On Product Formulas for Nonlinear Semigroups

J. Marsden*

Institut des Hautes Études Scientifiques, 91 - Bures-sur-Yvette, France and Department of Mathematics, University of California, Berkeley, Berkeley, California 94720

Communicated by Tosio Kato
Received August 2, 1972

We study a number of sufficient conditions which guarantee the convergence of semigroup product formulas of the type

\[ H_t = \lim_{n \to \infty} (F_{t/n} \circ G_{t/n})^n \]

and its generalizations. Our hypotheses differ from those of other authors in that we do not assume in advance that the limit operator is a generator. Rather this is a consequence and hence the above formula yields an existence theorem (local in time) for nonlinear semigroups. A number of smoothness properties are studied as well. The results may be applied to and are motivated by the Navier–Stokes equations.

1. INTRODUCTION

This paper deals with the following situation: suppose a (nonlinear) operator \( X \) on a Banach space generates a semigroup \( F_t \) (which we shall also call a flow or semi-flow) and \( Y \) generates a semigroup \( G_t \). Then the semigroup for \( X + Y \) ought to be

\[ H_t = \lim_{n \to \infty} (F_{t/n} \circ G_{t/n})^n. \]  \hspace{1cm} (1)

Results centering around formula (1) have been given by Trotter [22] for the linear case and Brezis–Pazy [1] for the nonlinear case in the setting of contractive semigroups. In addition to formula (1), we shall be dealing with an important generalization of (1) due to Chernoff [3], and also treated by Brezis–Pazy [1]. Namely if \( K(t) \)

* Some of the ideas in this paper were obtained in collaboration with David G. Ebin, SUNY at Stony Brook, N.Y. 11790. Partially supported by NSF grant GP 11735.

Copyright © 1973 by Academic Press, Inc.
All rights of reproduction in any form reserved.
is a curve of nonlinear mappings, then the semi-group for $K'(0)$ ought to be

$$H_t = \lim_{n \to \infty} K(t/n)^n.$$  \hspace{1cm} (2)

A good introduction to the formulas (1) and (2) is found in Nelson [16]. Usually one makes hypotheses of the following kind. One assumes that $X + Y$ or $K'(0)$ has a semigroup associated with it. Our approach is rather different; we want to prove the existence of a semigroup for $X + Y$ (or $K'(0)$). The reason for this is that we shall be working with general semigroups, not necessarily contractive ones. For such semigroups, there is no general criterion for determining whether or not $X + Y$ is a generator. (For the contractive case, one has various generalizations of the Hille–Yosida theorem; cf. [1, 2, 9].)

Because of the strength of the conclusions, we put on fairly restrictive hypotheses. In the linear case it amounts to $Y$ being a bounded perturbation of $X.$ Despite this restriction, the theorems have several important applications.

In connection with our theorem concerning the existence of a semigroup for $X + Y,$ we mention some related work of Segal [17]. He shows $X + Y$ generates a semigroup if $Y$ is Lipschitzian and $X$ is linear. Our result allows both $X$ and $Y$ to be nonlinear, as well as establishing formula (1).

The results presented here are motivated by certain applications in hydrodynamics. See Ebin–Marsden [10] for an application of (1) and (2). In that case, both $X$ and $Y$ were nonlinear and we knew a priori, both generated semigroups (certainly not contractive ones). We then wanted to show that $X + Y$ generates a semigroup, and to study the singular perturbation problem $\nu X + Y$ in the limit $\nu \to 0.$ This was used to study the Navier–Stokes equations in a region with no boundary. This paper refines those techniques and eliminates some important hypotheses. In addition, we obtain sufficient conditions for the convergence of Chernoff’s formula (2) as a more general case.

Some of the delicacies in the proof center around the regularity of solutions. Thus for formula (2), one wishes to know if a solution $x(t)$ of $x'(t) = K'(0) [x(t)]$ has initial data from a space with a certain degree of differentiability, then the solution has this same property. This involves then, some kind of a priori estimates. This property was verified for the Euler equations in hydrodynamics in [10]. See also [14]. We want to include here, an abstraction of this regularity property as well as establishing the convergence of formulas (1) and (2).
The main results of the paper are contained in Theorems 2.1, 2.10, 5.1, 5.2, 6.1.

Finally we wish to point out the utility of formula (1) in a number of other applied areas. For example, we cite Nelson [15] for quantum theory, Chorin [7, 8] and Temam [21] for numerical work on the Navier–Stokes equations, and Segal [18] and Simon and Hoegh Krohn [19] in quantum field theory.

In later work we hope to establish the validy of certain product formulas for the Navier–Stokes equations in regions with boundary that are suggested by the recent successful work of A. Chorin [8].

We than P. Chernoff, J. P. Penot and the referee for comments and corrections.

2. Sufficient Conditions for the Convergence of Chernoff's Formula (2)

In order to adequately deal with the case $X + Y$, it is necessary, as we shall see, to introduce metrics other than Banach space norms. Therefore, we deal with formula (2) in a more general context than Banach spaces, namely Banach manifolds with a certain distance metric specified. Furthermore, instead of working with domains of operators, we have found it necessary to work with chains of Banach manifolds: $M = M_0 \supset M_1 \supset M_2 \supset \cdots$ where each $M_i$ is densely included in $M_{i-1}$. For example $M_1$ plays the role of the domain of an operator with the graph topology (or a stronger one). Actually we need only $M_0$ through $M_4$ to obtain convergence on $M$.

In our first main result, 2.1 below, we give a theorem corresponding to the globally contractive case. In hydrodynamics examples, one of the hypotheses ((iii) below) is unrealistic. The regularity conclusion in the theorem runs deeper than that. Therefore, we give in 2.10 some other conditions which yield the same result but get to the heart of the regularity question. Theorem 2.10 represents, we feel, an abstraction of what is going on in the case of the hydrodynamics (Navier–Stokes) equations. The result can then be used to prove, for example, the various product formulas for the Navier–Stokes equations (see [10], [14], and Temam [21]).

