

## STABILITY CONDITIONS FOR NONLINEAR PRODUCTS AND SEMIGROUPS

ERIC SCHECHTER

Let  $D$  be a subset of a Banach space. Suppose  $R(t): D \rightarrow D (t \geq 0)$  is an "almost-semigroup," in the sense that  $R(t)R(s)$  is close to  $R(t+s)$ . If  $R$  also satisfies certain stability conditions, then  $R(t/n)^n$  converges to some semigroup  $S(t)$  as  $n \rightarrow \infty$ . The stability conditions are motivated by several examples involving nonlinear partial differential equations.

1. Introduction. Let  $D$  be a subset of a Banach space. By a *transformation* on  $D$  we shall mean a one-parameter family of mappings

$$S(t): D \longrightarrow D \quad (t \geq 0)$$

(not necessarily linear or continuous), such that

$$S(0)f = f \quad \text{for all } f \in D.$$

$S$  is a *semigroup* if in addition

$$S(t+s)f = S(t)S(s)f \quad (t, s \geq 0; f \in D).$$

(This notation is essentially that of [4]. Numerous variants exist. For instance, in [3], such an  $S$  is referred to as a "semiflow"; the term "semigroup" is reserved for the linear case in that paper.)

In this paper and in [9], we shall consider the following problem: Let  $S_1$  and  $S_2$  be semigroups on  $D$ . Give sufficient conditions for the existence of a semigroup  $S$  on  $D$  such that

$$(1.1) \quad S(t)f = \lim_{n \rightarrow \infty} [S_1(t/n)S_2(t/n)]^n f \quad (t \geq 0, f \in D).$$

Equation (1.1) is commonly known as the *Trotter product formula*, because of Trotter's results for linear semigroups in [12].

The product problem generalizes as follows: Investigate the conditions under which a transformation  $R$  and a semigroup  $S$  satisfy

$$(1.2) \quad S(t)f = \lim_{n \rightarrow \infty} R(t/n)^n f \quad (t \geq 0, f \in D).$$

The Lax Equivalence Theorem of numerical analysis takes the form (1.2) with  $R$  and  $S$  linear; see [5]. The Crandall-Liggett exponential formula for nonlinear semigroups is of the form (1.2) with  $R(t)$  equal to  $(I + tA)^{-1}$ , the resolvent of an operator; see [4]. Equation

(1.2) was investigated in greater generality by Chernoff in [2].

In most applications of (1.1), the semigroups  $S_1$ ,  $S_2$  and  $S$  are generated (in some sense) by some operators  $-A_1$ ,  $-A_2$ , and  $-A_1 - A_2$ , respectively. Then the initial value problems for the differential equations

$$(1.3) \quad u'(t) + A_1 u(t) = 0$$

$$(1.4) \quad v'(t) + A_2 v(t) = 0$$

$$(1.5) \quad w'(t) + (A_1 + A_2)w(t) = 0$$

have solutions

$$(1.6) \quad u(t) = S_1(t)u(0), \quad v(t) = S_2(t)v(0), \quad w(t) = S(t)w(0).$$

Many equations arising in physical applications are of the form (1.5) where  $A_1 + A_2$  is the sum of two simpler operators, e.g., a linear operator and a continuous operator. Therefore it is desirable to reduce some questions about (1.5), e.g., existence and regularity of solutions, to questions about the simpler equations (1.3) and (1.4). This can often be accomplished using (1.1).

In [6] and in Theorem 5.3 of [3], the following result (roughly) is shown: Let  $R(t): D \rightarrow D$  be a transformation satisfying certain stability conditions. Suppose that

$$(1.7) \quad R'(0) = -A \text{ (in an appropriate sense), and}$$

$$(1.8) \quad R \text{ is an almost-semigroup, i.e., } R(t+s) \text{ is close to } R(t)R(s).$$

Then there exists a semigroup  $S$  such that  $S'(0) = -A$  and  $R(t/n)^n \rightarrow S(t)$ . In the present paper, we shall show that hypothesis (1.7) can be dropped, and (1.8) can be weakened slightly.

For practical purposes it is preferable to drop (1.8) and assume some variant of (1.7). In [6] it is shown that if  $S_1(t)$  and  $S_2(t)$  are sufficiently smooth semigroups satisfying

$$(1.7') \quad S_1'(0) + S_2'(0) = -A,$$

then  $R(t) = S_1(t)S_2(t)$  satisfies (1.7) and (1.8), and hence (1.2). In [9] we shall weaken the smoothness hypothesis on  $S_1$  and  $S_2$ .

This is the first of several papers dealing with the above problems and several closely related problems in one abstract framework. In the present paper, we shall deal with the simplest part of the theory, i.e., that part which does not involve generators.

In §§ 2-4 below, we give some examples to motivate our stability conditions. In § 5 we state the main results of this paper; the abstract framework given there will also be used in [9] and [10].

Proofs are given in § 6.

In [9] we shall introduce generators, and prove a version of (1.1) which can be used to show existence and regularity of solutions to (1.5). Also in [9], we shall give some additional historical background and applications.

The framework developed here has other uses as well. In [10] we shall investigate the relation between autonomous equations such as (1.3) – (1.6) and time-dependent equations such as

$$x'(t) + A(t, x(t)) = 0 .$$

In [11], similar techniques are used to prove local existence and regularity results and the optimality of some convergence rates.

Much of the recent research in nonlinear semigroups in Banach spaces has been done in the framework established by Crandall and Liggett in [4]. We shall not follow that framework. Our departure from it is discussed in § 2, below, and in [9] and [10].

Part of this research was conducted as part of a Ph. D. thesis at the University of Chicago. Also, part of the research was performed while the author was affiliated with the Applied Mathematics Divisions at Brookhaven and Argonne National Laboratories, supported by the U. S. Department of Energy. The results in this paper and in [9], [10], and [11] benefited greatly from conversations with Jerry Bona, Michael Crandall, Simeon Reich, and others.

2. Error stability and nested Banach spaces. The term “stability” has several different meanings in the literature of differential equations. We shall say that an initial value problem or the corresponding semigroup is *error-stable*, or has *continuous dependence on the data*, if a small error in the initial data results in only a small error in the solution. Most useful semigroups must be error-stable in some sense, since no physical measurement is exact.

A transformation  $S(\cdot)$  on a subset  $D$  of a Banach space  $(X, \|\cdot\|)$  is *quasicontractive* if there is some constant  $\omega$  such that

$$(2.1) \quad \|S(t)f - S(t)g\| \leq e^{\omega t} \|f - g\| \quad (t \geq 0, f, g \in D) .$$

This condition implies error-stability: if  $f$  and  $g$  are close and  $t$  is bounded, then  $S(t)f$  and  $S(t)g$  are close. Every semigroup in the Crandall-Liggett framework is quasicontractive ([4]; see also [10]).

Some semigroups of physical interest are not quasicontractive, however. For instance, the semigroup  $S(\cdot)$  corresponding to the Korteweg-deVries equation

$$u_t + uu_x + u_{xxx} = 0$$

is not quasicontractive for any obvious choice of the Banach space  $(X, \|\cdot\|)$ , as is shown later in this section. Still, it is easy to show that this semigroup  $S$  does satisfy the following weaker condition: For any  $t \geq 0$  and  $f, g \in H^2(\mathbf{R})$ ,

$$(2.2) \quad \|S(t)f - S(t)g\|_{L^2(\mathbf{R})} \leq \exp [t \cdot \beta(t, f, g)] \|f - g\|_{L^2(\mathbf{R})}$$

where

$$\beta(t, f, g) = \frac{1}{4} \sup_{0 \leq s \leq t} \max \{ \|S(s)f\|_{H^2(\mathbf{R})}, \|S(s)g\|_{H^2(\mathbf{R})} \}.$$

It is shown in [1] that

$$(2.3) \quad \beta(t, f, g) \text{ is bounded when } t, \|f\|_{H^2(\mathbf{R})}, \|g\|_{H^2(\mathbf{R})} \text{ are bounded.}$$

Hence  $S$  is error stable, in a certain weak sense.

