The Annals of Probability
1984, Vol. 12, No. 2, 480–513

A CENTRAL LIMIT PROBLEM IN RANDOM EVOLUTIONS[1]

BY JOSEPH C. WATKINS

University of British Columbia

Let $\{T_n\}_{n\geq 1}$ be a sequence of independent and identically distributed strongly continuous semigroups on a separable Banach space. The corresponding generators $\{A_n\}_{n\geq 1}$ satisfy $E[A_n] = 0$. Conditions are given to guarantee that the weak limit $Y(t) = \text{limit}_{n\to\infty} \prod_{i=1}^{[n^2t]} T_i(1/n) Y_n(0)$ exists, and is characterized as the unique solution of a martingale problem. Transport phenomena, random classical mechanics, and families of bounded operators are the featured examples.

1. Introduction. Griego and Hersh [7] were the first to employ the term random evolution. Their ingredients were a stationary Markov chain $v(s)$ with state space $\{1, 2, \cdots, n\}$ and a set of infinitesimal generators $\{A^i\}$. In their work a random evolution is defined to be a product

$$(1.1) \qquad M(t) = \exp(t - \tau_N)A^{v(\tau_N)} \cdots \exp(\tau_2 - \tau_1)A^{v(\tau_1)}\exp\tau_1 A^{v(0)}$$

where τ_j is the time of the jth jump, and $N(t, \omega)$ is the number of jumps up to time t. They suggested this probabilistic tool in order to obtain existence, representation, and asymptotic theorems and formulas for initial value partial differential equations of both hyperbolic and parabolic type. Asymptotic theorems in random evolutions are ideally suited to modeling systems having rapid, but small fluctuations. These theorems separate into two types. First order theorems, generalizations of the law of large numbers, state that if the evolution coefficients are multiplied by $1/n$ and if the process is sped up by dividing the time parameter by n, then as $n \to \infty$, the average of the evolution is propagated by $\exp t\bar{A}$, where $\bar{A}$ is the closure of the average of the generators of the various modes of evolution. If $\bar{A} = 0$, then the first order limit is the identity. In this case if we keep the fluctuations small, multiplying the evolution coefficient again by $1/n$, then we must speed up the process even more, dividing the time parameter by n^2. A second order limit theorem proved in this manner generalizes the central limit theorem. This paper establishes second order theorems for random evolutions occurring in discrete time according to independent factors.

Section 6 details three examples which are typical areas of investigation for limit theorems in random evolutions.

1. *Bounded operators*: B is an arbitrary separable Banach space and $A(\xi)$ is a bounded operator.

Received August 1982; revised October 1983.

[1] This paper forms a part of the author's dissertation. I would like to take this opportunity to thank Tom Kurtz and the many others in the Department of Mathematics at the University of Wisconsin who were so helpful during the course of this work.

AMS 1980 *subject classifications.* Primary 60F17; secondary 60B10, 60B12, 60F05, 60G44.

Key words and phrases. Central limit problem, random evolution, weak convergence, martingale problem, duality.

2. *Transport phenomena*: $B = C_0(\mathbb{R}^d)$ and $A(\xi) = a(\xi)\cdot\nabla$ where a is a measurable mapping from Ξ to functions on $\mathbb{R}^d$.

3. *Random classical mechanics*: $B = L^2(\mathbb{R}^{6N})$ and $A(\xi)f = \{f, H(\xi)\}$ where H is a measurable mapping from Ξ to a family of N particle Hamiltonians.

The case of bounded operators on a separable Banach space B is not motivated by any specific example, but is introduced to highlight the structures in this paper free from the additional demands that unbounded operators incur. The bounded operator case provides a clean theory, relatively free from additional hypotheses. Besides the assumption that the $A(\xi)$ are uniformly bounded, we assume the existence of a Hilbert space H that compactly imbeds in B with the $A(\xi)$ remaining uniformly bounded. As we shall see, if B is itself a Hilbert space, the assumption of an additional Hilbert space is unnecessary.

The limit theorems for bounded operators have received attention for the most part on finite dimensional Banach spaces. The initial results for products of random matrices were set by Kesten and Furstenberg [10] as first order limit theorems. More recently, Marc Berger [3] has obtained central limit theorems for random matrices.

Random classical mechanics arises from the Hilbert space approach to classical mechanics [14, page 313]. A brief outline of this approach follows.

