilon|^2 - \varepsilon(u_\varepsilon, x) \leq C
\]

and therefore \( \sqrt{\varepsilon}|u_\varepsilon| \) remains bounded as \( \varepsilon \to 0 \).

Next we prove that \( \text{Int}[R(A)+R(B)] = \text{Int}[R(A+B)] \). It is sufficient to check that \( \text{Int}[R(A)+R(B)] \subset R(A+B) \). Let \( f \in \text{Int}[R(A)+R(B)] \), so that a ball \( B(f, \rho) \) is contained in \( R(A)+R(B) \). For every \( h \in H \) with \( |h| < \rho \), there exist \( \xi \)
and \( \eta \) (depending on \( h \)) such that \( f + h \in A\xi + B\eta \). Going back to (2) and (3) and adding them we obtain now
\[
(f - \varepsilon u_\varepsilon, u_\varepsilon - x) + C(h) - (f + h, u_\varepsilon) \geq 0
\]
where \( C(h) \) depends on \( h \), but is independent of \( \varepsilon \).
Hence \( (h, u_\varepsilon) \leq C(h) \) for every \( h \in H \) with \( |h| < \rho \). It follows from the uniform boundedness principle that \( \{ u_\varepsilon \} \) remains bounded as \( \varepsilon \to 0 \). Passing to the limit in (1) we conclude by standard methods that \( f \in R(A+B) \).

**Theorem 2** We suppose now that only \( A \) is the subdifferential of a convex function, but \( D(B) \subset D(A) \). Then \( R(A+B) \approx R(A) + R(B) \).

**Proof** We proceed as in the proof of Theorem 1.
First let \( f \in R(A+B) \) i.e. \( f \in A\xi + B\eta \); let \( u_\varepsilon \) be the solution of (1). We have
\[
(Au_\varepsilon, u_\varepsilon - \eta) + (A\eta, \eta - \xi) + (A\xi, \xi - u_\varepsilon) \geq 0
\]
(5) \[
(Bu_\varepsilon, u_\varepsilon - \eta) + (B\eta, \eta - u_\varepsilon) \geq 0.
\]
By adding (4) and (5) we obtain
\[
(f - \varepsilon u_\varepsilon, u_\varepsilon - \eta) + C - (f, u_\varepsilon) \geq 0
\]
and hence
\[
\varepsilon |u_\varepsilon|^2 - \varepsilon (u_\varepsilon, \eta) \leq C'.
\]
Next suppose \( f \in \text{Int}[R(A)+R(B)] \); we obtain now, as in the proof of Theorem 1
\[
(f - \varepsilon u_\varepsilon, u_\varepsilon - \eta) + C(h) - (f + h, u_\varepsilon) \geq 0
\]
i.e. \( (h, u_\varepsilon) \leq C'(h) \).

**Theorem 3** Suppose \( A \) is a subdifferential of a convex
function $\varphi$ and let $B$ be a maximal monotone operator such that

$$\varphi((I + \lambda B)^{-1}x) \leq \varphi(x) \quad \forall \lambda > 0, \forall x \in D(\varphi).$$

Then $R(A + B) \simeq R(A) + R(B)$.

**Remark** We know (see [4]) that (6) implies that $A + B$ is maximal monotone.

**Proof** Let $f \in R(A) + R(B)$ and let $u_\varepsilon$ be the solution of (1). It follows easily from (6) that $\varepsilon |u_\varepsilon|$, $|Au_\varepsilon|$ and $|Bu_\varepsilon|$ remain bounded as $\varepsilon \to 0$. Next we have

$$\begin{align*}
(Au_\varepsilon - A\bar{\xi}, u_\varepsilon - \bar{\xi}) & \geq 0 \\
(Bu_\varepsilon - B\eta, u_\varepsilon - \eta) & \geq 0.
\end{align*}$$

Hence, by adding (7) and (8) we obtain

$$\langle f - \varepsilon u_\varepsilon, u_\varepsilon \rangle - \langle f, u_\varepsilon \rangle + C \geq 0$$

i.e. $\varepsilon |u_\varepsilon|^2 \leq C$. Suppose now that $f \in \text{Int}[R(A) + R(B)]$, with the same argument as above we have

$$\langle f - \varepsilon u_\varepsilon, u_\varepsilon \rangle - \langle f + h, u_\varepsilon \rangle + C(h) \geq 0$$

i.e. $(h, u_\varepsilon) \leq C(h)$ for $|h| < \rho$.

**Some applications**

Let $\Omega \subset \mathbb{R}^N$ be a bounded domain with smooth boundary $\partial \Omega$. Let $\beta : \mathbb{R} \to \mathbb{R}$ be a monotone nondecreasing continuous function such that $\beta(0) = 0$. Consider the equation (for a given $f \in L^2(\Omega))$:

$$-\Delta u + \beta(u) = f \quad \text{on } \Omega, \quad \frac{\partial u}{\partial n} = 0 \quad \text{on } \partial \Omega.$$

**Theorem 4** A necessary condition for the existence of a
solution of (9) is that \( \frac{1}{|\Omega|} \int_{\Omega} f(x) dx \in R(\beta) \). A sufficient condition is that \( \frac{1}{|\Omega|} \int_{\Omega} f(x) dx \in \text{Int } R(\beta) \).