The Crandall-Liggett school of semigroups works entirely in one Banach space. However, many semigroups have useful properties which are best described using more than one space. One such property is given above. A more important property is given in [9]: by using two spaces we can often treat a discontinuous operator (e.g., the generator of a semigroup) as if it were continuous. (This is a variant of techniques using five, four, and three spaces in [6], [8], and [3], respectively.)

We now abstract the situation described in (2.2) and (2.3). The situation we shall work with is roughly as follows:  $(X, \|\cdot\|)$  and  $(D, \|\cdot\|_D)$  are two fixed Banach spaces, with  $D \subset X$ . (Later we shall need a slightly more general choice of  $(D, \|\cdot\|_D)$ .) We consider transformations  $R(\cdot)$  on  $D$  such that

$$(2.4) \quad \begin{aligned} &\text{if } f, g \in D \text{ are } \|\cdot\|_D\text{-close and } \|\cdot\|_D\text{-bounded and } t \text{ is} \\ &\text{bounded, then } R(t)f \text{ and } R(t)g \text{ are } \|\cdot\|_D\text{-close and } \|\cdot\|_D\text{-} \\ &\text{bounded.} \end{aligned}$$

We close this section with an example showing that a one-space approach is not adequate for the KdV equation:

2.5. EXAMPLE. The semigroup  $S(\cdot)$  corresponding to the Korteweg-deVries equation is not quasicontractive in  $L^p(\mathbf{R})$ ,  $1 \leq p \leq \infty$ . In fact, for any  $t > 0$ ,  $S(t)$  is not Lipschitz in  $L^p(\mathbf{R})$ .

*Proof.* Among the solutions  $u(t, x)$  of the KdV equation are the *solitons* or *solitary waves*,

$$u^a(t, x) = 3a^2 \operatorname{sech}^2 \left[ \frac{1}{2} a(x - a^2 t) \right],$$

for constant  $a$ . The wave  $u^a$  maintains a constant shape and velocity. These solutions were known to Korteweg and deVries; see [1] for references.

Let  $f^a(x) = u^a(0, x) = 3a^2 \operatorname{sech}^2(ax/2)$ . Then  $u^a(t, \cdot) = S(t)f^a(\cdot)$  in semigroup notation. It suffices to show that

$$(2.6) \quad \begin{aligned} & \|u^a(t, \cdot) - u^b(t, \cdot)\|_{L^p} / \|f^a - f^b\|_{L^p} \text{ is unbounded as} \\ & a, b \longrightarrow \infty \text{ and } b/a \downarrow 1. \end{aligned}$$

Fix any  $t > 0$  and  $p \in [1, \infty)$ . (We omit the case of  $p = \infty$ , which is similar.)

We shall make several uses of the inequalities

$$(2.7) \quad \begin{cases} \frac{1}{2} \exp(|\theta|) \leq \cosh \theta \leq \exp(|\phi|), \\ |\sinh \theta| \leq \cosh \theta, \quad 1 + |\theta| \leq \exp(|\theta|). \end{cases}$$

From (2.7) we obtain

$$(2.8) \quad 3a^2 \exp[-a|x - a^2t|] \leq |u^a(x, t)| \leq 12a^2 \exp[-a|x - a^2t|].$$

Let  $0 < a < b$ . The waves  $u^a$  and  $u^b$  travel at different velocities; hence their overlap becomes small and their difference becomes large. Using (2.8) we can estimate

$$\begin{aligned} \|u^a(t, \cdot) - u^b(t, \cdot)\|_{L^p(\mathbb{R})} & \geq \|u^a(t, \cdot) - u^b(t, \cdot)\|_{L^p(-\infty, a^2t)} \\ & \geq \|a^a(t, \cdot)\|_{L^p(-\infty, a^2t)} - \|u^b(t, \cdot)\|_{L^p(-\infty, a^2t)} \\ & \geq 3a^2(ap)^{-1/p} - 12b^2(bp)^{-1/p} \exp[-bt(b^2 - a^2)]. \end{aligned}$$

The initial values  $f^a$  and  $f^b$  are close together if  $a$  and  $b$  are close. To get a precise estimate, note that

$$\frac{\partial f^a(x)}{\partial x} = 6a \cosh^{-3}\left(\frac{1}{2}ax\right) \left[ \cosh\left(\frac{1}{2}ax\right) - \left(\frac{1}{2}ax\right) \sinh\left(\frac{1}{2}ax\right) \right];$$

hence, using (2.7),

$$\begin{aligned} \left| \frac{\partial f^a(x)}{\partial x} \right| & \leq 6a \cosh^{-2}\left(\frac{1}{2}ax\right) \left(1 + \left|\frac{1}{2}ax\right|\right) \\ & \leq 24a \exp\left(-\frac{1}{2}|ax|\right). \end{aligned}$$

Therefore, for  $0 < a < b$ ,

$$\begin{aligned} |f^a(x) - f^b(x)| & \leq 24 \int_a^b r \exp\left(-\frac{1}{2}r|x|\right) dr \\ & \leq 24(b - a)b \exp\left(-\frac{1}{2}a|x|\right) \end{aligned}$$

and so

$$(2.10) \quad \|f^a - f^b\|_{L^p} \leq 24(b-a)b \cdot \left(\frac{1}{4}ap\right)^{-1/p}.$$

Combine (2.9) and (2.10), and substitute  $a = \varepsilon^{-1}$ ,  $b = a + 1 = \alpha(1 + \varepsilon)$ . This yields

$$\begin{aligned} & \|u^a(t, \cdot) - u^b(t, \cdot)\|_{L^p} / \|f^a - f^b\|_{L^p} \\ & \geq 2^{(2/p)-3} \varepsilon^{-1} (1 + \varepsilon)^{-1} \{1 - 4(1 + \varepsilon)^{2-(1/p)} \exp[-\varepsilon^{-2}t(1 + \varepsilon)(2 + \varepsilon)]\} \end{aligned}$$

which tends to  $\infty$  as  $\varepsilon \downarrow 0$ . This proves (2.6).

**3. Product stability.** Let  $R$  be a transformation on  $D$  (as defined in Section 1, above). We would like the products  $R(t/n)^n$  to be error-stable, uniformly in  $n$ . Such a condition seems to be an essential ingredient of any proof of (1.2). For instance, Theorem 2.5.3 in [2], Theorem I in [4], and the Equivalence Theorem in [5] each assume a condition of the following type:

$$(3.1) \quad \begin{aligned} & \text{for each } t, \text{ the operators } R(t/n)^n \quad (n = 1, 2, \dots) \\ & \text{are uniformly } |\cdot| \text{-Lipschitz.} \end{aligned}$$

Condition (3.1) can be replaced by stronger, but simpler and more convenient assumptions. For instance, (3.1) holds if  $R$  is quasicontractive (defined in (2.1)). Moreover,  $R(t) = S_1(t)S_2(t)$  is quasicontractive if both  $S_1$  and  $S_2$  are, as was assumed by Trotter in [12].

Let us now replace quasicontractiveness with the more general type of condition indicated in (2.2)-(2.4). Specifically, assume that

$$(3.2) \quad |R(t)f - R(t)g| \leq |f - g| \exp[t \cdot \beta(t, \|f\| \vee \|g\|)]$$

for all  $t \geq 0$  and  $f, g \in D$ , where  $\beta: \mathbf{R}_+^2 \rightarrow \mathbf{R}_+$  is some function which is increasing, hence bounded on bounded sets. It follows that

$$(3.3) \quad |R(t/n)^n f - R(t/n)^n g| \leq |f - g| \exp\{t \cdot \beta[t, \gamma(t, \|f\| \vee \|g\|)]\}$$

uniformly in  $n$ , if  $\gamma: \mathbf{R}_+^2 \rightarrow \mathbf{R}_+$  is some increasing function satisfying

$$(3.4) \quad \sup_{0 \leq k \leq n < \infty} \|R(t/n)^k f\| \leq \gamma(t, \|f\|).$$

Condition (3.3) and (3.4) together imply that the products  $R(t/n)^n$  satisfy a version of (2.4), uniformly in  $n$ . Hypothesis (3.2) is easy to verify in many applications. In the remainder of this section and in §4, below, we shall concentrate on the a priori bounds (3.4), which are harder to verify. We would like to replace (3.4) with a hypothesis which is simpler but not too much stronger.