Starting from a Hamiltonian

$$H(q_1, \cdots, q_N; p_1, \cdots, p_N) \tag{1.2}$$

we obtain Hamilton's equations of motion

$$\frac{\partial p_i}{\partial t} = \frac{-\partial H}{\partial q_i}, \quad \frac{\partial q_i}{\partial t} = \frac{\partial H}{\partial p_i}. \tag{1.3}$$

Let $\omega(p_0, q_0; t) = (q(t), p(t))$ denote the vector in $\mathbb{R}^{6N}$ that is the solution to equation 1.3 with initial data $q(0) = q_0$ and $p(0) = p_0$. For $f \in C_0^\infty(\mathbb{R}^{6N})$ let

$$(T(t)f)(p, q) = f(\omega(p, q; t)). \tag{1.4}$$

Then $T(t)$ defines a group of operators. By the chain rule,

$$\begin{aligned} Af = \frac{d}{dt} T(t)f\,|_{t=0} &= \textstyle\sum_{i=1}^{3N} \frac{\partial f}{\partial q_i}\frac{\partial q_i}{\partial t} + \frac{\partial f}{\partial p_i}\frac{\partial p_i}{\partial t} \\ &= \textstyle\sum_{i=1}^{3N} \frac{\partial f}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial H}{\partial q_i} = \{f, H\}, \end{aligned} \tag{1.5}$$

the Poisson bracket. Whenever H is a C^1 function, the Liouville operator A is defined on the core $C_0^\infty(\mathbb{R}^{6N})$ by equation (1.5). The assertion that $T(t)$ is a unitary group in $L^2(\mathbb{R}^{6N})$ is known as Liouville's theorem. Thus a random classical mechanics can result from the addition of a stochastic parameter to H. A becomes a random family of first order differential operators on the phase space of the N particles.

The following two physical models illustrate transport phenomena.

(i) A particle moves with a certain velocity along an integral curve until it collides with another particle; then it changes its speed and moves along another curve. Both the change in speed and in direction are assumed to be random.

(ii) A paramecium twiddles, that is, it tumbles in place sampling its environment. It comes out of its tumble in some direction, assumed to be random, and runs for a certain distance based upon its conclusions about the environment. Each of these events is taken to be independent and is represented by one of the operators $A(\xi)$.

The majority of the effort in the research on random evolutions has been directed to transport phenomena. The above cited Reuben Hersh review paper [8] along with the conference proceedings in [13] provide an excellent overview to this area. Indeed, Mark Pinsky [13] formulates his questions in terms of a martingale problem. This is the point of view taken in this article.

Random evolutions have been used primarily as a probabilistic method in differential equations. As a consequence, the limit theorems have been restricted in scope to showing that the average behavior is propagated by $S(t) = \exp(t\overline{C})$, where C is one half of some averaging of the generators of the possible modes of evolution. As an example of a second order theorem, we consider the Hersh-Pinsky [9] paper on particle transport. In their example, the transition mechanism was determined by a finite state ergodic Markov chain. $A(v(t))$ is a random first order differential operator and $1/n$ is the time scale for the mean free path between collisions. The averaging is given by

$$C = \text{limit}_{t\to\infty} \frac{1}{t} \int_0^t \int_0^s EA(v(s))A(v(r))\, dr\, ds, \tag{1.6}$$

a limit that was presumed to exist. C is a second order elliptic differential operator. If its closure is assumed to be a generator, then they were able to conclude that the limit of the average of the random semigroups is $S(t)$.

These second order theorems place the statement that transport phenomena have a diffusion limit on solid ground. The goal of the authors was to find a comprehensive theory that included noncommutating families of unbounded operators, while leaving the stochastic structure that determined the random variations in the equation of state as general as possible. This method was quite successful in representing $S(t)$ for a wide variety of structures with independence, finite Markov chains, or some mixing condition controlling the mechanism that picks the successive modes of evolution.

2. The assumptions, the results, and the strategy.

First we state the general set up: A is a family of generators of strongly continuous semigroups on a separable Banach space B, indexed by a probability space $(\Xi, \mathscr{X}, \mu)$, i.e., A: $(\Xi, \mathscr{X}, \mu) \to \mathscr{G}(B)$. In addition,

(i) $\mathscr{D} = \bigcap_{\xi\in\Xi} \mathscr{D}(A^2(\xi))$ is dense in B,

(2.1) and $\bigcap \{\exp sA(\xi)(\mathscr{D})\colon s > 0, \xi \in \Xi\} \subseteq \mathscr{D}$.

(2.2) (ii) $\xi \to A(\xi)y$ is measurable for each $y \in \mathscr{D}$.