**Proof** The necessary condition is clear by integrating (9) on \( \Omega \). In order to prove the sufficient condition we apply Theorem 1 in \( H = L^2(\Omega) \) with

\[
A = -\Delta, \quad D(A) = \left\{ u \in H^2(\Omega); \quad \frac{\partial u}{\partial n} = 0 \quad \text{on } \partial \Omega \right\}
\]

\[
B = \beta, \quad D(B) = \left\{ u \in L^2(\Omega); \quad \beta(u) \in L^2(\Omega) \right\}.
\]

Both \( A \) and \( B \) are subdifferentials of convex functions; also \( A+B \) is maximal monotone. It is well known that \( R(A) = \left\{ f \in L^2(\Omega); \quad \int_{\Omega} f(x) dx = 0 \right\} \). Finally if \( \frac{1}{|\Omega|} \int_{\Omega} f(x) dx \in \text{Int } R(\beta) \), then \( f \in \text{Int}[R(A)+R(B)] \). Indeed for \( g \in L^2(\Omega) \) we have

\[
g = (g - \frac{1}{|\Omega|} \int_{\Omega} g(x) dx) + \frac{1}{|\Omega|} \int_{\Omega} g(x) dx.
\]

And so it is clear that \( g \in R(A)+R(B) \) as soon as

\[
\frac{1}{|\Omega|} \int_{\Omega} g(x) dx - \frac{1}{|\Omega|} \int_{\Omega} f(x) dx \leq |\Omega|^{-\frac{1}{2}} \| f - g \|_{L^2} \quad \text{is small enough.}
\]

**Remark** Theorem 4 is related to a number of results of Schatzman [22], Hess [13], Landesman - Lazer [17], Nirenberg [19] etc...

The method used in the proofs of Theorems 1 - 3 can be easily extended to include most results known about "semi coercive" problems.

Let \( \mathcal{H} \) be a Hilbert space and let \( \varphi \) be a convex function on \( \mathcal{H} \). Given \( f \in L^2(0, T; \mathcal{H}) \) consider the equation
\begin{align}
(10) \quad \frac{du}{dt} + \partial \varphi(u) \in f \quad \text{on} \quad (0, T), \quad u(0) = u(T).
\end{align}

**Theorem 5**  A necessary condition for the existence of a solution of (10) is that \( \frac{1}{T} \int_0^T f(t) dt \in \overline{R(\partial \varphi)} \). A sufficient condition is that \( \frac{1}{T} \int_0^T f(t) dt \in \text{Int } R(\partial \varphi) \).

**Proof** Since \( \overline{R(\partial \varphi)} \) is convex, the necessary condition follows from the integration of (10). For the sufficient condition we apply Theorem 3 in \( H = L^2(0, T; \mathcal{H}) \) with \( A = \partial \varphi \) i.e. \( f \in Au \) provided \( f, u \in H \) and \( f(t) \in \partial \varphi(u(t)) \) a.e. and with \( B = \frac{d}{dt} \),

\( D(B) = \{ u \in H, \quad \frac{du}{dt} \in H \text{ and } u(0) = u(T) \} \). It is well known that \( A \) is a subdifferential of a convex function in \( H \), that \( B \) is maximal monotone and that (6) holds. The assumption

\( \frac{1}{T} \int_0^T f(t) dt \in \text{Int } \overline{R(\partial \varphi)} \) implies that \( f \in \text{Int}[R(A) + R(B)] \).

Indeed, note that \( R(B) = \{ f \in H; \int_0^T f(t) dt = 0 \} \). For \( g \in H \) we can write

\( g = (g - \frac{1}{T} \int_0^T g(t) dt) + \frac{1}{T} \int_0^T g(t) dt \in R(A) + R(B) \)

provided \( \| g-f \|_H \) is small enough.

**Theorem 6**  Let \( H \) be a Hilbert space and let \( K \) be a maximal monotone operator in \( H \) with \( D(K) = H \). Let \( F \) be the subdifferential of a convex function on \( H \) with \( D(F) = H \). Then \( R(I +KF) = H \).

**Proof**  Given \( f \in H \) we want to solve \( u + KFu = f \) i.e.
We apply Theorem 2 with \( A = F \) and \( B = -K^{-1}(f-u) \) so that \( B \) is maximal monotone; it follows that \( R(A+B) = R(A)+R(B) \). However \( R(B) = -D(K) = H \) and therefore \( R(A+B) = H \).

**Remark** Results related to Theorem 6 were obtained in [6].

§ II.1 Comparative behavior of \((I+tA)^{-1}\) and \(S(t)\) near \( t = 0 \)

1. **The Hilbert space case**

Suppose \( H \) is a Hilbert space and let \( A \) be a maximal monotone operator; let \( S(t) \) be the semigroup generated by \(-A\) in the sense of Kato-Komura (see e.g. [23] or [4]).

For \( x \in \overline{D(A)} \) and \( y \in \overline{D(A)} \) we have

\[
| x - S(t)x | \leq 2|x - y| + |y - S(t)y| \leq 2|x - y| + t|A^o y|.
\]

Choosing \( y = J_{\lambda} x = (I + \lambda A)^{-1} x \) we get

\[
(11) \quad | x - S(t)x | \leq (2 + \frac{t}{\lambda}) | x - J_{\lambda} x |
\]

and in particular, for \( \lambda = t \), we obtain

\[
(12) \quad | x - S(t)x | \leq 3|x - J_t x |.
\]

In case \( A = D \) we can show (see [5]) that

\[
(13) \quad | x - J_t x | \leq (1 + \frac{1}{\sqrt{2}}) | x - S(t)x |
\]

(the best constants are not known).

For general monotone operators an inequality of the kind (13) does not hold (consider for example in \( H = \mathbb{R}^2 \), \( A \) = a rotation...
by $\pi/2)$. However one can obtain a "substitute" for (13) in the general case as follows:

**Theorem 7** Let $A$ be a general maximal monotone operator; then we have

\[ |x - J_t x| \leq \frac{2}{t} \int_0^t |x - S(\tau)x| d\tau, \quad \forall x \in D(A), \forall t > 0. \quad (14) \]

**Remark** It is clear that the constant 2 in (14) can not be improved. Otherwise we would have for $x \in D(A)$, $|x - J_t x| \leq \frac{C}{t} \int_0^t |A^o x| d\tau = \frac{C}{2} |A^o x| t$ and as $t \to 0$, $|A^o x| \leq \frac{C}{2} |A^o x|$ with $C < 2$.