As a first guess, assume

$$(3.5) \quad \|R(t)f\| \leq e^{\psi t} \|f\| \quad (t \geq 0, f \in D)$$

for some constant  $\psi \geq 0$ . Then (3.4) holds trivially with  $\gamma(t, h) = e^{\psi t}h$ . But (3.5) is not appropriate for nonlinear semigroups. For instance, the initial value problem for the differential equation

$$u'(t) + au(t) + b = 0$$

has solution  $u(t) = R(t)u(0)$ , where  $R$  is a semigroup defined by

$$(3.6) \quad R(t)f = e^{-at}f - \int_0^t e^{-as}b ds \quad (t \geq 0, f \in R).$$

This semigroup does not satisfy (3.5) if  $f = 0, b \neq 0$ , and  $t > 0$ .

For a second guess, replace the linear semigroup  $h \rightarrow e^t h$  with a nonlinear semigroup  $h \rightarrow \sigma(t)h$ . That is, assume

$$(3.7) \quad \|R(t)f\| \leq \sigma(\psi t) \|f\|,$$

where  $\psi$  is some nonnegative constant and  $\sigma(t): R_+ \rightarrow R_+(t \geq 0)$  is an increasing semigroup on  $R_+$ , not necessarily linear. That is,  $\sigma$  satisfies

$$(3.8) \quad \begin{cases} \sigma(0)h = h \\ \sigma(t + t')h = \sigma(t)\sigma(t')h \\ \sigma(t)h \geq \sigma(t')h' \quad \text{if } t \geq t', h \geq h'. \end{cases}$$

Then (3.4) holds trivially with  $\gamma(t, h) = \sigma(\psi t)h$ . Note that  $R(t) = S_1(t)S_2(t)$  satisfies (3.7) if

$$(3.9) \quad \|S_j(t)f\| \leq \sigma(\psi_j t) \|f\| \quad (t \geq 0; f \in D; j = 1, 2)$$

and we take  $\psi = \psi_1 + \psi_2$ . Condition (3.7) generalizes (3.5), since  $\sigma(t)h \equiv e^t h$  satisfies (3.8). Condition (3.7) is sufficiently general for the nonlinear semigroup given in (3.6), since we can take

$$\sigma(t)h \equiv e^t h + e^t - 1, \quad \psi = |a| \vee |b|.$$

Condition (3.7) is still not sufficiently general for many applications. Following is a simple example of a semigroup  $R$  which cannot satisfy (3.7). The semigroup  $R$  given in (3.11) is trivial, but in some respects it is typical of a class of nontrivial semigroups connected with partial differential equations, as we shall see in Section 4.

**EXAMPLE 3.10.** Define a semigroup  $R(t): R^2 \rightarrow R^2$  by

$$(3.11) \quad R(t) \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 + tx_1^2 \end{bmatrix}.$$

Then there do not exist a norm  $\| \cdot \|$  on  $\mathbf{R}^2$  and an increasing semi-group  $\sigma$  on  $\mathbf{R}_+$  such that

$$(3.12) \quad \|R(t)f\| \leq \sigma(t) \|f\| \quad (t \geq 0; f \in \mathbf{R}^2).$$

*Proof.* Suppose that such  $\| \cdot \|$  and  $\sigma$  do exist. Let

$$\phi(s) = \left\| \begin{bmatrix} 1 \\ s \end{bmatrix} \right\|.$$

Then  $\phi$  is convex, and  $\phi(s) \rightarrow \infty$  as  $s \rightarrow \infty$ . Hence there exist some positive constants  $r$  and  $c$  such that  $\phi(r) > 0$  and

$$\phi(s) \geq \phi(r) + (s - r)c \text{ if } s \geq r.$$

Let  $b = c/\phi(r)^2$ . We claim that

$$(3.13) \quad \sigma(t)h \geq h + tbh^2 \quad (t, h \geq 0).$$

Indeed, let  $x = h/\phi(r)$ . Then

$$\begin{aligned} \sigma(t)h &= \sigma(t) \left\| \begin{bmatrix} x \\ xr \end{bmatrix} \right\| \geq \left\| R(t) \begin{bmatrix} x \\ xr \end{bmatrix} \right\| \\ &= x \cdot \phi(r + xt) \geq x \cdot [\phi(r) + xtc] = h + tbh^2, \end{aligned}$$

which proves (3.13).

Now let  $u(s) = \sigma(s)(b^{-1})(s \geq 0)$ . From (3.13) with  $h = u(s)$  we obtain

$$\frac{1}{u(s)} - \frac{1}{u(s+t)} \geq \frac{bt}{1 + btu(s)}.$$

Take  $s = k/n$  and  $t = 1/n$ . Since  $u$  is increasing, we obtain

$$\frac{1}{u\left(\frac{k}{n}\right)} - \frac{1}{u\left(\frac{k+1}{n}\right)} \geq \frac{b/n}{1 + \frac{b}{n}u\left(\frac{k}{n}\right)} \geq \frac{b/n}{1 + \frac{b}{n}u(1)}$$

for integers  $0 \leq k \leq n$ . Summing over  $k$ ,

$$\frac{1}{u(0)} - \frac{1}{u(1)} \geq \frac{b}{1 + \frac{b}{n}u(1)}.$$

Now take limits as  $n \rightarrow \infty$ , and recall that  $u(0) = b^{-1}$ . This yields  $-1/u(1) \geq 0$ , contradicting the fact that  $u(1) \geq u(0) > 0$ . Hence



(3.12) cannot be satisfied.

The semigroup given in (3.11) only grows linearly for large time. But if we try to estimate its size with a semigroup on  $\mathbf{R}_+$ , that semigroup must grow at least quadratically, as in (3.13). We need to measure how “big” a vector is with something more sophisticated than just an element of  $\mathbf{R}_+$ .

*Hypotheses 3.14.*  $\mathcal{H}$  is a partially ordered set. If  $h_1, h_2$  are any two elements of  $\mathcal{H}$ , then their least upper bound  $h_1 \vee h_2$  exists in  $\mathcal{H}$ .

$\sigma(t): \mathcal{H} \rightarrow \mathcal{H} (t \geq 0)$  is an increasing semigroup on  $\mathcal{H}$ ; i.e.,  $\sigma$  satisfies (3.8) for all  $h, h' \in \mathcal{H}$  and  $t, t' \geq 0$ .

$(X, \|\cdot\|)$  is a Banach space.  $D$  is a subset of  $X$ .  $\|\cdot\|: D \rightarrow \mathcal{H}$  is some mapping.

A set  $B \subset D$  is  $\|\cdot\|$ -bounded if  $\{\|b\|: b \in B\}$  is bounded above in  $\mathcal{H}$ . With this terminology, (2.4) still makes sense. The existence of least upper bounds in  $\mathcal{H}$  implies that the union of any two  $\|\cdot\|$ -bounded sets is  $\|\cdot\|$ -bounded.

Where appropriate, we interpret “ $\leq$ ” to be the partial ordering of  $\mathcal{H}$ . Note that if (3.7) holds, then the product stability condition (3.4) holds with  $\gamma(t, h) \equiv \sigma(\psi t)h$ . Also note that if (3.9) holds, then  $R(t) = S_1(t)S_2(t)$  satisfies (3.7) with  $\psi = \psi_1 + \psi_2$ .