(2.3) (iii) $E\|Ay\| = \int_\Xi \|A(\xi)y\| \mu(d\xi) < \infty$, and

(2.4) (iv) $E\|A^2y\| = \int_\Xi \|A^2(\xi)y\| \mu(d\xi) < \infty$ for each $y \in \mathscr{D}$.

(2.5) (v) $\|\exp sA(\xi)\| \leq M \exp \gamma s$ for some $M, \gamma \in \mathbb{R}$.

(2.6) (vi) $EAy = \int_\Xi A(\xi)y \, \mu(d\xi) = 0$.

Statements (2.3) and (2.4) are the moment conditions on $\{A(\xi)\}$.

On the dual space B^*, we assume the existence of a separable subspace $\mathscr{D}'$ with the following properties:

(2.7) (i) $\|y\| = \sup\{(\theta, y): \theta \in \mathscr{D}', \|\theta\| = 1\}$

(2.8) (ii) $\cap_{\xi\in\Xi} \overline{\mathscr{D}(A^{*2}(\xi))} \subseteq \mathscr{D}'$, $\cap \{\exp sA^*(\xi)(\mathscr{D}'): s > 0, \xi \in \Xi\} \subseteq \mathscr{D}'$

(2.9) (iii) $E\|A^*\theta\| = \int_\Xi \|A^*(\xi)\theta\| \mu(d\xi) < \infty$, and

(2.10) (iv) $E\|A^{*2}\theta\| = \int_\Xi \|A^{*2}(\xi)\theta\| \mu(d\xi) < \infty$ for each $\theta \in \mathscr{D}'$.

Condition (2.7) is a mild condition so that we can determine the norm of an element from information about its pairing with elements in the dual space. This condition will be essential in the proof of Theorem 3.1. Conditions (2.8), (2.9), and (2.10) mirror conditions (2.1), (2.3), and (2.4), and are designed to permit us to construct a process on B^* similar to Y, the limiting random evolution. This dual process is the key to the uniqueness of Y.

Let $A, A_1, A_2, \cdots$ be a sequence of independent, identically distributed generators of strongly continuous semigroups as described above. The purpose of this work is to investigate the limiting behavior of the random evolutions

$$Y_n(t) = \exp\left(\frac{1}{n} A_{[n^2t]}\right) \cdots \exp\left(\frac{1}{n} A_2\right)\exp\left(\frac{1}{n} A_1\right) Y_n(0), \tag{2.11}$$

where $Y_n(0)$ converges weakly to $Y(0)$.

In order to become more comfortable with the notation, we present an example. For each i, let

$$A_i = \begin{cases} \dfrac{d}{dx} \text{ with probability } \dfrac{1}{2} \\ -\dfrac{d}{dx} \text{ with probability } \dfrac{1}{2}. \end{cases}$$

In other words, $A_i = r_i(d/dx)$ where $\{r_i\}$ is an independent sequence of Rademacher random variables. Then $\exp(sA_i)$ is a semigroup of uniform motion either to the left or to the right, depending upon the coefficient of d/dx. Let $Y_n(0) = f$ for all $n, f \in C_0(R)$. Then

$$\left[\exp\left(\frac{1}{n} A_1\right) Y_n(0)\right](x) = f\left(x + \frac{1}{n} r_1\right)$$

$$\left[\exp\left(\frac{1}{n} A_2\right)\exp\left(\frac{1}{n} A_1\right) Y_n(0)\right](x) = f\left(x + \frac{1}{n}(r_1 + r_2)\right).$$

Continuing in this manner we find that $Y_n(t)$ is given by

$$\prod_{i=1}^{[n^2t]} \exp\left(\frac{1}{n} A_i\right) Y_n(0)(x) = f\left(x + \frac{1}{n} \sum_{i=1}^{[n^2t]} r_i\right). \tag{2.12}$$

As $n \to \infty$, $(1/n) \sum_{i=1}^{[n^2t]} r_i$ converges weakly to standard Brownian motion $B(t)$. The limit process $Y(t)$ is

$$Y(t)(x) = f(x + B(t)). \tag{2.13}$$

This will be a good example to keep in mind throughout. The essential features to notice are:

(i) Each of the random evolutions Y_n has its sample paths in a Banach space.
(ii) The sample paths for Y_n are right continuous with left limits.
(iii) The limiting random evolution Y is the weak limit of the Y_n, and Y has continuous sample paths.

Thus, we are working with the weak convergence of a sequence of processes in $D_B[0, \infty)$ to a process in $C_B[0, \infty)$. This paper has two major results.