**Proof** Clearly, it is sufficient to prove (14) for $x \in D(A)$. Let $u(t) = S(t)x$; by the monotonicity of $A$, we have for $v \in D(A)$

\[ (Av + \frac{du}{dt}(t), v - u(t)) \geq 0. \quad (15) \]

Integrating (15) on $[0, t)$ we obtain

\[ \frac{1}{2} |u(t) - v|^2 - \frac{1}{2} |x - v|^2 \leq \int_0^t (Av, v - u(\tau)) d\tau = t(Av, v - x) + \int_0^t (Av, x - u(\tau)) d\tau. \quad (16) \]

Thus

\[ \frac{1}{2} |u(t) - v|^2 - \frac{1}{2} |x - v|^2 \leq t(Av, v - x) + |Av| \int_0^t |x - u(\tau)| d\tau. \]

Choosing $v = J_t x$ we get

\[ \frac{1}{2} |u(t) - J_t x|^2 - \frac{1}{2} |x - J_t x|^2 \leq -|x - J_t x|^2 + \frac{|x - J_t x|}{t} \int_0^t |x - u(\tau)| d\tau, \]

and (14) follows.
Remark Combining (12) and (14) we see that \(|x - J_\lambda x|\) and 
\(|x - S(t)x|\) have the same modulus of continuity at \(t = 0\).
Also, using Hardy's inequality we can deduce that for \(1 \geq \alpha > 0\) and \(1 \leq p \leq \infty\)
\[
\left\| \frac{x - J_\lambda x}{t^\alpha} \right\|_{L_p^*} \leq 3 \left\| \frac{x - S(t)x}{t^\alpha} \right\|_{L_p^*}
\]
and
\[
\left\| \frac{x - J_\lambda x}{t^\alpha} \right\|_{L_p^*} \leq \frac{2}{1+\alpha} \left\| \frac{x - S(t)x}{t^\alpha} \right\|_{L_p^*}
\]
where \(L_p^* = L^p([0, 1], H; \frac{dt}{t})\). These inequalities are useful
in the study of nonlinear interpolation classes (see [3]).

In a "similar spirit" we have the following

**Theorem 8** Let \(A\) be a general maximal monotone operator.
For \(x \in \overline{D(A)}\), \(\lambda > 0\) and \(t > 0\) we set
\[
y_{\lambda,t} = (I + \frac{\lambda}{t} (I - S(t)))^{-1} x.
\]
Then
\[
(17) \quad \left| y_{\lambda,t} - J_\lambda x \right|^2 \leq \left| x - J_\lambda x \right| \frac{2}{t} \int_0^t \left| x - S(\tau)x \right| d\tau.
\]

**Remark** Let \(\omega(t) = \text{Sup} \left| x - S(\tau)x \right|\). By a result of Kato
[14] (see also [4] Lemma 4.2) we know that for every integer \(n\)
\[
\left| y_{\lambda,t} - y_{\lambda,t/n} \right|^2 \leq 2 \omega(t) \left| y_{\lambda,t/n} - x \right|.
\]
Using the fact that \(y_{\lambda,s} \to J_\lambda x\) as \(s \to 0\) (see e.g. [4]
Proposition 4.1) we obtain as \(n \to \infty\)
\[
(18) \quad \left| y_{\lambda,t} - J_\lambda x \right|^2 \leq 2 \omega(t) \left| J_\lambda x - x \right|.
\]
Such an inequality follows also directly from (17).
Proof. We apply (16) with $x$ replaced by $y_{\lambda, t}$ and $v$ by $J_{\lambda} x$. Thus

$$
\frac{1}{2} |S(t)y_{\lambda, t} - J_{\lambda} x|^2 - \frac{1}{2} |y_{\lambda, t} - J_{\lambda} x|^2 \\
\leq \int_0^t \left( \frac{x - J_{\lambda} x}{\lambda}, J_{\lambda} x - S(\tau)y_{\lambda, t} \right) d\tau.
$$

However $S(t)y_{\lambda, t} = (1 + \frac{t}{\lambda})y_{\lambda, t} - \frac{t}{\lambda} x$ and so

$$
|S(t)y_{\lambda, t} - J_{\lambda} x|^2 \geq |y_{\lambda, t} - J_{\lambda} x|^2 + \frac{2t}{\lambda} (y_{\lambda, t} - J_{\lambda} x, y_{\lambda, t} - x).
$$

On the other hand

$$
(x - J_{\lambda} x, J_{\lambda} x - S(\tau)y_{\lambda, t}) = -|x - J_{\lambda} x|^2 + (x - J_{\lambda} x, x - S(\tau)y_{\lambda, t}) \\
\leq -|x - J_{\lambda} x|^2 + |x - J_{\lambda} x||x - S(\tau)x| + |x - y_{\lambda, t}|.
$$

We deduce from (19), (20) and (21) that

$$
\frac{t}{\lambda}(y_{\lambda, t} - J_{\lambda} x, y_{\lambda, t} - x) \leq \frac{t}{\lambda}|x - J_{\lambda} x|^2 + \frac{t}{\lambda}|x - J_{\lambda} x||x - y_{\lambda, t}| \\
+ \int_0^t \frac{|x - S(\tau)x|}{\lambda} \, d\tau.
$$

Therefore

$$
|x - J_{\lambda} x|^2 + (y_{\lambda, t} - J_{\lambda} x, y_{\lambda, t} - x) \leq |x - J_{\lambda} x||x - y_{\lambda, t}| \\
+ |x - J_{\lambda} x| \int_0^t \frac{|x - S(\tau)x|}{\lambda} \, d\tau.
$$

i.e. $|a|^2 + (b - a, b) \leq |a||b| + |x - J_{\lambda} x| \int_0^t |x - S(\tau)x| \, d\tau$

with $a = x - J_{\lambda} x$ and $b = x - y_{\lambda, t}$. Hence

$$
\frac{1}{2}|a-b|^2 = \frac{1}{2}|a|^2 + \frac{1}{2}|b|^2 - (a,b) \leq \\
- \frac{1}{2}|a|^2 - \frac{1}{2}|b|^2 + |a||b| + |x - J_{\lambda} x| \int_0^t |x - S(\tau)x| \, d\tau
$$

and $\frac{1}{2}|a-b|^2 \leq |x - J_{\lambda} x| \int_0^t |x - S(\tau)x| \, d\tau$. 