The semigroup condition  $\sigma(t + t')h = \sigma(t)\sigma(t')h$  in (3.8) can be replaced with the weaker condition  $\sigma(t + t')h \geq \sigma(t)\sigma(t')h$ . However, in applications  $\sigma$  is generally constructed as the solution of an autonomous ordinary differential equation, and so  $\sigma$  satisfies the semigroup condition.

In the examples in this paper, we shall take  $\mathcal{H}$  to be a Euclidean space,  $\mathbf{R}^k$ , or its positive cone,  $\mathbf{R}_+^k$ , partially ordered by

$$(3.15) \quad \begin{aligned} (x_1, x_2, \dots, x_k) &\geq (y_1, y_2, \dots, y_k) \\ &\text{if and only if } x_1 \geq y_1, \quad x_2 \geq y_2, \dots, x_k \geq y_k. \end{aligned}$$

However, in §5 of [9] we shall use a quite different choice of  $\mathcal{H}$ , taking advantage of the greater generality of Hypotheses 3.14.

For the semigroup  $R$  given in (3.11), take  $X = D = \mathbf{R}^2$  and  $\mathcal{H} = \mathbf{R}_+^2$ . Define  $\|\cdot\|: D \rightarrow \mathcal{H}$  by

$$\left\| \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right\| = \begin{bmatrix} |x_1| \\ |x_2| \end{bmatrix},$$

and let  $\sigma$  be the restriction of  $R$  to  $\mathbf{R}_+^2$ . Then (3.12) follows trivially. In the next section we give some less trivial examples.

#### 4. Examples from partial differential equations. In both of

the following examples, semigroups are related to initial value problems as in (1.6). Our interest in the  $H^2(\mathbf{R})$ -norm is motivated in part by the discussion in § 2, above.

EXAMPLE 4.1. Let  $S_1$  and  $S_2$  be the semigroups corresponding to the initial value problems for

$$(4.2) \quad u_t - u_{xx} = 0 \quad (\text{heat equation}),$$

$$(4.3) \quad v_t + v^3 = 0,$$

respectively. Then  $\| [S_1(t/n)S_2(t/n)]^n f \|_{H^2(\mathbf{R})}$  is bounded uniformly for bounded  $\|f\|_{H^2(\mathbf{R})}$  and  $t$  and all positive integers  $n$ .

*Proof.* We shall take  $\mathcal{H} = \mathbf{R}_+^3$ , partially ordered as in (3.15). Define  $\| \cdot \|: H^2(\mathbf{R}) \rightarrow \mathbf{R}_+^3$  by

$$\|f\| = (\|f\|_{L^2}, \|f_x\|_{L^2}, \|f_{xx}\|_{L^2}).$$

Then a set in  $D = H^2(\mathbf{R})$  is  $\| \cdot \|$ -bounded if and only if it is  $\| \cdot \|_{H^2(\mathbf{R})}$ -bounded. It suffices to exhibit an increasing semigroup  $\sigma$  on  $\mathbf{R}_+^3$  satisfying (3.9). Let

$$(4.4) \quad \sigma(t) \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 + tx_2^3 \end{bmatrix}.$$

The reader can easily verify that  $\sigma$  satisfies (3.8) for all  $t, t' \geq 0$  and  $h, h' \in \mathbf{R}_+^3$ . (Note the similarity of this semigroup  $\sigma$  to the semigroup  $R$  given in (3.11).) Condition (3.9) holds for  $j = 1$  with  $\psi_1 = 0$ , by some well-known properties of the heat equation. We assert that (3.9) holds for  $j = 2$  with  $\psi_2 = 1$ ; i.e., that any solution  $v$  of (4.3) satisfies

$$(4.5) \quad \begin{bmatrix} \|v(t)\|_{L^2} \\ \|v_x(t)\|_{L^2} \\ \|v_{xx}(t)\|_{L^2} \end{bmatrix} \leq \begin{bmatrix} \|v(0)\|_{L^2} \\ \|v_x(0)\|_{L^2} \\ \|v_{xx}(0)\|_{L^2} + t \|v_x(0)\|_{L^2}^3 \end{bmatrix}.$$

We shall only show the third inequality in (4.5) in detail; the first two inequalities follow by simpler versions of the same argument. Differentiate (4.3) twice to obtain

$$v_{xxt} + 6vv_x^2 + 3v^2v_{xx} = 0.$$

At any  $t$  where  $\|v_{xx}(t)\|_{L^2} > 0$ , we can integrate by parts and interpolate as follows:

$$\begin{aligned} \|v_{xx}\|_{L^2} \frac{d}{dt} \|v_{xx}\|_{L^2} &= \frac{1}{2} \frac{d}{dt} \|v_{xx}\|_{L^2}^2 = \frac{1}{2} \int v_{xxt} v_{xx} \\ &= -3 \int v v_x^2 v_{xx} - \frac{3}{2} \int v^2 v_{xx}^2 = \int v_x^4 - \frac{3}{2} \int v^2 v_{xx}^2 \leq \int v_x^4 \\ &\leq \|v_x\|_{L^\infty}^2 \|v_x\|_{L^2}^2 \leq \|v_{xx}\|_{L^2} \|v_x\|_{L^2}^3 \leq \|v_{xx}\|_{L^2} \|v_x(0)\|_{L^2}^3 . \end{aligned}$$

Divide both ends of the above inequality by  $\|v_{xx}\|_{L^2}$ , and integrate over time. Thus we obtain the third inequality in (4.5).

In the above argument, we used estimates on  $\|v_x(t)\|_{L^2}$  to help obtain estimates on  $\|v_{xx}(t)\|_{L^2}$ . This is an example of “bootstrapping of energy estimates.” Note also that the above argument reduces certain questions about the partial differential equations (4.2) and (4.3), and ultimately also the equation

$$w_t - w_{xx} + w^3 = 0 ,$$

to questions about the ordinary differential equation

$$\frac{d}{dt} \begin{bmatrix} \xi_1(t) \\ \xi_2(t) \\ \xi_3(t) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \xi_2(t)^3 \end{bmatrix}$$

which corresponds to the semigroup in (4.4). In the next example we shall use a more complicated ordinary differential equation, and a more complicated choice of  $\| \cdot \|$ . This greater complexity apparently is unavoidable, as is shown in an example in [9].

EXAMPLE 4.6. Let  $S_1$  and  $S_2$  be the semigroups corresponding to the initial value problems for

$$(4.7) \quad u_t - u_{xx} = 0 \quad (\text{heat equation}) ,$$

$$(4.8) \quad v_t + v v_x + v_{xxx} = 0 \quad (\text{KdV equation}) ,$$

respectively. Then  $\|[S_1(t/n)S_2(t/n)]^n f\|_{H^2(\mathbf{R})}$  is bounded uniformly for bounded  $\|f\|_{H^2(\mathbf{R})}$  and  $t$  and all positive integers  $n$ .