**Theorem 2.1.** Let $M = M_0 \supset M_1 \supset M_2 \supset M_3 \cdots$ be Banach manifolds with $M = M_0 \supset M_1 \supset M_2 \supset \cdots$ such that each inclusion is continuous and dense. Let $d_0 = d, d_1, d_2, \cdots$ be complete metrics for $M, M_1, M_2, \cdots$ respectively; $d \leq d_1 \leq d_2 \cdots$. 
Let \( x_0 \in M \), let \( V \subset U \) be open neighborhoods of \( x_0 \) and let
\[ V_k = M_k \cap V, \quad U_k = M_k \cap U. \]
Suppose we are given a curve of mappings
\[ t \in [0, T] \mapsto K(t) : V \to U \]
such that \( K(t) \) maps \( V_k \) to \( U_k \), and \( K(t) \) satisfies the following conditions.

(i) \( K(0) x = x, \quad x \in V. \)

(ii) For \( x \in V_k \), \( t \mapsto K(t) x \in U_k \) is continuous, \((k \geq 0)\),
\( t \mapsto K(t) x \in U_{k-1} \) is \( C^1 \) \((k \geq 1)\) and \( K'(t) : V_k \to TU_{k-1} \) is continuous
\((TU\) is the tangent space of \( U)\).

(iii) There are constants \( \beta = \beta_0, \beta_1, \ldots \geq 0 \) such that for
\( x, y \in V_k \)
\[ d_k(K(t) x, K(t) y) \leq e^{\beta t} d_k(x, y). \] (3)

(iv) For \( x \in V_{k+2} \) and \( B \) a bounded subset of \( V_{k+2} \) containing \( x \),
there is a constant \( C_0 \) such that
\[ d_k(K(t + s) y, K(t) K(s) y) \leq C_0 t s \] (4)
for all \( y \in B \) (the constant may depend on \( x, B, \) and \( k) \).

Then there exists a neighborhood \( W \) of \( x_0 \) in \( M \) and a \( \tau > 0 \) such that
for \( x \in W, \quad 0 \leq t \leq \tau \)
\[ H_{x} x = \lim_{n \to \infty} K(t/n)^n x \] (2)
is defined and exists. Furthermore: (Regularity) If \( x \in W \cap M_k \), then
\( H_{x} x \in M_k \) for all \( 0 \leq t \leq \tau \) (that is, the time of existence \( \tau \) is independent of \( k) \).

Moreover, we have

(a) \( d(H_x x, H_y y) \leq e^{\beta t} d(x, y) \)

(b) \( t \mapsto H_x x \) is continuous

(c) \( H_{t+s} x = H_t \circ H_s x, \text{ if } x \in W, \quad 0 \leq t + s \leq \tau. \)

(d) for \( x \in M_1, \quad H_x x \in M_1, \) and as a curve in \( M, \)
\[ H_0(x) = x, \quad \frac{d}{dt} H_x x = K'(0) (H_x x), \quad 0 \leq t \leq \tau. \] (5)

Moreover, \( H_x x \) is the unique solution of the differential Eq. (5).

(See Lemma 2.8 below for the precise meaning of (d).)
Remark 1. It is in the sense (d) that \( H_{\cdot \cdot} \) is the semigroup with generator \( K'(0) \) (see Section 5 below).

Remark 2. In [22, 3, 1] the key hypothesis is that \( K'(0) \) be a generator. Notice that we conclude this. Instead, we make the a priori estimate (4) which says that \( K(t) \) is, for small \( t \), an approximate semigroup. In Section 3, below we shall see how to verify the estimates (3) and (4) in the context of formula (1). A weakening of condition (iii) is given in Theorem 2.10 below.

Remark 3. We establish (2) and (5), only for small \( t \)-intervals. However (2) is valid as long as one has solutions to (5); in other words, global convergence of (2) is equivalent to global solutions of (5).

Lemma 2.2. There is a neighborhood \( W \) of \( x_0 \) in \( M \) such that for \( x \in W \), \( K(t/n)^n x \) is defined and remains in \( V \) for all \( n = 1, 2, 3, \ldots \) and \( t \) in some interval \( [0, \tau] \), \( \tau > 0 \).

Proof. Choose \( \epsilon > 0 \) such that \( D(x_0, 2\epsilon) \) the disc of radius \( 2\epsilon \) with center \( x_0 \), lies in \( V \). Define \( \delta > 0 \) by

\[
    e^{\tau \delta} < \epsilon/2, \quad \delta < \epsilon
\]  

(6)

and (by denseness) choose \( y_0 \in V_1 \) with \( d(y_0, x_0) < \delta/2 \). By assumption (ii), there is a constant \( C \) such that

\[
    d(K(t)y_0, y_0) \leq Ct, \quad t \in [0, T].
\]  

(7)

Define \( \tau \) by

\[
    Ce^{\tau \delta} < \epsilon/2
\]  

(8)

we assert, inductively, that \( K(t/n)^n y_0 \in D(y_0, \epsilon/2) \). Indeed this follows from (7), (8) for \( n = 1 \) and the following:

\[
    d(K(t/n)^n y_0, y_0) \leq \sum_{j=1}^{n} d(K(t/n)^j y_0, K(t/n)^{j-1} y_0)
\]

\[
    \leq \sum_{j=1}^{n} e^{\delta(j-1)t/n} d(K(t/n)y_0, y_0)
\]

\[
    \leq \sum_{j=1}^{n} e^{\delta t} Ct/n = Cte^{\delta t} < \epsilon/2.
\]
Let \( W = D(x_0, \delta/2) \). For \( x \in W \),
\[
d(x, y) \leq d(x, x_0) + d(x_0, y_0) < \delta,
\]
so
\[
d(K(t/n)^n x, K(t/n)^n y_0) \leq (e^{\beta t/n})^n \cdot d(y, y_0)
\]
\[
\leq e^{\beta t} \delta < \epsilon/2
\]
by (6). Thus, inductively
\[
K(t/n)^n x \in D(K(t/n)^n y_0, \epsilon/2) \subset D(y_0, \epsilon) \subset D(x_0, 2\epsilon)
\]
which proves our assertion.

Let \( W_k = W \cap V_k \), a neighborhood of \( x_0 \) in \( M_k \).

**Lemma 2.3.** Let \( B \subset W_{k+2} \) be a bounded set. Then there is a constant \( C_1 \) such that
\[
d_k(K(t) x, K(t/l)^l x) \leq C_1 t^2
\]
for \( x \in B \), \( 0 \leq t \leq \tau \), \( l = 1, 2, 3, \ldots \).