- 10 -
II.2 The Banach space case

Let $X$ be a general Banach space and let $A$ be an $m$-accretive operator on $X$. Let $S(t)$ be the semigroup generated by $-A$ in the sense of Crandall-Liggett (see [10] or [23]). Clearly we have as in §II.1

$$\|x - S(t)x\| \leq (2 + \frac{t}{\lambda}) \|x - J_\lambda x\|.$$ \hfill (22)

We don't know whether the exact analogue of (14) holds true. However we can prove the following

**Theorem 9** For every $x \in D(A)$, $t > 0$ and $\lambda > 0$ we have

$$\|x - J_\lambda x\| \leq (1 + \frac{\lambda}{t}) \int_0^t \|x - S(\tau)x\| d\tau \quad \hfill (23)$$

and in particular

$$\|x - J_t x\| \leq \frac{4}{t} \int_0^t \|x - S(\tau)x\| d\tau \quad \hfill (24)$$

**Proof** As usual we denote for $x, y \in X$

$$\tau(x, y) = \lim_{\lambda \to 0} \frac{1}{\lambda} (\|x + \lambda y\| - \|x\|) = \inf_{\lambda > 0} \frac{1}{\lambda} (\|x + \lambda y\| - \|x\|).$$

The analogue of (16) becomes now (see [10] or [2] for equivalent forms):

$$\|S(t)x - v\| - \|v - x\| \leq \int_0^t \tau(v - S(s)x, Av) ds \quad \hfill (25)$$

for every $v \in D(A)$.

However we have for every $\lambda > 0$

$$\tau(v - S(s)x, Av) \leq \frac{1}{\lambda} (\|v - S(s)x + \lambda Av\| - \|v - S(s)x\|). \quad \hfill (26)$$

If we choose in (26) $v = J_\lambda x$ we obtain
(27) \[ \tau (J_\lambda x - S(s)x, A_\lambda x) \leq \frac{1}{\lambda} (\|x - S(s)x\| - \|J_\lambda x - S(s)x\|) \]
and by (25) we get

(28) \[ \|S(t)x - J_\lambda x\| - \|J_\lambda x - x\| \leq \frac{1}{\lambda} \int_0^t (\|x - S(s)x\| - \|J_\lambda x - S(s)x\|) \, ds. \]

But \(-\|J_\lambda x - S(s)x\| \leq \|x - S(s)x\| - \|x - J_\lambda x\|\) and therefore (28) leads to

\[ -\|x - S(s)x\| \leq \frac{1}{\lambda} \int_0^t \|x - S(s)x\| \, ds + \frac{1}{\lambda} \int_0^t (\|x - S(s)x\| \| - \frac{t}{\lambda} \| x - J_\lambda x \| \]

i.e.

(29) \[ \|x - J_\lambda x\| \leq \frac{\lambda}{t} \|x - S(t)x\| + \frac{2}{t} \int_0^t \|x - S(s)x\| \, ds. \]

Finally note that

(30) \[ \|x - S(t)x\| \leq \frac{2}{t} \int_0^t \|x - S(s)x\| \, ds ; \]

indeed

\[ \|S(t)x - \frac{1}{t} \int_0^t S(s)x \, ds\| \leq \frac{1}{t} \int_0^t \|S(t)x - S(s)x\| \, ds \]

\[ \leq \frac{1}{t} \int_0^t \|S(t-s)x - x\| \, ds = \frac{1}{t} \int_0^t \|S(s)x - x\| \, ds , \]

and so

\[ \|x - S(t)x\| \leq \|x - \frac{1}{t} \int_0^t S(s)x \, ds\| + \frac{1}{t} \int_0^t \|S(s)x - x\| \, ds \leq \frac{2}{t} \int_0^t \|x - S(s)x\| \, ds . \]

Combining (29) and (30) we obtain (23).

Remarks:

1) I would like to thank Prof. M. Crandall, Y. Konishi and I. Miyadera for stimulating discussions concerning Theorem 9.

After our first result was obtained \( (\|x - J_\tau x\| \leq \frac{2}{t} \int_0^2 \|x - S(t)x\| \, dt) \),

I. Miyadera showed that \( \|x - J_\tau x\| \leq \frac{6}{t} \int_0^t \|x - S(\tau)x\| \, d\tau \) and
Y. Konishi got \( \|x - J_t x\| \leq \frac{4}{t} \int_0^t \|x - S(\tau)x\| d\tau \).

2) Using (22) and (23) one can prove directly the following result of M. Crandall [9]:

\[
\limsup_{t \to 0} \frac{\|x - S(t)x\|}{t} = \lim_{\lambda \to 0} \frac{\|x - J_\lambda x\|}{\lambda}.
\]

Indeed let \( \alpha = \limsup_{t \to 0} \frac{\|x - S(t)x\|}{t} \); and so \( \forall \varepsilon > 0 \exists \delta > 0 \)
such that \( 0 < t < \delta \)

\[
\|x - S(t)\| \leq t(\alpha + \varepsilon).
\]

From (23) we have for \( 0 < t < \delta \) and every \( \lambda > 0 \)

\[
\|x - J_\lambda x\| \leq (1 + \frac{\lambda}{t})^{\frac{2}{t}}(\alpha + \varepsilon) \int_0^t \tau d\tau = (\lambda + t)(\alpha + \varepsilon).
\]

It follows that \( \|x - J_\lambda x\| \leq \lambda(\alpha + \varepsilon) \) for every \( \lambda > 0 \) and

\( \varepsilon > 0 \). Next let \( \beta = \lim_{\lambda \to 0} \frac{\|x - J_\lambda x\|}{\lambda} \); and so \( \forall \varepsilon > 0 \exists \delta > 0 \)
such that for \( 0 < \lambda < \delta \)

\[
\|x - J_\lambda x\| \leq \lambda(\beta + \varepsilon).
\]

From (22) we get for \( 0 < \lambda < \delta \) and every \( t > 0 \)

\[
\|x - S(t)x\| \leq (2 + \frac{t}{\lambda}) \lambda(\beta + \varepsilon) = (t + 2\lambda)(\beta + \varepsilon).
\]

Hence \( \|x - S(t)x\| \leq t\beta \) for every \( t > 0 \).