THEOREM 4.1. *Given the random evolutions* $\{Y_n\}$ *satisfying*

(i) *conditions* (2.1)–(2.8).
(ii) *the compact containment criterion* (3.1),
(iii) $Y_n(0)$ *converges weakly to* $Y(0)$.

Then $\{Y_n\}$ *is relatively compact in the topology of weak convergence. Furthermore, any limit process* Y *lies in* $C_B[0, \infty)$ *and for each* $\theta \in \mathscr{D}'$ *(defined in* (2.7) *and* (2.8)).

$$M^\theta(t) = \left(\theta,\ Y(t) - Y(0) - \frac{1}{2}\int_0^t \int_\Xi A^2(\xi) Y(s)\mu(d\xi)\, ds\right) \tag{4.1}$$

is a martingale, and $M^\theta(t)^2 - V^\theta(t)$ *is a martingale where*

$$V^\theta(t) = \int_0^t \int_\Xi (\theta, A(\xi) Y(s))^2 \mu(d\xi)\, ds. \tag{4.2}$$

THEOREM 5.6. *Under the hypothesis* (2.1)–(2.10), *a (dual) process* Θ *exists (on* $D_{B^*}[0, \infty)$) *for each starting point* $\theta \in \mathscr{D}'$ *and satisfies* (5.22) *and* (5.23) *(requirements much like* (4.1) *and* (4.2), *and hence the process* Y *is unique in law.*

The techniques involved in the proofs are a combination of facts concerning martingales along with the theory of weak convergence in Banach space. Some general results in this direction are set in Section 3 and are applied in Section 4 to prove Theorem 4.1. Section 5 uses a duality argument to prove the uniqueness theorem stated above. These methods apply to justify the following martingale difference schemes.

Returning to the process $Y_n(t)$, we see that

$$\left(\theta,\ Y_n\left(\frac{i}{n^2}\right) - Y_n\left(\frac{i-1}{n^2}\right) - E\left(\left(\exp\frac{1}{n}A\right) - I\right)Y_n\left(\frac{i-1}{n^2}\right)\right) \tag{2.14}$$

has zero conditional expectation with respect to $\mathscr{F}_{(i-1)/n^2} = \sigma(A_1, \cdots, A_{i-1})$, the sigma algebra of information up to time $(i-1)/n^2$. Therefore,

$$M_n^\theta(t) = \left(\theta,\ Y_n(t) - Y_n(0) - \textstyle\sum_{i=1}^{[n^2t]} E\left(\exp\left(\frac{1}{n}A\right) - I\right)Y_n\left(\frac{i-1}{n^2}\right)\right) \tag{2.15}$$

is a martingale for every $\theta \in \mathscr{D}'$. By Taylor's theorem recalling that $EA = 0$, we find that,

$$E\left(\exp\frac{1}{n}A - I\right) = \frac{1}{2n^2}EA^2 + O\left(\frac{1}{n^3}\right). \tag{2.16}$$

Placing this into equation (2.15), we have

$$M_n^\theta(t) = \left(\theta,\ Y_n(t) - Y_n(0) - \frac{1}{2}\textstyle\sum_{i=1}^{[n^2t]} EA^2Y_n\left(\frac{i-1}{n^2}\right)\frac{1}{n^2}\right) + O\left(\frac{1}{n}\right) \tag{2.17}$$

so if Y_n converges to some process Y, and if the Riemann sum converges to the appropriate integral

$$M^\theta(t) = \left(\theta,\ Y(t) - Y(0) - \frac{1}{2}\int_0^t \int_\Xi A^2(\xi)Y(s)\mu(d\xi)\,ds\right) \tag{2.18}$$

is a martingale. For continuous processes, the quadratic variation process is necessary to describe the process Y. The formal calculations proceed, guided by the construction outlined in the proof of the Doob-Meyer decomposition for discrete parameter martingales, in order to determine what this process ought to be. The procedure is straightforward and amounts to creating a process $V_n^\theta(t)$ by summing the increments necessary to center $M_n^\theta(t)^2$. In other words, we are finding a process $V_n^\theta(t)$ so that $M_n^\theta(t)^2 - V_n^\theta(t)$ is a martingale. At the kth stage, we add