**Proof.** By (iv),
\[
d_k(K(t+s) x, K(t) K(s) x) \leq C_0 t s
\]
for all \( x \in B \), \( t, s \in [0, \tau] \). Thus using the triangle inequality
\[
d_k(K(t) x, K(t/l)^l x)
\]
\[
\leq \sum_{j=0}^{l-1} d_k(K(t/l)^j K(t - jt/l) x, K(t/l)^j K(t/l) K(t - (j + 1) t/l) x)
\]
\[
\leq \sum_{j=0}^{l-1} e^{\beta t/l} d_k(K(t - jt/l) x, K(t/l) K(t - (j + 1) t/l) x) \quad \text{by (iii)}
\]
\[
\leq \sum_{j=0}^{l-1} e^{\beta t} C_0(t/l) (t - (j + 1) t/l) \quad \text{by (iv)}
\]
\[
\leq C_0 e^{\beta t} (t^2/2(l - (l + 1)/2) \leq C_1 t^2.
\]

In particular, we may conclude from (9) and compactness of \( \{K(t) x: 0 \leq t \leq \tau\} \) for fixed \( x \), that \( \{K(t/l)^l x: 0 \leq t \leq \tau, \, l = 1, 2, 3, \ldots \} \) is a bounded set in \( V_{k+2} \) if \( x \in W_{k+4} \). (The bound \( C_1 \) of course depends on \( B \) and \( k \)). Thus we obtain the following.

**Corollary 2.4.** Let \( x \in W_{k+4} \). Then there is a constant \( C_1 \) such that
\[
d_k(K(t) y, K(t/l)^l y) \leq C_1 t^2
\]
for all \( 0 \leq t \leq \tau \), \( y = K(t/j)^j x \), \( 0 \leq t_1 \leq \tau_1 \), \( j = 1, 2, 3, \ldots \).
We also assert that:

**Corollary 2.4'.** The above result is valid if \( k + 4 \) is replaced by \( k + 2 \).

**Proof.** Let \( x \in W_{k+2}. \) We need only show that in \( W_{k+2}, \) \( \{K(t_{ij})^j x \mid 0 \leq t_j \leq \tau, i = 1, 2, 3, \ldots\} \) is a bounded set. Pick \( \bar{x} \in W_{k+4} \) so we have this assertion for \( \bar{x}. \) But then

\[
d_{k+2}(K(t_{ij})^j x, K(t_{ij})^j \bar{x}) \leq e^{\theta_{k+2}t_1} d_{k+2}(x, \bar{x})
\]

so that if \( K(t_{ij})^j \bar{x} \) lies in a ball of radius \( R, K(t_{ij})^j x \) will lie in a ball of radius \( R + e^{\theta_{k+2}t} d_{k+2}(x, \bar{x}). \)

Now we are ready to make the key estimate.

**Lemma 2.5.** Let \( x \in W_{k+2}. \) Then there is a constant \( C_2 \) such that for all \( m \geq n \) we have

\[
d_k(K(t/n)^n x, K(t/m)^m x) \leq C_2 t^2/n, \tag{11}
\]

for all \( 0 \leq t \leq \tau. \) The constant \( C_2 \) depends on \( x \) and \( k. \)

**Proof.** First, suppose \( m = nl. \) Then write, by the triangle inequality,

\[
d_k(K(t/n)^n x, K(t/nl)^{nl} x)
\]

\[
\leq \sum_{j=0}^{n-1} d_k(K(t/n)^{n-j} K(t/nl)^j x, K(t/n)^{n-j-1} K(t/nl)^{j+1} x)
\]

\[
\leq \sum_{j=0}^{n-1} e^{\theta_k t(n-j-1)/n} d_k(K(t/n) K(t/nl)^j x, K(t/nl)^j K(t/nl)^{j+1} x).
\]

We now employ Corollary 2.4' to obtain that the above is

\[
\leq e^{\theta \tau} C_1 t^2/n.
\]

To obtain the result for general \( m, \) write

\[
d_k(K(t/n)^n x, K(t/m)^m x)
\]

\[
\leq d_k(K(t/n)^n x, K(t/m)^{nm} x) + d_k(K(t/m)^m x, K(t/mn)^{mn} x)
\]

\[
\leq C_1 e^{\theta \tau} t^2((1/n) + (1/m)) \leq 2C_1 e^{\theta \tau} t^2/n.
\]

We get our result by taking \( C_2 = 2C_1 e^{\theta \tau}. \)
Observe that the constant $C_2$ depends on the point $x \in W_{k+2}$, but that $\tau > 0$ is uniform, independent of $x$ and $k$.

**Lemma 2.6.** $K(t/n)^n x$ converges in $V$ for each $x \in W$ uniformly in $t$. In particular, for $x \in W$

$$t \mapsto H_t x = \lim_{n \to \infty} K(t/n)^n x$$

is continuous and so $H_t x : [0, \tau] \times W \to W$ is jointly continuous.

**Proof.** By 2.5, $K(t/n)^n x$ converges uniformly if $x \in W_2$. By denseness, for $x \in W$, find $x_j \to x$, $x_j \in W_2$. Then the inequality

$$d(K(t/n)^n x, K(t/m)^m x) \leq e^{\epsilon t} d(x, x_j) + d(K(t/n)^n x_j, K(t/m)^m x_j) + e^{\epsilon t} d(x_j, x)$$

shows that $K(t/n)^n x$ converges uniformly in $t$ as well. We clearly have $d(H_t(x), H_t(y)) \leq e^{\epsilon t} d(x, y)$ so the rest of 2.6 holds.

Now the same argument shows that $K(t/n)^n x$ converges in $V_k$ if $x \in W_k$. Hence since the inclusions are continuous we can conclude that $H_t x \in V_k$ if $x \in W_k$. This establishes the regularity property.

**Remark.** 1. This proof depends crucially on the fact that the $\beta_k$ are uniform on all of $V_k$ even though for $k \geq 1$, $V_k$ will, in general, be an unbounded set. In hydrodynamics the $\beta_k$ will be bounded on bounded sets. However, without additional assumptions, the result does not seem to be true with this weakening. One only obtains for $x \in W_{k+2}$, an integral curve $H_t(x)$ lying in $V_k$. Thus we have a "loss of smoothness" which is an unfortunate property for dynamical systems. (The flow property (c) then does not make much sense.) We shall see in 2.10 how to rectify the situation.

**Remark 2.** The convergence in 2.6 is probably not locally uniform in $x$ (except from $W_{k+2}$ to $V_k$).

We continue the proof. Condition (a) being obvious, and (b) proven in 2.6, we turn to (c).