3) In general for \( x \in D(A), \frac{\|x - S(t)x\|}{\|x - J_t x\|} \) does not necessarily converge to 1 as \( t \to 0 \).

Consider for example in \( H = \mathbb{R} \), \( Au = \frac{-1}{u} \) for \( u > 0 \) and \( Au = \emptyset \)
for \( u \leq 0 \). In this case \( J_0^t = \sqrt{t} \) and \( S_0^t = \sqrt{2t} \) (slightly more complicated examples were built previously by A. Plant and L. Veron).

4) In view of the example built by Crandall-Liggett in [11]
one can not expect to extend Theorem 8 to Banach spaces (or even to $\mathbb{R}^3$ with some Banach norm) since $y_{\lambda,t}$ does not necessarily converge to a limit as $t \to 0$.

II.3 **An application to the characterization of compact semi-groups.**

Let $A$ be an m-accretive operator in a general Banach space $X$ and let $S(t)$ be the semigroup generated by $-A$.

**Theorem 10.** The following properties are equivalent.

1. For every $t > 0$, $S(t)$ is compact i.e. $S(t)$ maps bounded sets of $\overline{D(A)}$ into compact sets of $X$.

2. For every $\lambda > 0$, $(I + \lambda A)^{-1}$ is compact i.e. maps bounded sets of $X$ into compact sets of $X$.

3. For every bounded set $B$ in $\overline{D(A)}$ and every $t_0 > 0$ the mappings $t \mapsto S(t)x$ are equicontinuous at $t = t_0$ as $x \in B$.

**Remarks**

1. Theorem 10 is due to A. Pazy [20] in the linear case and to Y. Konishi [15] in the nonlinear Hilbert case (his proof relies on a consequence of (18) and could not be extended to Banach spaces).

2. It is obvious that (32a) is equivalent to

   (32a') $(I + A)^{-1}$ is compact

   and also to

   (32a'') For every $M > 0$ the set
\{ x \in D(A); \| x \| \leq M \text{ and } \| y \| \leq M \text{ for some } y \in Ax \}\)

is relatively compact in \( X \).

**Proof** (31) \( \implies \) (32a)

Let \( \lambda \) be fixed and let \( x \in X \); we have for every \( t \geq 0 \)
\[ \| J_\lambda x - S(t) J_\lambda x \| \leq t \| A_\lambda x \| = \frac{t}{\lambda} \| x - J_\lambda x \|. \]

Let \( B \) be a bounded set in \( X \); given \( \varepsilon > 0 \), choose \( t_0 \) so small that
\[ \frac{t_0}{\lambda} \| x - J_\lambda x \| < \varepsilon /2 \quad \text{for } x \in B. \]

Since \( J_\lambda(B) \) is bounded in \( D(A) \), it follows from (31) that \( S(t_0)J_\lambda(B) \) is relatively compact. Thus \( S(t_0)J_\lambda(B) \) can be covered by a finite union \( \bigcup_i B(x_i, \varepsilon/2) \). Hence \( J_\lambda(B) \subseteq \bigcup_i B(x_i, \varepsilon) \)

and consequently \( J_\lambda(B) \) is precompact.

(31) \( \implies \) (32b)

Using (31) we have only to prove that the mappings \( t \mapsto S(t)x \)
are equicontinuous at \( t = \frac{t_0}{2} \) as \( x \in K, K \) compact

( \( K = S(\frac{t_0}{2})B \)). This follows directly from the fact that for each fixed \( x \), \( t \mapsto S(t)x \) is continuous and that \( x \mapsto S(t)x \)
is a contraction.

(32a) + (32b) \( \implies \) (31)

Fix a \( t_0 > 0 \) and let \( B \) be a bounded set in \( D(A) \). By (32b), for every \( \varepsilon > 0 \) there exists \( \delta > 0 \) such that
\[ \| S(t)x - S(t_0)x \| < \varepsilon \quad \text{for } |t - t_0| < \delta \text{ and } x \in B. \]

We deduce from (23) that for \( x \in B \) and \( \lambda > 0 \),
\[ \| S(t_0)x - J_\lambda S(t_0)x \| \leq (1 + \lambda t) \frac{2}{t} \int_0^t \| S(t)x - S(t + t_0)x \| \, dt \]
\[ \leq (1 + \frac{\lambda}{t}) 2 \varepsilon \quad \text{for every} \quad 0 < t \leq \delta. \]

In particular for \( 0 < \lambda \leq \delta \) and \( x \in B \) we have

\[ \| S(t_0)x - J_\lambda S(t_0)x \| \leq 4 \varepsilon. \]

Since \( J_\delta S(t_0)B \) is relatively compact it can be covered by a finite union \( \bigcup_i B(x_i, \varepsilon) \). Hence \( S(t_0)B \) can also be covered by a finite union of balls of radius \( 5 \varepsilon \) and thus \( S(t_0)B \) is precompact.