$$
\begin{aligned}
&= E\left\{\left(M_n^\theta\left(\frac{k+1}{n^2}\right) - M_n^\theta\left(\frac{k}{n^2}\right)\right)^2 \,\Big|\, \mathscr{F}_{k/n^2}\right\} \\
&= E\left\{\left(\theta,\, Y_n\left(\frac{k+1}{n^2}\right) - Y_n\left(\frac{k}{n^2}\right) - E\left(\left(\exp\frac{1}{n}A\right) - I\right)Y_n\left(\frac{k}{n^2}\right)\right)^2 \,\Big|\, \mathscr{F}_{k/n^2}\right\} \\
&= E\left\{\left(\theta,\, Y_n\left(\frac{k+1}{n^2}\right) - E\left(\exp\frac{1}{n}A\right)Y_n\left(\frac{k}{n^2}\right)\right)^2 \,\Big|\, \mathscr{F}_{k/n^2}\right\} \\
&= E\left\{\left(\theta,\, Y_n\left(\frac{k+1}{n^2}\right)\right)^2 - 2\left(\theta,\, Y_n\left(\frac{k+1}{n^2}\right)\right)\left(\theta,\, E\left(\exp\frac{1}{n}A\right)Y_n\left(\frac{k}{n^2}\right)\right)\right. \\
&\qquad \left. + \left(\theta,\, E\left(\exp\frac{1}{n}A\right)Y_n\left(\frac{k}{n^2}\right)\right)^2 \,\Big|\, \mathscr{F}_{k/n^2}\right\} \\
&= E\left\{\left(\theta,\, \exp\left(\frac{1}{n}A_{k+1}\right)Y_n\left(\frac{k}{n^2}\right)\right)^2\right. \\
&\qquad - 2\left(\theta,\, \exp\left(\frac{1}{n}A_{k+1}\right)Y_n\left(\frac{k}{n^2}\right)\right)\left(\theta,\, E\left(\exp\frac{1}{n}A\right)Y_n\left(\frac{k}{n^2}\right)\right) \\
&\qquad \left. + \left(\theta,\, E\left(\exp\frac{1}{n}A\right)Y_n\left(\frac{k}{n^2}\right)\right)^2 \,\Big|\, \mathscr{F}_{k/n^2}\right\} \\
&= E\left\{\left(\theta,\, \exp\left(\frac{1}{n}A_{k+1}\right)Y_n\left(\frac{k}{n^2}\right)\right)^2 - \left(\theta,\, E\left(\exp\frac{1}{n}A\right)Y_n\left(\frac{k}{n^2}\right)\right)^2 \,\Big|\, \mathscr{F}_{k/n^2}\right\} \\
&= E\left\{\left(\theta,\, \left(\exp\left(\frac{1}{n}A_{k+1}\right) + E\left(\exp\frac{1}{n}A\right)\right)Y_n\left(\frac{k}{n^2}\right)\right)\right. \\
&\qquad \left. \times \left(\theta,\, \left(\exp\left(\frac{1}{n}A_{k+1}\right) - E\left(\exp\frac{1}{n}A\right)\right)Y_n\left(\frac{k}{n^2}\right)\right) \,\Big|\, \mathscr{F}_{k/n^2}\right\} \\
&= E\left\{\left(\theta,\, \left(2I + \frac{1}{n}A_{k+1} + \frac{1}{2n^2}(A_{k+1}^2 + EA^2)\right)Y_n\left(\frac{k}{n^2}\right)\right)\right. \\
&\qquad \left. \times \left(\theta,\, \frac{1}{n}A_{k+1} + \frac{1}{2n^2}(A_{k+1}^2 - EA^2)\right)Y_n\left(\frac{k}{n^2}\right) \,\Big|\, \mathscr{F}_{k/n^2}\right\} + O\left(\frac{1}{n^3}\right). \\
&= E\left\{\left(\theta,\, 2Y_n\left(\frac{k}{n^2}\right)\right)\left(\theta,\, \frac{1}{n}A_{k+1}Y_n\left(\frac{k}{n^2}\right)\right) \,\Big|\, \mathscr{F}_{k/n^2}\right\} \\
&\quad + E\left\{\left(\theta,\, 2Y_n\left(\frac{k}{n^2}\right)\right)\left(\theta,\, \frac{1}{2n^2}(A_{k+1}^2 - EA^2)Y_n\left(\frac{k}{n^2}\right)\right) \,\Big|\, \mathscr{F}_{k/n^2}\right\} \\
&\quad + E\left\{\left(\theta,\, \frac{1}{n}A_{k+1}Y_n\left(\frac{k}{n^2}\right)\right)\left(\theta,\, \frac{1}{n}A_{k+1}Y_n\left(\frac{k}{n^2}\right)\right) \,\Big|\, \mathscr{F}_{k/n^2}\right\} + O\left(\frac{1}{n^3}\right).
\end{aligned}
\tag{2.19}
$$