*Proof.* Let  $\mathcal{H} = \mathbf{R}^3$ , partially ordered as in (3.15). For any  $f \in D = H^2(\mathbf{R})$ , let

$$\begin{aligned} I_1(f) &= \int f^2 , \\ I_2(f) &= \int f_x^2 - \frac{1}{3} \int f^3 , \\ I_3(f) &= \frac{1}{4} \int f^4 - 3 \int f f_x^2 + \frac{9}{5} \int f_{xx}^2 , \end{aligned}$$

and define  $\| \cdot \| : D \rightarrow \mathcal{H}$  by taking  $\|f\| = (I_1(f), I_2(f), I_3(f))$ . It suffices to show that

(4.9) a subset of  $D$  is  $\| \cdot \|$ -bounded if and only if it is  $\| \cdot \|_{H^2(\mathbb{R})}$ -bounded,

and also exhibit an increasing semigroup  $\sigma$  on  $\mathcal{H}$  satisfying (3.9). Using standard interpolation results, we have

$$\begin{aligned} \|f\|_{L^\infty}^2 &\leq \|f\|_{L^2} \|f_x\|_{L^2} \leq \frac{1}{2} \|f\|_{L^2}^2 + \frac{1}{2} \|f_x\|_{L^2}^2, \\ \|f_x\|_{L^2}^2 &\leq I_2(f) + \frac{1}{3} \int |f|^3 \leq I_2(f) + \frac{1}{2} \int |f|^3, \\ \int |f|^3 &\leq \|f\|_{L^2}^2 \|f\|_{L^\infty} \leq \frac{1}{2} \|f\|_{L^2}^4 + \frac{1}{2} \|f\|_{L^\infty}^2. \end{aligned}$$

Add the above three lines. This leads to

(4.10)  $\|f\|_{L^\infty}^2 + \|f_x\|_{L^2}^2 + \int |f|^3 \leq \|f\|_{L^2}^2 + \|f\|_{L^2}^4 + 2I_2(f).$

(Note: the computations in this proof are chosen for simplicity, not for sharpness.) Inequality (4.10), together with a few straightforward interpolations, yields (4.9).

The numbers  $I_1, I_2, I_3$  are among the well-known invariants of the Korteweg-deVries equation; see [7]. Thus, if  $v$  is any solution of (4.8),  $\|v(t)\|$  does not depend on  $t$ . Therefore (3.9) holds for  $j = 2$  with  $\psi_2 = 0$ , if  $\sigma$  is any increasing semigroup on  $\mathbb{R}^3$ . It suffices to choose  $\sigma$  to satisfy (3.9) for  $j = 1$ .

For any real number  $r$ , let  $r^+ = \max\{r, 0\}$ . Note that  $r^+$  is a nonnegative, continuous, increasing function of  $r$ .

Define an operator  $A: \mathbb{R}^3 \rightarrow \mathbb{R}_+^3$  by

$$A \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} 0 \\ a(w_1)w_2^+ + b(w_1) \\ c(w_1, w_2)w_3^+ + d(w_1, w_2) \end{bmatrix}$$

where  $a, b, c, d$  are nonnegative, continuous, increasing functions which will be specified later. Then for any  $x \in \mathbb{R}^3$ , the initial value problem

(4.11)  $w'(t) = Aw(t), \quad w(0) = x$

has a unique solution defined for all  $t \geq 0$ . (The function  $w_1(t)$  is defined and continuous for all  $t \geq 0$ ; hence so is  $w_2(t)$ ; hence so is  $w_3(t)$ .) Since the system (4.11) is autonomous,  $\sigma(t)x = w(t)$  defines a semigroup on  $\mathbb{R}^3$ . The function  $\sigma(t)x$  is increasing in  $t$  and in  $x$ ,

since  $a, b, c, d$  are nonnegative and increasing. Thus (3.8) holds. We shall verify (3.9) for  $j = 1$  with  $\gamma_1 = 1$ . It suffices to choose  $a, b, c, d$  so that any solution  $u$  of (4.2) satisfies

$$(4.12) \quad \frac{d}{dt} \|u(t)\| \leq A \|u(t)\| .$$

It is well-known that any solution  $u$  of the heat equation satisfies

$$\frac{d}{dt} \int u^2 \leq 0 ,$$

the first component of the vector inequality (4.12).

From  $(P - Q)^4 + 3(P^2 - Q^2)^2 \geq 0$  we obtain

$$(4.13) \quad PQ^3 \leq P^3Q + PQ^3 \leq P^4 + Q^4$$

for all nonnegative  $P$  and  $Q$ . For the second component of (4.12) we use integration by parts, interpolation, (4.13), and (4.10) in the following computation:

$$\begin{aligned} \frac{d}{dt} I_2(u) &= -2 \int u_{xx}^2 + 2 \int uu_x^2 \\ &\leq -2 \|u_{xx}\|_{L^2}^2 + 2 \|u_x\|_{L^\infty} \|u_x\|_{L^2} \|u\|_{L^2} \\ &\leq -2 \|u_{xx}\|_{L^2}^2 + 2 (\|u_{xx}\|_{L^2}^{1/2}) (\|u_x\|_{L^2}^{1/2} \|u\|_{L^2}^{1/3})^3 \\ &\leq -2 \|u_{xx}\|_{L^2}^2 + 2 [(\|u_{xx}\|_{L^2}^{1/2})^4 + (\|u_x\|_{L^2}^{1/2} \|u\|_{L^2}^{1/3})^4] \\ &= 2 \|u_x\|_{L^2}^2 \|u\|_{L^2}^{4/3} \\ &\leq 2[I_1(u) + I_1(u)^2 + 2I_2(u)]I_1(u)^{2/3} \\ &\leq 2I_1(u)^{2/3}I_2(u)^+ + 2[I_1(u)^{5/3} + I_1(u)^{8/3}] . \end{aligned}$$

Note that  $I_1(u)^+ = I_1(u)$ , since  $I_1(u) \geq 0$ . Hence the second component of (4.12) is satisfied with

$$\begin{aligned} a(r) &= 2(r^+)^{2/3} , \\ b(r) &= 2[(r^+)^{5/3} + (r^+)^{8/3}] . \end{aligned}$$

These functions are nonnegative, continuous and increasing.

Define a nonnegative, continuous, increasing function  $e$  by

$$e(r, s) = [r^+ + (r^+)^2 + 2s^+]^{1/2} .$$

Then it follows from (4.10) that

$$\max \{ \|f\|_{L^\infty}, \|f_x\|_{L^2} \} \leq e(I_1(f), I_2(f)) ,$$

hence

$$\begin{aligned} \|f_{xx}\|_{L^2}^2 &\leq \frac{5}{9} \left[ I_3(f) + 3 \int f_x^2 \right] \\ &\leq \frac{5}{9} I_3(f) + \frac{5}{3} \|f\|_{L^\infty} \|f_x\|_{L^2}^2 \\ &\leq \frac{5}{9} I_3(f) + \frac{5}{3} e(I_1(f), I_2(f))^3 . \end{aligned}$$

Hence, for the third component of (4.12),

$$\begin{aligned} \frac{d}{dt} I_3(u) &= -3 \int u^2 u_x^2 - \frac{18}{5} \int u^3_{xxx} + 6 \int u u_{xx}^2 \\ &\leq 6 \|u\|_{L^\infty} \|u_{xx}\|_{L^2}^2 \\ &\leq 6 \cdot e(I_1(u), I_2(u)) \cdot \left[ \frac{5}{9} I_3(u) + \frac{5}{3} e(I_1(u), I_2(u))^3 \right] . \end{aligned}$$

Therefore we can take  $c(r, s) = 10e(r, s)/3$ ,  $d(r, s) = 10e(r, s)^4$ .

5. Statement of main results. We assume the hypotheses of Section 3.14. In addition we assume the mapping  $\| \cdot \| : D \rightarrow \mathcal{H}$  satisfies the following condition:

(5.1) Suppose  $f \in X$ ,  $h \in \mathcal{H}$ , and  $\{f_n\}$  is a sequence in  $D$ . Suppose  $f_n \rightarrow f$  in  $X$ , and  $\|f_n\| \leq h$  for all  $n$ . Then  $f \in D$ , and  $\|f\| \leq h$ .

Condition (5.1) says that  $D$  is “closed” in a certain sense. This condition will contribute greatly to the simplicity of our results. It is assumed only as a convenience of notation, for it can always be satisfied by modifying  $D$ ,  $\mathcal{H}$ , and  $\| \cdot \|$  slightly (i.e., by replacing  $D$  by its “closure”). This is discussed in detail in § 5 of [9]. For motivation, note that (5.1) follows by a weak compactness argument if  $\mathcal{H} = \mathbf{R}_+$  and  $(D, \| \cdot \|)$  is a reflexive Banach space continuously included in  $(X, | \cdot |)$ .