Remark Suppose \( H \) is a Hilbert space, \( \varphi \) is a convex function on \( H \) and let \( A = \partial \varphi \). In this case (31) is equivalent to (32a) since (32b) is satisfied automatically. Indeed we have

\[ |S(t)x - S(t_0)x| = |S(t - \frac{t_0}{2})y - S(\frac{t_0}{2})y| \leq |t - t_0| |A^\circ y| \]

where \( y = S(\frac{t_0}{2})x \). On the other hand (see e.g. [4] Théorème 3.2) we know that

\[ |A^\circ S(\frac{t_0}{2})x| \leq |A^\circ v| + \frac{2}{t_0} |x - v| \quad \text{for every} \quad v \in D(A). \]

Therefore the mappings \( t \mapsto S(t)x \) are equicontinuous at \( t = t_0 \) as \( x \) remains bounded.

In this case property (32a) is also equivalent to (32a'') For every \( M \) the set

\[ \{ x \in D(\varphi); \quad |x| \leq M \quad \text{and} \quad \varphi(x) \leq M \} \]

is relatively compact in \( H \).

Indeed (32a'') \( \implies \) (32a):

Let \( E = \{ x \in D(A); \quad |x| \leq M \quad \text{and} \quad |A^\circ x| \leq M \}; \) for a fixed \( v_0 \in D(\varphi) \) we have
\[ \varphi(v_0) - \varphi(x) \geq (A^\circ x, v_0 - x) \]

and so
\[ \varphi(x) \leq \varphi(v_0) + M|v_0| + M = M' \text{ when } x \in E. \]

Conversely \((32a) \implies (32a'''):\)

Let
\[ F = \{ x \in D(\varphi); \ |x| \leq M \text{ and } \varphi(x) \leq M \}; \]

for \( x \in F \) we have
\[ \varphi(x) - \varphi(J_{\lambda}x) \geq (A_{\lambda}x, x - J_{\lambda}x) = \frac{1}{\lambda} |x - J_{\lambda}x|^2. \]

Therefore, since \( \varphi \) is bounded below by some affine function, we get for \( x \in F, \)
\[ \frac{1}{\lambda} |x - J_{\lambda}x|^2 \leq M + C_1 |J_{\lambda}x| + C_2 \leq M + C_1 |x - J_{\lambda}x| + C_1 M + C_2. \]

Thus
\[ |x - J_{\lambda}x| \leq \sqrt{\lambda(C_3 \lambda + C_4)} \text{ for } x \in F. \]

Given \( \varepsilon > 0 \) we choose \( \lambda_0 > 0 \) so small that
\[ \sqrt{\lambda_0(C_3 \lambda_0 + C_4)} < \varepsilon. \]

Since \( J_{\lambda_0}(F) \) is relatively compact, it can be covered by a finite union \( \bigcup_i B(x_i, \varepsilon) \) and then \( F \subset \bigcup_i B(x_i, 2\varepsilon). \)

### III. A convergence theorem for nonlinear semigroups

Let \( H \) be a Hilbert space; let \( \{A_n\}_{n \geq 1} \) and \( A \) be maximal monotone operators. Let \( \{S_n(t)\}_{n \geq 1} \) and \( S(t) \) be the corresponding semigroups.

Our next result is a nonlinear version of the Theorem of Trotter-Kato-Neveu. A number of related results have been obtained previously by Miyadera-Oharu [18], Brezis-Pazy [8], Benilan [1], Goldstein [12], Kurtz [16] etc...

**Theorem 11.** The following properties are equivalent.
(33) \( \forall x \in D(A), \ \forall \lambda > 0 \ (I + \lambda A_n)^{-1}x \rightarrow (I + \lambda A)^{-1}x \)

(34) \( \forall x \in D(A) \ \exists x_n \in D(A_n) \text{ such that } x_n \rightarrow x \text{ and } A^o_n x_n \rightarrow A^o x \)

(35) \( \forall x \in D(A) \ \exists x_n \in D(A_n) \text{ such that } x_n \rightarrow x \text{ and } \forall t \geq 0 \ S_n(t)x_n \rightarrow S(t)x . \)

In addition the convergence in (33) (resp. (35)) is uniform for bounded \( \lambda \) (resp. bounded \( t \)).

The proof of Theorem 11 is divided into four parts

Part A \( (33) \Rightarrow (34) \)

Part B \( (34) \Rightarrow (33) \)

Part C \( (33) \Rightarrow (35) \)

Part D \( (35) \Rightarrow (33) \).

Part A \( (33) \Rightarrow (34) \)

Let \( x \in D(A) \); given \( \varepsilon > 0 \) there is a \( \lambda > 0 \) such that
\[
|x - (I + \lambda A)^{-1}x| < \varepsilon / 2
\]
\[
|A^o x - A^o x| < \varepsilon / 2 .
\]

Next, by (33) there is an integer \( N \) such that for \( n \geq N \)
\[
|(I + \lambda A_n)^{-1}x - (I + \lambda A)^{-1}x| < \varepsilon / 2
\]
\[
|(A_n^o x - A^o x| < \varepsilon / 2 .
\]

Combining these estimates we see that given \( \varepsilon > 0 \) there is an integer \( N(\varepsilon) \) and sequences \( u_n(\varepsilon) = (I + \lambda A_n)^{-1}x \) and \( f_n(\varepsilon) = (A_n^o x \) such that \( [u_n(\varepsilon), f_n(\varepsilon)] \in G(A_n) \) and for \( n \geq N(\varepsilon), \ |u_n(\varepsilon) - x| < \varepsilon , \ |f_n(\varepsilon) - A^o x| < \varepsilon . \) Let \( N_k = N(1/k) \);

we can always assume that \( N_k \) is increasing to \( \infty \).
We define the sequences $x_n$ and $g_n$ by $x_n = u_n(\frac{1}{k})$ and $g_n = f_n(\frac{1}{k})$ for $N_k \leq n < N_{k+1}$. Therefore $[x_n, g_n] \in G(A_n)$ and for $N_k \leq n < N_{k+1}$ we have $|x_n - x| < \frac{1}{k}$ and $|g_n - A^o x| < \frac{1}{k}$.