By joint continuity of $H_t x$, we can shrink $W$, $\tau$ down to $\hat{W}$, $\hat{\tau}$ such that

$$0 \leq t \leq \hat{\tau}, \quad x \in \hat{W} \Rightarrow H_t x \in W.$$

Let us still call the new neighborhood and time interval by $W$, $\tau$. This additional shrinking is probably superfluous, but it guarantees that $K(t/n)^n H_t x$ converges as $n \to \infty$ (by 2.6) and we use this fact in the following.
Lemma 2.7. \( H_t \) is a semigroup. That is, for \( x \in W, 0 \leq s + t \leq \tau, 0 \leq s, t \) we have
\[
H_{s+t}x = H_s(H_t x).
\]

Proof. First, suppose \( s \) and \( t \) are rationally related, say \( s = lt/m \). Then
\[
H_{t+s}x = \lim_{n \to \infty} K((s+t)/n)x
= \lim_{n \to \infty} K(s/k') K(t/k)x
\]
where \( k' = ln/(l + m), k = nm/(l + m) \) and the limit is taken through multiples of \( l + m \) so that \( k, k' \) are integers. Here \( k, k' \) are chosen so that
\[
s/k' = t/k = (s + t)/n \quad \text{and} \quad k + k' = n
\]
which makes the above equality clear. Now write
\[
d(K(s/k') K(t/k)x, H_s H_t(x))
\]
\[
\leq d(K(s/k') K(t/k)x, K(s/k') K(t/k)x) + d(K(s/k') K(t/k)x, H_s(H_t(x))
\]
\[
\leq e^{\theta} d(K(t/k)x, H_t(x)) + d(K(s/k') K(t/k)x, H_s(H_t(x))
\]
and observe that each of these terms \( \to 0 \) as \( k', k \to \infty \). Thus \( H_{t+s}x = H_s(H_t x) \) holds for a dense set of \( s, t \) and so by continuity in \( t \) (Lemma 2.6) it holds for all \( s, t \).

Lemma 2.8. Let \( x \in M_1 \). Then \( H_t x \in V \) is right differentiable on \([0, \tau)\), continuously differentiable on \((0, \tau)\) and we have
\[
H'_t x = K'(0)(H_t x), \quad 0 \leq t < \tau.
\]

Proof. Let \( x \in W_2 \). Then, by Lemma 2.3 and letting \( l \to \infty \), we have
\[
d(K(t) x, H_t x) = O(t^\theta).
\]
Thus \( H_t x \) is right differentiable at \( 0 \) with derivative \( K'(0) \). Now
\[
H_{t+s}x = H_s(H_t(x))
\]
and \( H_s(x) \in V_s \) by regularity. Thus \( H_t \) is right differentiable at each \( s \) with \( H'_s x = K'(0)(H_s x) \). But this right derivative is clearly continuous. Hence by standard arguments (Yosida [23], Kato [12]) \( H_t(x) \) is differentiable on \((0, \tau)\).
Hence we have, going to a coordinate chart,

\[ H_t(x) = x + \int_0^t K'(0)(H_s(x)) \, dt \]
on a dense set, namely for \( x \in V_2 \cap W_1 \). By continuity, this equation holds then for all \( x \in W_1 \) and the conclusion therefore follows. \( \blacksquare \)

To complete the proof of our theorem, we quote the following.

**Lemma 2.9.** Let \( X: D \subset M \to TM \) be a densely defined vector field and suppose \( X \) has a \( C^0 \) semigroup \( H_t \) locally defined on \( M \), leaving \( D \) invariant and \( H_t'(x) = X(H_t(x)) \) for \( x \in D \). Suppose (locally) \( d(H_t(x), H_t(y)) \leq Mt \, d(x, y) \) where \( M_t \) is locally bounded. Then integral curves of \( X \) are unique.

This is proved by standard techniques; see Chernoff–Marsden [5, 6].

Now we wish to show how the same results can be obtained under more specialized hypotheses. These can actually be verified for the Navier–Stokes equations in the context of [10] and Section 5 below.

**Theorem 2.10.** In Theorem 2.1, replace (iii) by the weaker assumption:

(iii)' for each bounded set \( B \subset V_{k-1} \), there is a constant \( \beta \), (depending on \( B, k \)) such that

\[ d_k(K(t)x, K(t)y) \leq e^{\beta t} d_k(x, y) \]

for \( x, y \in B \cap V_k \).

Then the same conclusions, including regularity, hold.

**Proof.** We proceed as in 2.1 with Lemma 2.2 choosing a (bounded) \( W \) and \( \tau \) so that \( K(t/n)^n \) \( x \) are all defined. The key thing is the following lemma.

**Lemma 2.11.** For \( x \in B_k \), a bounded set in \( W_k \), \( \{K(t/n)^n x: x \in B_k, 0 \leq t \leq \tau, n = 1, 2, 3, \ldots\} \) is a bounded set in \( V_k \).

**Proof.** We proceed by induction \( k \). It holds by construction for \( k = 1 \). Assume it is true for \( k - 1 \). Then by (iii)' and the argument in Lemma 2.3, we see that the estimate (9) remains valid. (All quantities lie in a bounded set of \( V_{k-1} \) by assumption.) Thus there is some particular \( y \in W_k \) so that \( K(t/n)^n y \) lies in a bounded set. Then as in 2.4' for any \( x \in B_k \) we have

\[ d_k(K(t/n)^n x, K(t/n)^n y) \leq e^{\beta t} d(x, y) \]
where $\beta$ is the constant from (iii)' for the bounded set $B_k \cup \{K(t/n)^n y\}$ in $V_{k-1}$. Thus if $K(t/n)^n y \in D(y, R)$ we have $K(t/n)^n x \in D(y, R + e^{at} d(x, y))$, so our assertion follows. 

Once we have observed this result, the rest of the argument goes through exactly as before.

Verifying (iii)' for hydrodynamics rests on a basic nonlinear estimate done in [10] or [14, part II, Section 3]. That statement also yields a direct proof of the regularity conclusion of the theorem for those equations.

Of course one could also make the more general assumptions:

(iii)" the $\beta_k$ of (iii) are assumed bounded on bounded sets of $V_k$ and for each $x \in V_k$, \{K(t/n)^n x: 0 \leq t \leq T, n = 1, 2, \ldots\} is a bounded set in $V_k$.