NOTATION 5.2.

(a) All topological considerations, e.g., convergences and continuity, will be in the topology of  $(X, | \cdot |)$ .

(b) Let  $\beta$  be a function of several arguments, each of which has domain  $\mathbf{R}_+$  or  $\mathcal{H}$ . Assume  $\beta$  has range in  $\mathbf{R}_+$ , and  $\beta$  is increasing in each argument. Then we shall call  $\beta$  a *bounding function*, since  $\beta$  is bounded on bounded sets.

(c) Let  $S$  be a transformation on  $D$  (as defined in §1). Let  $\psi \in \mathbf{R}_+$ , and let  $\omega : \mathcal{H} \rightarrow \mathbf{R}_+$  be a bounding function. The triple  $(S, \psi, \omega)$  is a *restrained transformation* (with respect to  $\mathcal{H}, \sigma, X, | \cdot |, D, \| \cdot \|$ ) if

$$\begin{aligned} \|S(t)f\| &\leq \sigma(\psi t) \|f\| \\ |S(t)f - S(t)g| &\leq |f - g| \exp \{t \cdot \omega[\sigma(\psi t)(\|f\| \vee \|g\|)]\} \end{aligned}$$

for all  $t \geq 0$  and  $f, g \in D$ .  $(S, \psi, \omega)$  is a *restrained semigroup* if in addition

$$S(t + s)f = S(t)S(s)f \quad (t, s \geq 0; f \in D).$$

In [9] we shall use the above definitions, plus a few more definitions involving generators, to prove results about equations (1.3)-(1.5). However, the above framework is already adequate for an interesting result about (1.2).

Theorem 5.3, below, is similar to the results in [6] and in Theorem 5.3 of [3]. The proofs are also related. The results below are simpler in that (i) generators are not mentioned, in the hypotheses or the conclusions, and (ii) only the two spaces  $D \subset X$  are used, rather than three or five spaces as in [3] or [6].

The result below also differs in the “almost-semigroup” condition (5.4). The analogous hypothesis in [3] or [6] (translated to the notation of the present paper) is

$$(5.4') \quad |R(t)R(s)f - R(t + s)f| \leq ts \cdot \gamma(t + s, \|f\|).$$

This is similar to (5.4) for  $\lambda = 1$ , but is stronger. Some justification of the above condition is given in [6] for products of  $C^2$ -semigroups. The weaker condition (5.4) seems to be more natural; it is satisfied by a class of  $C^1$ -semigroups considered in [9].

It is mostly because of this weakened hypothesis that the proof in Section 6 is so long and complicated. Inequality (5.4) is not very strong except when  $t$  and  $s$  are of nearly the same order—e.g., when  $1/2 < t/s < 2$  and hence  $(t + s)^2 = \mathcal{O}(ts)$ . This restriction is violated if we take  $s = t/n$ , and so the simple, direct proof of (5.5) in [3] and [6] fails if we replace (5.4') with the weaker (5.4).

**THEOREM 5.3.** *Let  $(R, \psi, \omega)$  be a restrained transformation. Suppose  $R$  is an “almost-semigroup,” in the sense that there exists a constant  $\lambda > 0$  and a bounding function  $\gamma: \mathbf{R}_+ \times \mathcal{H} \rightarrow \mathbf{R}_+$  such that*

$$(5.4) \quad |R(t)R(s)f - R(t + s)f| \leq (t + s)^{1+\lambda} \gamma(t + s, \|f\|)$$

for all  $f \in D$  and  $t, s \geq 0$ .

Then there exists  $S$  such that  $(S, \psi, \omega)$  is a restrained semigroup and

$$(5.5) \quad R(t/n)^n f \longrightarrow S(t)f \text{ as } n \longrightarrow \infty,$$

uniformly for bounded  $\|f\|$  and  $t$ . In fact, we have the following rate of convergence: Define a bounding function  $\beta: \mathbf{R}_+ \times \mathcal{H} \rightarrow \mathbf{R}_+$  by

$$(5.6) \quad \beta(T, h) = T \exp \{T \cdot \omega[\sigma(\psi T)h]\} \cdot \gamma[T, \sigma(\psi T)h].$$

Then

$$(5.7) \quad |R(t/n)^n f - S(t)f| \leq (1 - 2^{-\lambda})^{-1} (t/n)^\lambda \beta(t, \|f\|).$$

More generally, let  $s_1 + s_2 + \cdots + s_n = t$ . Then

$$(5.8) \quad R(s_1)R(s_2)\cdots R(s_n)f \longrightarrow S(t)f \text{ as } \max_i s_i \longrightarrow 0,$$

uniformly for bounded  $t$  and  $\|f\|$ . We have this convergence rate:

$$(5.9) \quad \begin{aligned} & |R(s_1)R(s_2)\cdots R(s_n)f - S(t)f| \\ & \leq (1 - 2^{-\lambda})^{-1} (\max_i s_i)^\lambda \beta(t, \|f\|). \end{aligned}$$

**THEOREM 5.10.** *In addition to the above hypotheses, assume  $R$  is strongly continuous at 0; i.e., for each  $f \in D$ ,*

$$R(t)f \longrightarrow f \quad \text{as } t \downarrow 0.$$

*Then  $S$  is strongly continuous; i.e., for each  $f \in D$ ,*

$$t \longrightarrow S(t)f \text{ is continuous from } \mathbf{R}_+ \text{ into } X.$$

Theorem 5.3 is sharp. We shall give two simple examples showing that it cannot be improved in certain directions. Neither of these examples is pathological. In fact, in both cases we shall take  $R(t)$  ( $t \geq 0$ ) to be a family of rotations of the complex plane. That is, take  $D = X = C$  and

$$(5.11) \quad R(t)f = f \exp [i\theta(t)] \quad (f \in C, t \geq 0)$$

for some function  $\theta: \mathbf{R}_+ \rightarrow \mathbf{R}$  satisfying  $\theta(0) = 0$ . In both cases we shall take  $\theta$  continuous on  $[0, +\infty)$  and smooth on  $(0, +\infty)$ .

**EXAMPLE 5.12.** Hypothesis (5.4) cannot be weakened much. There exists a restrained transformation satisfying the weakened "almost-semigroup" condition

$$(5.13) \quad |R(t)R(s)f - R(t+s)f| \leq \frac{t+s}{\log \frac{1}{t+s}} \gamma(t+s, \|f\|)$$

for small  $t+s$ , but such that the sequence  $\{R(t/n)^n f\}$  does not converge.



EXAMPLE 5.14. The convergence rate  $\mathcal{O}(n^{-\lambda})$  given in (5.7) is best possible. For any  $\lambda > 0$ , there exists a transformation  $R$  satisfying the hypotheses of the theorem and such that  $n^\lambda[R(t/n)^n f - S(t)f]$  converges to a nonzero limit.

6. Proofs of main results.

*Proof of Theorem 5.3.* Fix any  $T > 0$ , and consider all partitions

$$(6.1) \quad \pi: 0 = t_0 < t_1 < \dots < t_n = T$$

of the interval  $[0, T]$ . For each partition  $\pi$ , let  $M(\pi) = \max_i (t_i - t_{i-1})$ , and let

$$P(\pi) = R(t_n - t_{n-1})R(t_{n-1} - t_{n-2}) \cdots R(t_1 - t_0).$$

Most of the proof will be directed toward showing that  $|P(\pi)f - P(\pi')f|$  is small if  $M(\pi), M(\pi')$  are small, and hence that  $\{P(\pi)f\}$  is a Cauchy net in  $(X, |\cdot|)$ .

For simplicity of notation let

$$(6.2) \quad U_\pi(t_j, t_i) = R(t_j - t_{j-1})R(t_{j-1} - t_{j-2}) \cdots R(t_{i+1} - t_i)$$

if  $0 \leq i \leq j \leq n$ . In particular,  $U_\pi(T, 0) = P(\pi)$  and  $U_\pi(t_i, t_i)$  is the identity map on  $D$ .