Consequently $x_n \to x$ and $g_n \to A^o x$; we are going to prove now that $A_n^o x_n \to A^o x$. Indeed $|A_n^o x_n| \leq |g_n|$ and thus for a subsequence we get $A_n^o x_n \to h$. Let $v \in D(A)$; by the monotonicity of $A_n$ we have
\[
((A_n)v - A_n^o x_n, (1 + \lambda A_n)^{-1}v - x_n) \geq 0.
\]

At the limit as $n_j \to \infty$ we obtain
\[
(A_n^o v - h, (I + \lambda A)^{-1}v - x) \geq 0.
\]

Next we pass to the limit as $\lambda \to 0$:
\[
(A^o v - h, v - x) \geq 0 \quad \forall v \in D(A).
\]

Therefore $h \in Ax$ (see e.g. [4] Proposition 2.7). Since on the other hand $|h| \leq |A^o x|$ we have $h = A^o x$. By the uniqueness of the limit, and the fact that $\limsup |A_n^o x_n| \leq |A^o x|$ we conclude that $A_n^o x_n \to A^o x$.

Part B \hspace{1em} (34) \Rightarrow (33)

Without loss of generality we may assume that $\lambda = 1$. Let $x \in D(A)$ and let $u_n = (I + A_n)^{-1}x$. Given $y \in D(A)$, let $y_n \in D(A_n)$ be the sequence given by (34) so that $y_n = (I + A_n)^{-1}(y + A_n^o y_n)$. Therefore $|u_n - y_n| \leq |x - y_n - A^o y_n|$ and thus $u_n$ is bounded. For a subsequence $u_{n_j} \to u$; by the monotonicity of $A_n$ we have
\[
(x - u_n - A_n^o y_n, u_n - y_n) \geq 0.
\]

Passing to the limit in (36) we obtain
(37) \((x - u - A^0 y, u - y) \geq 0 \forall y \in D(A)\).

In (37) we choose \(y = (I + \lambda A)^{-1} u\) and so

\[(x - u, u - J_{A^0} u) \geq \lambda (A^0 J_{A^0} u, A^0 u) \geq 0.\]

As \(\lambda \to 0\) we see that

\[(x - u, u - \text{Proj}_{D(A)} u) \geq 0.\]

On the other hand since \(x \in \overline{D(A)}\) we have

\[(\text{Proj}_{D(A)} u - x, u - \text{Proj}_{D(A)} u) \geq 0\]

and consequently \(u = \text{Proj}_{D(A)} u\) i.e. \(u \in \overline{D(A)}\). Going back to (37) we deduce now from [4] Proposition 2.7 that \(x - u \in A u\) i.e. \(u = (I + A)^{-1} x\). By the uniqueness of the limit we have in fact \(u_n \to (I + A)^{-1} x\).

It follows from (36) that for every \(y \in D(A)\)

\[
\lim \sup |u_n|^2 \leq (x, u - y) + (u, y) + (A^0 y, y - u).
\]

In particular if we take \(y = u\) we get

\[
\lim \sup |u_n|^2 \leq |u|^2 \quad \text{and thus} \quad u_n \to u.
\]

The convergence in (33) is uniform in \(\lambda\) as \(\lambda\) remains bounded:

Without loss of generality we may assume that \(x \in D(A)\) and let

\(x_n \in D(A_n)\) with \(x_n \to x\) and \(A_n^0 x_n \to A^0 x\). We have

\[
|(I + \lambda A_n)^{-1} x_n - (I + \lambda A_n)^{-1} x_n| \leq |\lambda - \mu| |A_n^0 x_n|.
\]

Therefore the functions \(f_n(\lambda) = (I + \lambda A_n)^{-1} x_n\) are uniformly lipschitz continuous on \([0, +\infty)\). Since they converge simply to \((I + \lambda A)^{-1} x\) as \(n \to +\infty\), we conclude that the convergence is uniform in \(\lambda\) as \(\lambda\) remains in a bounded interval.

**Part C** \((33) \Rightarrow (35)\)

Without loss of generality we may assume that \(x \in D(A)\). By (34)
we have a sequence \( x_n \in D(A_n) \) such that \( x_n \to x \) and \( A_n^\circ x_n \to A^\circ x \). We are going to prove that \( S_n(t)x_n \to S(t)x \). It is known (see e.g. [4] Corollaire 4.4) that

\[
|S_n(t)x_n - (I + \frac{t}{k}A)^{-k}x_n| \leq \frac{2k}{\sqrt{k}} |A_n^\circ x_n| \leq \frac{2tM}{\sqrt{k}}
\]

and

\[
|S(t)x - (I + \frac{t}{k}A)^{-k}x| \leq \frac{2k}{\sqrt{k}} |A^\circ x| \leq \frac{2tM}{\sqrt{k}}
\]

where \( M = \sup_{n} |A_n^\circ x_n| \). Given \( \varepsilon > 0 \), we first fix \( k \) large enough so that \( \frac{2tM}{\sqrt{k}} < \varepsilon \). Next observe, by induction, that for every integer \( N \) and for every sequence \( u_n \to u \) with \( u \in \overline{D(A)} \) then \( (I + \lambda A_n)^{-N}u_n \to (I + \lambda A_n)^{-N}u \), as \( n \to +\infty \). Thus

\[
|S_n(t)x_n - S(t)x| \leq 2\varepsilon + |(I + \frac{t}{k}A)^{-k}x_n - (I + \frac{t}{k}A)^{-k}x| \leq 3\varepsilon
\]

provided \( n \) is large enough.

Finally (35) holds true uniformly in \( t \) as \( t \) remains bounded since (33) holds true uniformly in \( \lambda \) as \( \lambda \) remains bounded.