Then with this assumption replacing (iii) it is not difficult to see that the conclusions remain valid. The trouble is of course that in examples it is difficult to check directly the boundedness of $K(t/n)^n x$.

3. Smoothness of the Flow $H_t$

Sometimes one expects more than just continuity of $H_t: W \to V$, and it is important to have criteria for establishing this. The natural hypotheses are in terms of smoothness of the $K(t)$, for fixed $t$. We observe that if $H_t$ is $C^r$ for each fixed $t$, then derivatives up to order $r$ will automatically be strongly continuous functions of $t$. This is a general fact about semi-groups proved in Chernoff–Marsden [5].

It is an amusing fact that the Navier–Stokes equations have this smoothness property in Lagrangian coordinates, but not necessarily in Eulerian coordinates. The results in [10] and this section establish this fact.

Let $M \supset M_1 \supset \cdots$, etc. be as in 2.1. We shall need some additional structure on these spaces. Namely we assume that on $V$ we have a $C^0$ structure enabling us to parallel translate vectors at $x \in V$ to our reference point $x_0$, and have a norm on the fiber $T_{x_0}M$ (thus we inherit a $C^0$ Finsler structure on $V$). Thus we can subtract vectors at different points and we get a (locally) complete metric on $TV$ by setting

$$d(v_x, v_y) = d(x, y) + \|v_x - v_y\|.$$ 

For example we may suppose we are working in a chart, that $M$ is a Banach space to start with, or has some group structure or has a
Riemannian or Finsler structure admitting an exponential chart at $x_0$. We demand regularity in that this same operation should restrict to each $V_k$.

In the following we treat the $C^1$ case. The $C^r$ case is similar.

**Proposition 3.1.** Let the hypotheses of 2.1 (or 2.10) hold and let there be a metric on $TU$ as described above. Suppose each map $K(t)$ is $C^1$ so that

$$\hat{K}(t) \equiv TK(t) \colon TV \to TU.$$ 

Assume that the hypotheses of 2.1 hold for these maps $\hat{K}(t)$ with $V = TV$ etc.

Then for each $t$, $H_t$ is a $C^1$ map of $TW_k \to TM_k$, $k = 0, 1, 2, \ldots$. Moreover, $(t, v_x) \mapsto TH_t \cdot v_x$ is jointly continuous.

**Proof.** It suffices to prove the conclusion for $k = 0$. From the hypotheses and the proof of 2.1, we see that, with the same $\tau$, $W$ as in 2.1,

$$[TK(t/n)]^n v_x = T(K(t/n))^n \cdot v_x$$

converges uniformly in $t$, as $n \to \infty$ for each $v_x \in T_x W$. Call the limit $g_t(x) \cdot v$. We have

$$d(g_t(x) \cdot v_x, g_t(y) \cdot v_y) \leq e^{\alpha}[d(x, y) + ||v_x - v_y||]$$

so that $x \mapsto g_t(x)$ is continuous in operator norm.

Our result now is an immediate consequence of an elementary lemma, and our introductory remarks.

**Lemma 3.2.** Let $E$, $F$ be Banach spaces, $U \subset E$ open. Let $f_n : U \subset E \to F$ be a sequence of $C^1$ functions such that: $Df_n(x)$ is locally uniformly bounded,

$$f_n \to f \quad \text{pointwise on} \quad U,$$

$$Df_n \to g \quad \text{pointwise} \quad U \times E$$

and $x \mapsto g(x) \in L(E, F)$ is continuous (in norm topology). Then $f$ is $C^1$ and $Df = g$.

**Proof.** Write

$$f_n(x + v) - f_n(x) = \int_0^1 Df_n(x + sv) \cdot v \, ds.$$
Now by dominated convergence we may let \( n \to \infty \) and get

\[
f(x + v) - f(v) = \int_0^1 g(x + sv) \cdot v \, ds.
\]

The result is an easy consequence of this formula and norm continuity of \( g \).

4. Towards Verifying Hypothesis (iii)

We shall now give a nonlinear generalization of a remark apparently due to W. Feller and H. Trotter (see Hille-Phillips [11] or Trotter [22]). This remark is that if \( F_t \) is a linear \( C^0 \) semigroup on a Banach space \( \mathbb{E} \), then \( \mathbb{E} \) can be renormed in such a way that \( F_t \) is a quasi-contraction:

\[
\| F_t \cdot x - F_t \cdot y \| \leq e^{\delta t} \| x - y \|.
\]

For nonlinear semigroups this is not possible. In fact there is a large (and very important) gap between the developed theory of quasincontractive nonlinear semigroups and general semigroups which are, say, just locally lipschitzian for each \( t \).

We show that if one is willing to move out of the confines of Banach spaces to manifolds, then by a simple adaptation of the linear argument we can recover the above remark. The Banach space norm is replaced by a distance function associated to a certain Finsler (= norm on each tangent space) structure. This is evidently useful in view of Theorem 2.1 which is valid for these more general spaces.

The hope is that these more general ideas will enable one to deal with the semigroups not covered by the quasincontractive theory on Banach spaces (i.e., many interesting nonlinear partial differential equations). For the concrete type of differential equations to which the contractive theory applies, see for instance Browder [2] and [6].

**Theorem 4.1.** Let \( M \) be a Banach manifold admitting partitions of unity and \( \| \cdot \|_x \) a given Finsler structure on \( M \). Let \( F_t(x) \) be a jointly continuous, (perhaps locally defined) flow on \( M \); \( t \geq 0 \).

Assume that for each \( t \), \( x \mapsto F_t(x) \) is a \( C^1 \) mapping and there is a constant \( M_t \) such that

\[
\| TF_t \cdot v_x \|_{F_t(x)} \leq M_t \| v_x \|_x
\]

(12)
for $v_x \in T_xM$ and $t \mapsto M_t$ is locally bounded. Thus, if $\rho$ is the metric on $M$ corresponding to $\| \cdot \|_x (\rho(x, y) = \inf \{ \int_0^1 \| C'(t) \| dt : C(t) is a C^1 curve joining x, y \})$ it follows that

$$\rho(F_t(x), F_t(y)) \leq M_t \rho(x, y) \quad (13)$$

Then there are constants $\alpha, \beta$ such that

$$\rho(F_t(x), F_t(y)) \leq \alpha e^{\beta t} \rho(x, y), \quad t \geq 0 \quad (14)$$

and there is an equivalent Finsler structure $\| \cdot \|_x$ on $M$ with associated metric $d$ for which

$$d(F_t(x), F_t(y)) \leq e^{\beta t} d(x, y). \quad (15)$$

**Remark.** Under the assumption that $F_t(x)$ is $C^1$ for fixed $t$, assuming (12) is the same as its consequence (13). We have done this for simplicity. If $F_t$ is just assumed Lipschitzian the same theorem holds, with some additional effort, by replacing $C^1$ norms by Lipschitz norms where appropriate (see the remark of P. Chernoff below).