By abuse of notation let  $\pi$  also represent the set  $\{t_0, t_1, \dots, t_n\}$ . Then partition  $\pi'$  is a refinement of partition  $\pi$  if and only if the set  $\pi'$  includes the set  $\pi$ .

Let  $\pi$  be as in (6.1) and let  $\pi'$  be some refinement of  $\pi$ . Then  $\pi'$  is a *first-order refinement* of  $\pi$  if each open interval  $(t_{i-1}, t_i)$  contains at most one element of the set  $\pi'$ . In this case we can estimate  $|P(\pi)f - P(\pi')f|$  as follows:

For  $1 \leq i \leq n$  we have  $U_\pi(t_i, t_{i-1}) = R(t_i - t_{i-1})$ . Let

$$I = \{i: (t_{i-1}, t_i) \text{ contains some } s_i \in \pi'\}.$$

Then

$$U_{\pi'}(t_i, t_{i-1}) = \begin{cases} R(t_i - s_i)R(s_i - t_{i-1}) & \text{if } i \in I, \\ R(t_i - t_{i-1}) & \text{if } i \notin I. \end{cases}$$

Let  $f_i = U_{\pi'}(t_i, 0)f$  ( $0 \leq i \leq n$ ), with  $f_0 = f$  and  $f_n = P(\pi')f$ . The  $f_i$ 's are well-defined by (6.2) since  $\pi' \supset \pi$ . With these notations we can write a telescoping sum:

$$(6.3) \quad \begin{aligned} P(\pi)f - P(\pi')f &= \sum_{i \in I} [U_\pi(T, t_i)U_\pi(t_i, t_{i-1})f_{i-1} \\ &\quad - U_\pi(T, t_i)U_{\pi'}(t_i, t_{i-1})f_{i-1}]. \end{aligned}$$

Note that for any  $g, g' \in D$ ,

$$\begin{aligned} & |U_\pi(T, t_i)g - U_\pi(T, t_i)g'| \\ & \leq \exp \{T \cdot \omega[\sigma(\psi(T - t_i))(\|g\| \vee \|g'\|)]\} \end{aligned}$$

by repeated use of (3.8) and the inequalities in 5.2(c). Substitute this inequality into (6.3), with  $g = U_\pi(t_i, t_{i-1})f_{i-1}$  and  $g' = U_{\pi'}(t_i, t_{i-1})f_{i-1}$ . Also using the triangle inequality, (5.4), and (5.6), we obtain

$$\begin{aligned} & |P(\pi)f - P(\pi')f| \\ (6.4) \quad & \leq \exp \{T \cdot \omega[\sigma(\psi T) \|f\|]\} \cdot \sum_{i \in I} |U_\pi(t_i, t_{i-1})f_{i-1} - U_{\pi'}(t_i, t_{i-1})f_{i-1}| \\ & \leq \exp \{T \cdot \omega[\sigma(\psi T) \|f\|]\} \cdot \sum_{i \in I} (t_i - t_{i-1})^{1+\lambda\gamma} [t_i - t_{i-1}, \sigma(\psi T) \|f\|] \\ & \leq M(\pi)^\lambda \beta(T, \|f\|). \end{aligned}$$

A partition  $\pi'$  is a  $k$ th-order refinement of  $\pi$  if there exist refinements

$$\pi = \pi_0 \subset \pi_1 \subset \cdots \subset \pi_k = \pi'$$

such that each  $\pi_{j+1}$  is a first-order refinement of  $\pi_j$ . Then it follows from (6.4) and the triangle inequality that

$$(6.5) \quad |P(\pi)f - P(\pi')f| \leq \sum_{j=0}^{k-1} M(\pi_j)^\lambda \beta(T, \|f\|).$$

Since  $M(\pi) = M(\pi_0) \geq M(\pi_1) \geq \cdots \geq M(\pi_k)$ , we obtain the crude estimate

$$(6.6) \quad |P(\pi)f - P(\pi')f| \leq kM(\pi)^\lambda \beta(T, \|f\|),$$

which is useful for small integers  $k$ .

If in addition  $M(\pi_{j+1}) = M(\pi_j)/2$  for each  $j$ , we shall call  $\pi'$  a *dyadic refinement* of  $\pi$ . In this case we obtain from (6.5)

$$\begin{aligned} & |P(\pi)f - P(\pi')f| \leq \sum_{j=0}^{k-1} 2^{-\lambda j} M(\pi)^\lambda \beta(T, \|f\|) \\ (6.7) \quad & \leq \sum_{j=0}^{\infty} 2^{-\lambda j} M(\pi)^\lambda \beta(T, \|f\|) \\ & = (1 - 2^{-\lambda})^{-1} M(\pi)^\lambda \beta(T, \|f\|). \end{aligned}$$

Note that the right side does not depend on  $k$ .

For any partition  $\pi$  as in (6.1), let  $m(\pi) = \min_i (t_i - t_{i-1})$ . We now make several observations about refinements:

$$(6.8) \quad \text{Any partition } \pi \text{ has a first-order refinement } \pi' \text{ with } m(\pi') = m(\pi)/2 \text{ and } M(\pi') = M(\pi)/2.$$

(*Proof:* bisect each subinterval of  $\pi$ .)

Suppose  $\pi$  is a partition with  $M(\pi) \geq 2m(\pi)$ . Then  $\pi$  has  
 (6.9) a first-order refinement  $\pi'$  with  $m(\pi') = m(\pi)$  and  $M(\pi') = M(\pi)/2$ .

(*Proof:* bisect every subinterval of  $\pi$  having length at least  $2m(\pi)$ .)

Suppose  $\pi$  is a partition and  $0 < p \leq m(\pi)$ . Then  $\pi$  has  
 (6.10) a dyadic refinement  $\pi'$  satisfying  $1 \leq m(\pi')/p < 2$  and  $1 \leq M(\pi')/m(\pi') < 2$ .

(*Proof:* apply (6.9) several times, until  $1 \leq M(\pi')/m(\pi') < 2$  is satisfied. Then apply (6.8) several times, until  $1 \leq m(\pi')/p < 2$  is also satisfied.)

Let  $\pi'_1$  and  $\pi'_2$  be partitions, with  $2^k m(\pi'_1) \geq M(\pi'_2)$  for some  
 (6.11) nonnegative integer  $k$ . Then  $\pi' = \pi'_1 \cup \pi'_2$  is a refinement of  $\pi'_2$  of order at most  $k + 1$ .

(*Proof:* each open subinterval  $(t_{i-1}, t_i)$  of  $\pi'_2$  contains at most  $2^k$  points of  $\pi'_1$ , hence certainly not more than  $2^{k+1} - 1$  of them. We can add one point to  $(t_{i-1}, t_i)$  by taking a first-order refinement of  $\pi'_2$ ; then two more points by refining again; then four more; etc.)