**Part D** \( (35) \Rightarrow (33) \)

The proof relies on the following

**Lemma 1** Suppose (35) holds. Let \( f_n \in \overline{D(A_n)} \) be such that \( f_n \to f \) and \( f \in \overline{D(A)} \). Then \( \forall \lambda > 0, \forall t > 0 \)

\[
\begin{align*}
    u_n &= (I + \frac{\lambda}{t}(I - S_n(t)))^{-1}f_n \\
    &\to u = (I + \frac{\lambda}{t}(I - S(t)))^{-1}f.
\end{align*}
\]

**Proof of Lemma 1** By (35) there exists a sequence \( x_n \in \overline{D(A_n)} \)

such that \( x_n \to u \) and \( S_n(t)x_n \to S(t)u \). Writing the monotonicity of \( I - S_n(t) \) we have

\[
((u_n - S_n(t)u_n) - (x_n - S_n(t)x_n), u_n - x_n) \geq 0
\]
and therefore
\[
\left( \frac{u - u_n}{\lambda} + \delta_n, u_n - x_n \right) \geq 0
\]
where
\[
\delta_n = \frac{f_n - f}{\lambda} + \frac{u - x_n}{t} + \frac{S_n(t)x_n - S(t)u}{t}
\]
and \( \delta_n \to 0 \).

Hence
\[
\frac{1}{\lambda} |u_n - u|^2 \leq |\delta_n| |u_n - u| + |\delta_n| |u - x_n| + \frac{1}{\lambda} |u - u_n| |u - x_n|,
\]
and consequently \( u_n \to u \) as \( n \to \infty \).

**Lemma 2.** Let \( x_n \in D(A_n) \) be a sequence such that \( x_n \to x \) with \( x \in D(A) \) and \( S_n(t)x_n \to S(t)x \) for every \( t \geq 0 \). Then for every \( T \) there exists a constant \( K \) such that \( |(I + \lambda A_n)^{-1}x_n| \leq K \) and \( |S_n(t)x_n| \leq K \) for every \( 0 < \lambda < T \), for every \( 0 < t < T \) and every \( n \).

**Proof of Lemma 2.** Let \( M = \sup_{0 \leq t \leq 1} |S(t)x| \) and let
\[
E_n = \left\{ t \in [0, 1]; |S_p(t)x_p| \leq M + 1 \text{ for every } p \geq n \right\}.
\]
Clearly \( E_n \) is closed and \( \bigcup_{n=1}^\infty E_n = [0, 1] \); it follows from Baire's theorem that \( \text{Int } E_n \neq \emptyset \) for some \( N \). Let \( [t_0, t_0 + h] \subset E_n \) so that
\[
|S_p(t)x_p| \leq M + 1 \text{ for } n \geq N \text{ and } t_0 \leq t \leq t_0 + h.
\]
It follows from Theorem 9 that
\[
|S_n(t_0)x_n - (I + \lambda A_n)^{-1}S_n(t_0)x_n| \leq \frac{1 + \lambda}{h} \int_0^h |S_n(t_0)x_n - S_n(t_0 + \tau)x_n| \, d\tau.
\]
Choosing \( n \geq N \) we get
\[
|(I + \lambda A_n)^{-1}x_n| \leq |x_n - S_n(t_0)x_n| + |S_n(t_0)x_n| + \frac{2(1 + \lambda)}{h} \int_0^h |S_n(t_0)x_n - S_n(t_0 + \tau)x_n| \, d\tau.
\]
\[ |x_n| + 2(M+1) + 4(1 + \frac{A}{h})(M+1). \]

We conclude by using the fact that
\[ |x_n - S_n(t)x_n| \leq 3|x_n - (I + tA_n)^{-1}x_n|. \]

**Proof of (35) \implies (33)** In what follows \( \lambda \) is fixed. Using Theorem 8 we get
\[
| (I + \frac{\lambda}{t}(I - S_n(t)))^{-1}x_n - (I + \lambda A_n)^{-1}x_n |^2 \\
\leq |x_n - (I + \lambda A_n)^{-1}x_n| \cdot \frac{2}{t} \int_0^t |x_n - S_n(\tau)x_n| \, d\tau
\]

and
\[
| (I + \frac{\lambda}{t}(I - S(t)))^{-1}x - (I + \lambda A)^{-1}x |^2 \\
\leq |x - (I + \lambda A)^{-1}x| \cdot \frac{2}{t} \int_0^t |x - S(\tau)x| \, d\tau.
\]

Let \( P = 2|x - (I + \lambda A)^{-1}x| + 2 \sup_n |x_n - (I + \lambda A_n)^{-1}x_n| < \infty \) (by Lemma 2). We have
\[
\frac{1}{t} \int_0^t |x_n - S_n(\tau)x_n| \, d\tau \leq |x_n - x| + \frac{1}{t} \int_0^t |x - S(\tau)x| + \frac{1}{t} \int_0^t |S(\tau)x - S_n(\tau)x_n| \, d\tau
\]

and so
\[
| (I + \lambda A_n)^{-1}x_n - (I + \lambda A)^{-1}x | \leq |(I + \lambda \frac{A}{t}(I - S_n(t)))^{-1}x_n - (I + \lambda \frac{A}{t}(I - S(\tau)))^{-1}x | \\
+ \sqrt{P |x_n - x|} + 2 \sqrt{\frac{P}{t} \int_0^t |x - S(\tau)x| \, d\tau} + \sqrt{\frac{P}{t} \int_0^t |S(\tau)x - S_n(\tau)x_n| \, d\tau}
\]

\[ = X_1 + X_2 + X_3 + X_4. \]

Given \( \varepsilon > 0 \) we choose first \( t > 0 \) small enough so that \( X_3 < \varepsilon \)
and then we choose \( n \) large enough so that \( X_1 + X_3 + X_4 < \varepsilon \) (we use here Lemma 1 to make \( X_1 \) small and Lemma 2 combined with Lebesgue's Theorem to make \( X_2 \) small).
References


p.558-573.


Dept. de Mathématiques
Université de Paris VI
4 pl. Jussieu
75230 Paris 5e