**Proof of 4.1.** That one can replace $M_t$ by $\alpha e^{\beta t}$ is a classical argument in the linear case (Yosida [23]). It was observed to carry over to nonlinear semi-groups in Banach spaces by several people, for example Phillips and Chernoff. Exactly the same argument may be used here in the metric space context, so we shall omit it (cf. Chernoff–Marsden [6] and Crandall–Pazy [8]).

So we turn our attention to the proof of (15). We first remark that our local flows can be converted to globally defined ones without changing them on a given neighborhood $V_0$, where $V_0 \subset V_0 \subset V$. Namely, find a smooth function $f: M \rightarrow \mathbb{R}$ equaling 1 on $V_0$ and 0 on $M \setminus V$. Then setting

$$G_t x = F_{\int_0^t f(F_s(x)) ds} x,$$

$G$ clearly extends to a globally defined flow equaling $F_t$ on $V_0$, $0 \leq t \leq \tau$. Thus in what follows we may assume we are dealing with global flows.

The new Finsler structure is simply given by

$$\| v_x \|_x = \sup_{t \geq 0} \| e^{-\beta t} T F_t \cdot v_x \|_{F_t(x)} \cdot \quad (16)$$

Observe that formula (16) reduces to the linear renorming if $F_t$ is linear (Trotter [22]).
From (12) we have
\[ \| v_x \|_x \leq \alpha \| v_x \|_x \]
and clearly (take \( t = 0 \)),
\[ \| v_x \|_x \leq \| v_x \|_x \]
so we have an equivalent Finsler structure.

To demonstrate (15), let \( c(s) \) be a curve joining \( x \) and \( y \). Then \( F_t \circ c \) is a curve joining \( F_t(x), F_t(y) \); we have
\[
d(F_t(x), F_t(y)) \leq l(F_t \circ c) = \int_0^1 \| \frac{d}{ds} F_t(c(s)) \|_{F_t(c(s))} \, ds
\]
\[
= \int_0^1 \sup_{\tau \geq 0} \| e^{-\tau F_t(c(s))} \cdot [TF_t(c(s)) \cdot c'(s)] \| \, ds
\]
\[
= \int_0^1 \sup_{\tau \geq 0} \| e^{-\tau TF_{t+\tau}(c(s))} \cdot c'(s) \| \, ds
\]
by the Chain rule and the fact that \( F_t \) is a semigroup. The above is, taking \( \sigma = \tau + t \),
\[
\leq \int_0^1 \sup_{\sigma \geq t} \| e^{\sigma t} TF_{\sigma}(c(s)) \cdot c'(s) \| \, ds
\]
\[
= e^{\beta t} l(c) \quad \text{(length in the \( \| \cdot \| \) structure)}.
\]

Taking the inf. over all such \( c \) gives
\[
d(F_t(x), F_t(y)) \leq e^{\beta t} d(x, y).
\]

Chernoff has pointed out that estimates (14) and (15) also follow from the linear theorem and the application of his general "linearizing functor" (Chernoff [4]). However we shall also require some more detailed smoothness properties of the Finsler structure (16) (see 4.2 below).

Note that for finite \( t \)-intervals it is trivial that one can replace \( M_t \) by \( \alpha e^{\beta t} \), merely by choosing \( \alpha \) large and \( \beta = 0 \).

In what follows, we shall require some smoothness properties of our Finsler structure and we want to make sure \( \| \cdot \| \) inherits this property. We do this as follows.
Proposition 4.2. Let, in Theorem 4.1, \( V \) be a coordinate chart for \( M \) and suppose \( F_t : W \to V \) is defined for \( W \subset V \), \( 0 \leq t \leq \tau \). Assume there is a constant \( C \) such that
\[
\| \nu \|_x \leq C \| \nu \|_v \quad \text{and} \quad \| \nu \|_x - \| \nu \|_v \| \leq Cd(x, y) \| \nu \|_x
\]  \( (17) \)
for all \( x, y \in V \) and \( \nu \) in the coordinate linear space \( \mathcal{E} \). Assume that \( F_t \) satisfies the conditions in 4.1; In addition, assume the derivative of \( F_t \) for each \( t \) is Lipschitz
\[
\| DF_t(x) \cdot \nu - DF_t(y) \cdot \nu \| \leq \text{const} \| \nu \| d(x, y)
\]
for all \( x, y \in W, \nu \in \mathcal{E}, 0 \leq t \leq \tau \).

Then the new Finsler structure \( \| \cdot \| \) defined on \( W \) by Theorem 4.1 also satisfies the condition (17) (with possibly a larger constant).

Proof. We have from (16),
\[
\| \| \nu \|_x - \| \nu \|_v \| \\
= \sup_{t \geq 0} \| e^{-tG}DF_t(x) \cdot \nu \|_F \|x(x) - \sup_{t \geq 0} \| e^{-tG}DF_t(y) \cdot \nu \|_F \|y(y)\| \\
\leq \sup_{t \geq 0} \{ \| e^{-tG}DF_t(x) \cdot \nu \|_F \|x(x) - \| e^{-tG}DF_t(y) \cdot \nu \|_F \|y(y)\| \} \\
\leq \sup_{t \geq 0} \{ \| e^{-tG}DF_t(x) \cdot \nu \|_F \|x(x) - e^{-tG} \| DF_t(x) \cdot \nu \|_F \|x(x)\| \} \\
+ e^{-tG} \| DF_t(x) \cdot \nu - DF_t(y) \cdot \nu - DF_t(y) \cdot \nu \|_F \|y(y)\| \\
\leq \sup_{t \geq 0} \{ Ce^{-tG} \| DF_t(x) \cdot \nu \|_F \|d(F_t(x), F_t(y))\| + e^{-tG} \| \nu \| d(x, y)\} \\
\leq (\text{const}) \| \nu \| d(x, y). \]

5. Sufficient Conditions for the Convergence of Trotter's Formula (1)

We now consider some simple sufficient conditions which will enable us to derive formula (1) from (2) via Theorem 2.1. One of the chief problems is to verify the condition (iii) (resp. (iii)') where \( K(t) = F_t \circ G_t \). Basically, this can be done when both \( F_t \), \( G_t \) satisfy (iii) for the same \( d \). (Trotter [22] points out this same problem in the linear case). For contractive semigroups, this difficulty vanishes. The rest of the conditions seem to be reasonable in most concrete situations of interest.
The simplest case in which we can make $F_t$ and $G_t$ simultaneously quasicontractive is when the generator of one of them is "bounded" or more precisely is a smooth vector field.