We are now ready to estimate  $|P(\pi_1)f - P(\pi_2)f|$  for arbitrary partitions  $\pi_1$  and  $\pi_2$ . Fix some small  $p$ , with  $0 < p \leq \min \{m(\pi_1), m(\pi_2)\}$ . Choose dyadic refinements  $\pi'_1, \pi'_2$  as in (6.10). Then

$$(6.12) \quad M(\pi'_2) < 2m(\pi'_2) < 4p \leq 4m(\pi'_1) \leq 4M(\pi'_1).$$

By (6.11),  $\pi' = \pi'_1 \cup \pi'_2$  is a refinement of  $\pi'_2$  of order at most 3. Similarly,  $\pi'$  is a refinement of  $\pi'_1$  of order at most 3. Now use (6.6), (6.7), and the triangle inequality to obtain

$$\begin{aligned} & |P(\pi_1)f - P(\pi_2)f| \\ & \leq (1 - 2^{-\lambda})^{-1} [M(\pi_1)^\lambda + M(\pi_2)^\lambda] \beta(T, \|f\|) \\ & \quad + 3[M(\pi'_1)^\lambda + M(\pi'_2)^\lambda] \beta(T, \|f\|). \end{aligned}$$

By (6.12),  $M(\pi'_2) < 4p$ . Similarly,  $M(\pi'_1) < 4p$ . Let  $p \rightarrow 0$ ; thus

$$(6.13) \quad \|P(\pi_1)f - P(\pi_2)f\| \leq (1 - 2^{-\lambda})^{-1} [M(\pi_1)^\lambda + M(\pi_2)^\lambda] \beta(T, \|f\|).$$

Hence the net  $\{P\{\pi\}f\}$  is Cauchy in the Banach space  $(X, \|\cdot\|)$ . Denote the limit by  $S(T)f$ . Hold  $\pi_1$  fixed in (6.13) and let  $M(\pi_2) \rightarrow 0$  to obtain (5.9) and hence (5.7).

It follows easily from (3.8) and (5.1) that  $S(T)f$  is in  $D$  and

that  $(S, \psi, \omega)$  is a restrained transformation.

To show that  $S$  is a semigroup, fix any  $s, t \geq 0$  and  $f \in D$ . Let  $g = S(t)f$ . Then  $R(t/n)^n f \rightarrow g$  and  $R(s/n)^n \rightarrow S(s)g$ , by (5.5). Also, the set  $\{R(t/n)^n f: n=1, 2, \dots\} \cup \{g\}$  is  $\|\cdot\|$ -bounded, and the operators  $R(s/n)^n$  ( $n=1, 2, \dots$ ) are uniformly  $\|\cdot\|$ -Lipschitz on  $\|\cdot\|$ -bounded sets. It follows that  $R(s/n)^n R(t/n)^n f \rightarrow S(s)S(t)f$ . On the other hand,  $R(s/n)^n R(t/n)^n f \rightarrow S(s+t)f$  by (5.8). Hence  $S(s)S(t) = S(s+t)$ .

*Proof of Theorem 5.10.* For each fixed  $n$ ,  $R(t)^n$  is strongly continuous at 0; this follows by induction on  $n$ . Since the convergence (5.5) is uniform for bounded  $t$ ,  $S(t)$  is strongly continuous at 0. Since  $S$  is a semigroup,  $S(t)f$  is continuous from the right. It suffices to show that  $S(t-h)f \rightarrow S(t)f$  as  $h \downarrow 0$ , for fixed  $t > 0$  and  $f \in D$ . For  $0 \leq h \leq t$ ,  $S(t-h)$  is uniformly  $\|\cdot\|$ -Lipschitz on  $\|\cdot\|$ -bounded sets, and  $S(h)f$  is  $\|\cdot\|$ -bounded. Hence

$$\|S(t-h)f - S(t)f\| = \|S(t-h)f - S(t-h)S(h)f\| \leq \|S\|_{\text{Lip}} \|f - S(h)f\|,$$

which tends to 0 as  $h \downarrow 0$  since  $S$  is strongly continuous at 0.

*Proofs of Examples 5.12 and 5.14.* It follows from (5.11) that each  $R(t)$  is a linear isometry of  $C$ ; hence  $R$  is a restrained transformation. The product formula takes the form

$$R(t/n)^n f = f \cdot \exp \{in\theta(t/n)\},$$

which converges to some limit  $S(t)f = fe^{i\mu(t)}$  if and only if there exists a limit

$$(6.14) \quad \mu(t) = \lim_{n \rightarrow \infty} n\theta(t/n) \pmod{2\pi}.$$

Also from (5.11) we obtain

$$(6.15) \quad |R(t)R(s) - R(t+s)| \leq |\theta(t) + \theta(s) - \theta(t+s)|.$$

These considerations reduce our questions about  $R$  to questions about  $\theta$ .

We only need to consider the behavior of  $R$  and  $\theta$  for small  $s$  and  $t$ . For any continuous choice of  $R$ , we can always satisfy (5.4) for large  $t+s$  by taking  $\gamma$  sufficiently large.

For Example 5.12, it suffices to choose  $\theta: [0, T) \rightarrow \mathbf{R}$  for some  $T > 0$  so that

$$(6.16) \quad |\theta(t) + \theta(s) - \theta(t+s)| \leq c \cdot \frac{t+s}{\log\left(\frac{1}{t+s}\right)}$$

for some constant  $c$  and all small  $t+s$ , and so that

(6.17)  $n\theta(t/n)$  does not converge (mod  $2\pi$ ) as  $n \rightarrow \infty$ .

We shall show that these conditions are satisfied by

$$\theta(t) = t \log |\log t| \quad \left(0 < t < \frac{1}{e}\right).$$

To prove (6.16), temporarily fix any  $h$  in  $(0, 1/e)$ , and let  $p(s) = \theta(s) + \theta(h - s)$  ( $0 \leq s \leq h$ ). Then  $p$  is concave, nonnegative, and symmetric about  $s = h/2$ . It follows that

$$\begin{aligned} |\theta(s) + \theta(h - s) - \theta(h)| &= p(s) - p(0) \\ &\leq p\left(\frac{1}{2}h\right) - p(0) = h \log \left[1 - \frac{\log 2}{\log \frac{1}{h}}\right] \\ &\leq \frac{1}{1 - \log^2} \cdot \frac{h}{\log \frac{1}{h}} \end{aligned}$$

which proves (6.16).

For large  $n$ ,

$$(n + 1)\theta\left(\frac{t}{n + 1}\right) - n\theta\left(\frac{t}{n}\right) \sim \frac{t}{n \log n},$$

which decreases to 0 but is not summable. Hence (6.17).

For Example 5.14, take  $\theta(t) = t^{1+\lambda}$ . The details are straightforward and are omitted.

#### REFERENCES

1. J. L. Bona and R. Smith, *The initial-value problem for the Korteweg-de Vries equation*, Philos. Transac. Royal Soc. of London, **278A** (1975), 555-601.
2. P. R. Chernoff, *Product formulas, nonlinear semi-groups, and addition of unbounded operators*, Mem. Amer. Math. Soc., **140** (1974).
3. A. J. Chorin, T. J. Hughes, M. F. McCracken, and J. E. Marsden, *Product formulas and numerical algorithms*, Comm. Pure and Appl. Math., **31** (1978), 205-256.
4. M. G. Crandall and T. M. Liggett, *Generation of semi-groups of nonlinear transformations on general Banach spaces*, Amer. J. of Math., **93** (1971), 265-298.
5. P. D. Lax and R. D. Richtmyer, *Survey of the stability of linear finite difference equations*, Comm. Pure and Appl. Math., **9** (1956), 267-293.
6. J. Marsden, *On product formulas for nonlinear semigroups*, J. Functional Analysis, **13** (1973), 51-72.
7. R. M. Miura, C. S. Gardner, and M. D. Kruskal, *Korteweg-de Vries equation and generalizations, II: existence of conservation laws and constants of motion*, J. Math. Physics, **9** (1968), 1204-1209.
8. G. C. Papanicolaou and S. R. S. Varadhan, *A limit theorem with strong mixing in Banach space and two applications to stochastic differential equations*, Comm. Pure and Appl. Math., **26** (1973), 497-524.
9. E. Schechter, *Product formulas for continuously differentiable semigroups*, to appear.

10. E. Schechter, *Product formulas and limit theorems for differentiable evolutions*, to appear.
11. ———, *Well-Behaved Evolutions and Trotter Products*, Thesis, U. of Chicago, 1978.
12. H. F. Trotter, *On the product of semi-groups of operators*, Proc. Amer. Math. Soc., **10** (1959), 545-551.

Received November 7, 1978.

DUKE UNIVERSITY  
DURHAM, NC 27706