First we list some general notation. By a densely defined vector field on $M$ we mean a map $X: D \rightarrow TM$ from a dense set $D \subset M$ such that $X(x) \in T_x M$, the tangent space at $x \in M$. By a (local) flow for $X$ we mean a $C^0$ semi-group, (perhaps local) $F_t: M \rightarrow M$ such that $F_0 = I_D$ and for $x \in D$,

$$\frac{dF_t(x)}{dt} = X(F_t(x)).$$

Sometimes it may be convenient to choose $D$ smaller than the domain of the complete generator of $F_t$, and so we do not make this assumption in our results (cf. Chernoff–Marsden [5, 6]).

**Theorem 5.1.** Let $M \supset M_1 \supset M_2 \cdots$, be Banach manifolds with continuous and dense inclusions. Let $|| \cdot ||_k$ be a Finsler structure on $M_k$ with the associated metric $d_k$ making $M_k$ a complete metric space. Let $x_0 \in M$, $V \subset U$ be neighborhoods of $x_0$ in $M$ and $V_k = V \cap M_k$, $U_k = U \cap M_k$ as in Theorem 2.1. Assume $U$ gives a local chart for $M$ and that this restricts to $U_k$ to give charts for $M_k$. Assume that the Finsler structures in these charts satisfy, for each fixed $k$, the boundedness and Lipschitz properties (17).

Let $X: M_1 \rightarrow TM$ be a given densely defined vector with a local flow $F_t: V \rightarrow U$. Assume $X: V_k \rightarrow TV_k$ and is of class $C^1$ with bounded derivative on bounded sets. Suppose for each $t$, $F_t$ is of class $C^0$, $V_k \rightarrow U_k$ with its first and second derivatives uniformly bounded on $V_k$ for $0 \leq t \leq T$.

Let $Y: M \rightarrow TM$ be a vector field on $M$ such that $Y: V_k \rightarrow TV_k$ is of class $C^0$ with $Y$ and its first derivative uniformly bounded on $V_k$, and second derivative bounded on bounded sets. Let $G_t$ be the local flow of $Y$.

Then $X + Y$ has a unique local flow $H_t$ which is Lipschitz for each $t$. Moreover $H_t$ maps $W_k = W \cap V_k$ to $V_k$, and we have

$$H_t x = \lim_{n \to \infty} (F_{t/n} \circ G_{t/n})^n x$$

uniformly in $t$ for each $x \in W_k$, $0 \leq t \leq \tau$.

If all the degrees of differentiability and the hypotheses on them are increased by one, then $H_t$ will for each $t$ be of class $C^1$ with a locally Lipschitz derivative, etc.
Proof. We shall verify the hypotheses of Theorem 2.1. Let 
\( K(t) = F_t \circ G_t \), (and choose \( V \) suitably small). It is clear that we 
have hypotheses (i), (ii) on \( K \). The main job will be to verify (iii), 
so let us first dispose with (iv). We consider, for \( x \in V_{k+2} \), the curves 
\[
\begin{align*}
f(t,s) &= K(t+s) \cdot x = F_{t+s} \circ G_{t+s}(x) \\
g(t,s) &= K(t) \cdot K(s) = F_t \circ G_t \circ F_s \circ G_s(x).
\end{align*}
\]
From our hypotheses we see that 
\[
\begin{align*}
f(0,0) &= g(0,0), & (\partial f/\partial t)(0,0) &= (\partial g/\partial t)(0,0) \\
(\partial f/\partial t)(0,0) &= (\partial g/\partial s)(0,0),
\end{align*}
\]
and that \( f, g \) are \( C^2 \) with bounded derivatives as \( x \) ranges over a 
bounded set. For example 
\[
\begin{align*}
\partial f/\partial t &= X(F_{t+s} \circ G_{t+s}(x)) + DF_{t+s} \cdot Y(G_{t+s}(x)) \\
\partial g/\partial t &= X(F_t \circ G_t \circ F_s \circ G_s(x)) + DF_t \cdot Y(G_t \circ F_s \circ G_s(x)).
\end{align*}
\]
From this and Taylors theorem we obtain 
\[
d(K(t+s) \cdot x, K(t) \cdot K(s) \cdot x) = O(ts),
\]
which is (iv). [This is basically an estimate on the commutator 
\([X, Y]\) (as a densely defined vector field)].

Now we verify (iii). For this purpose, we define a new Finsler 
structure \( \| \cdot \| \) as in Theorem 4.1. Thus 
\[
d(F_t x, F_t y) \leq e^{\beta t} d(x, y).
\]
(It suffices to take the case \( k = 0 \); the others are the same). We want 
to verify the same hypothesis on \( G_t \). This is where the result 4.2 
comes in. So we can, by 4.2 assume the estimates (17) on \( \| \cdot \| \).

Now \( G_t \) is a smooth flow jointly in \( t, x \). We have 
\[
(d/\partial t) DG_t(x) \cdot v = DY(G_t(x) \cdot v) \cdot DG_t(x) \cdot v
\]
so 
\[
DG_t(x) \cdot v = DG_t(x) \cdot v + \int_s^t DY(G_t(x) \cdot v) \cdot DG_t(x) \cdot v
\]
and thus
\begin{align}
\| DG_t(x) \cdot v \|_{G_t(x)} - \| DG_s(x) \cdot v \|_{G_s(x)} \leq & \int_s^t \| DY(G_o(x) \cdot v) \|_{G_{t}(x)} \cdot \| DG_o(x) \cdot v \|_{G_{s}(x)} \, d\sigma.
\end{align}
(18)

Now
\begin{align}
\| DG_t(x) \cdot v \|_{G_t(x)} - \| DG_s(x) \cdot v \|_{G_s(x)} \leq & (\text{const}) \| DG_t(x) \cdot v \|_{G_s(x)} \cdot d(G_t(x), G_s(x)) \\
\leq & (\text{const}) \| DG_t(x) \cdot v \|_{G_s(x)} \int_s^t \| Y(G_o(x)) \|_{G_s(x)} \, d\sigma \\
\leq & | s - t | (\text{const}) \| DG_t(x) \cdot v \|_{G_s(x)},
\end{align}
by (17).

Thus, writing out a telescoping sum and employing the above, we get
\begin{align}
\| DG_t(x) \cdot v \|_{G_t(x)} - \| v \|_x \leq & \text{const} \cdot \int_0^t \| DG_o(x) \cdot v \|_{G_o(x)} \, d\sigma
\end{align}
and so
\begin{align}
\| DG_t(x) \cdot v \|_{G_t(x)} \leq \exp(\text{const} \cdot t) \| v \|_x
\end{align}
(Gronwall’s inequality). From this, it follows that
\begin{align}
d(G_t(x), G_t(y)) \leq \exp(\text{const} \cdot t) \, d(x, y)
\end{align}
as in 4.1.

Thus, if we let \( \beta \) be the sum of \( \beta_t \) and this constant, we get
\begin{align}
d(F_t \circ G_t, F_t \circ G_t) \leq e^{\beta t} \, d(x, y),
\end{align}
which is our condition (iii).

The last statement of the theorem follows from 3.1.

**Corollary 5.2.** Suppose the quantities \( DF_t, D^2 F_t, Y, DY \) which in 5.1 were assumed bounded on \( V_k \), are merely bounded on sets of the form \( B \cap V_k \) where \( B \subset V_{k-1} \) is bounded. Then the same conclusions are true.

**Proof.** Use 2.10 instead of 2.1.
6. Singular Perturbations

Typically, $X$ and $Y$ in Section 5 represent nonlinear differential operators. They may be of different orders. Thus it is not a priori obvious that the limit of the flow of $\nu X + Y$ as $\nu \to 0$ has anything to do with the flow of $Y$. In [9, 14] we verified that one does have the correct limit in the case of the hydrodynamic equations when no boundaries are present (see [10(II)] for the case of boundaries). This result is also due (independently) to Swann [20] and Kato [13].

One more observation: even though $X$ and $Y$ may have different orders, we can let them have the same domain by using the domain of the one with the highest order. In the context of 5.1 there is no problem as $Y$ has “order zero.”

Now we wish to abstract this situation.

**Theorem 6.1.** Let $K^\nu(t)$ be a family of maps for $\nu \in [0, A]$ each of which satisfies the hypotheses of 2.1 (or 2.10). Moreover, assume $V$, $U$, $T$ are independent of $\nu$ and

(i) \( K^\nu(t) x \to K(t) x \) uniformly in $t$ for each $x \in V_k$,

(ii) for $y \in V_{k+1}$, $(d/dt) K^\nu(t)$, is uniformly bounded in $V_k$, \( \nu \in [0, A], \ t \in [0, T] \),

(iii) the constants $\beta_k$ in 2.1(iii) (or 2.10(iii)') are uniformly bounded for $\nu \in [0, A]$,

(iv) the constant $C_0$ is independent of $\nu \in [0, A]$ in 2.1(iv).

Then $W$ and $\tau$ may be chosen so that each of the flows $H_\tau^\nu x$ are defined on $W$ (that is, $W, \tau$ are independent of $\nu$).

Furthermore, for each $x \in W_k$, $H_\tau^\nu x \to H_\tau x$ in $V_k$ as $\nu \to 0$ uniformly in $t$.

If \( d_k(K^\nu(t) x, K(t) x) = O(\nu) \) for $x \in W_{k+2}$, then the same thing is true for \( d_k(H_\tau^\nu x, H_\tau x) \).

**Note.** Because of assumption (iii), it suffices to check (i) on a dense set.

**Proof.** The various constants constructed in the proof of 2.1 are independent of $\nu$ so we have that $W, \tau$ are independent of $\nu$. For $x \in W_{k+2}$ we see from the estimate (11) that $K^\nu(t/n)^n x \to H_\tau^\nu x$ in $V_k$ as $n \to \infty$ uniformly in $\nu, t$. Thus it follows that $H_\tau^\nu x \to H_\tau x$ as $\nu \to 0$ (write

\[
d_k(H_\tau^\nu x, H_\tau x) \leq d_k(H_\tau^\nu(x), K^\nu(t/n)^n x) + d_k(K^\nu(t/n)^n x, K(t/n)^n x)
\]

\[
+ d_k(K(t/n)^n x, H_\tau x).
\]
Finally if \( x \in W_k \), let \( x_j \to x \), \( x_j \in W_{k+j} \). Then writing
\[
d_k(Hx, x) = d_k(Hx, x) + d_k(Hx, x) + d_k(Hx, x)
\leq 2d_k(x, x) + d_k(Hx, x)
\]
we see that \( Hx \to Hx \) as \( \nu \to 0 \) for \( x \in W_k \). The last statement of the theorem also follows from this proof. 

In particular, if \( X \) and \( Y \) are as in Theorem 5.1, then the hypotheses are satisfied for
\[
K_\nu(t) = F_\nu \circ G_t
\]
so that the generator is \( \nu X + Y \). Here \( K_\nu(t) \to G_t \) in \( W_k \) as \( \nu \to 0 \)
and the difference is \( O(\nu) \) on the spaces \( W_{k+j}, j \geq 1 \).

REFERENCES

1. H. BREZIS AND A. PAZY, Semigroups of Nonlinear Contractions on Convex Sets, 
   Analysis* 2 (1968), 238–242.
4. P. CHERNOFF, Semigroup product formulas and the addition of unbounded 
5. P. CHERNOFF AND J. MARSDEN, On Continuity and Smoothness of Group Actions, 
   (in preparation).
7. A. CHORIN, On the Convergence of Discrete Approximations to the Navier– 
9. M. CRANDALL AND A. PAZY, Semigroups of nonlinear Contractions and Dissipative 
10. D. EBIN AND J. MARSDEN, Groups of Diffeomorphisms and the Motion of an 
13. T. KATO, Nonstationary Flows of Viscous and Ideal Fluids in \( \mathbb{R}^3 \), *J. Functional 
    Analysis* 9 (1972), 296–305.
14. J. MARSDEN, D. EBIN, AND A. FISHER, "Diffeomorphism Groups, Hydrodynamics, 