The first two terms vanish since A_{k+1} is independent of $\mathscr{F}_{k/n^2}$ and

$$E\{A_{k+1} \mid \mathscr{F}_{k/n^2}\} = EA = 0.$$

Continuing, we have the above expression equal to

$$(2.20)\qquad \begin{aligned} &E\left\{\left(\theta, A_{k+1} Y_n\left(\frac{k}{n^2}\right)\right)^2 \,\middle|\, \mathscr{F}_{k/n^2}\right\} \frac{1}{n^2} + O\left(\frac{1}{n^3}\right) \\ &\qquad = \int_{\Xi} \left(\theta, A(\xi) Y_n\left(\frac{k}{n^2}\right)\right)^2 \mu(d\xi) \frac{1}{n^2} + O\left(\frac{1}{n^3}\right). \end{aligned}$$

Note that the integral acts only upon A, not upon Y_n. Therefore, the quadratic variation process

$$(2.21)\qquad V_n^\theta(t) = \sum_{k=0}^{[n^2t]} \int_{\Xi} \left(\theta, A(\xi) Y_n\left(\frac{k}{n^2}\right)\right)^2 \mu(d\xi) \frac{1}{n^2} + O\left(\frac{1}{n}\right)$$

which, if all goes well, converges to

$$(2.22)\qquad V^\theta(t) = \int_{\Xi} \int_0^t (\theta, A(\xi) Y(s))^2 \mu(d\xi)\, ds.$$

3. A primer on weak convergence in $D_B[0, \infty)$. This primer is designed to lead us to the following theorem.

THEOREM 3.1. *Let $\{Y_n\}$ be a sequence of processes with sample paths in $D_B[0, \infty)$ such that for every $\varepsilon > 0$ and $t > 0$ there exists a compact set $K \subseteq B$ for which*

$$(3.1)\qquad \liminf_{n\to\infty} P\{Y_n(s) \in K_\varepsilon \text{ for } 0 \le s \le t\} \ge 1 - \varepsilon.$$

Then $\{Y_n\}$ is relatively compact with all of its limit points in $C_B[0, \infty)$ if and only if for each $\theta \in \mathscr{D}'$, $\{(\theta, Y_n)\}$ is relatively compact with all of its limit points in $C_R[0, \infty)$.

$\mathscr{D}'$ is any subset of B^* with the property that $\| y \| = \sup\{(\theta, y): \theta \in \mathscr{D}', \| \theta \| = 1\}$. We shall refer to equation 3.1 as a *compact containment criterion.* In the final section of this paper, we shall show how three particular models satisfy this condition. This section is an adaptation of Billingsley [4] and is close in spirit to Kurtz [11, Section 4]. A much more complete description of weak convergence can be found in the forthcoming book by Ethier and Kurtz [6]. The Skorohod topology on $D_B[0, \infty)$ is an integrated version of the Skorohod topology on $D_B[0, T)$ for $T < \infty$. Therefore, weak convergence in $D_B[0, \infty)$ is precisely weak convergence in $D_B[0, T_K)$ for a sequence T_K with $\text{limit}_{K\to\infty} T_K = \infty$. Theorem 3.1 will be applied either to a separable Banach space B or to a closed and separable subset of its dual space B^*. In these situations, Prohorov's theorem assures us that relative compactness and tightness are equivalent notions. Because our limit processes are found in $C_B[0, \infty)$, the convergence is actually uniform on bounded subsets, i.e., the convergence is on $D_B[0, T)$ in the uniform topology for each T.

We begin with a well known characterization of weak convergence and indicate the steps needed to arrive at Theorem 3.1.

In $C_B[0, \infty)$ we can use the Arzela-Ascoli theorem to characterize compact sets. However, the space $D_B[0, \infty)$ contains functions that have discontinuities of the first kind and so we must generalize the definition of the modulus of continuity. Let

$$w'(y, \delta, t) = \inf_{\Pi \in \mathscr{P}(t)} \max_i \sup\{\| y(r) - y(s) \| : r, s \in [t_{i-1}, t_i), t_{i-1}, t_i \in \Pi\} \tag{3.2}$$

where $\mathscr{P}(t)$ is the set of all partitions $\prod = \{t_0, t_1, \cdots, t_n\}$ satisfying

$$\begin{aligned} t_i - t_{i-1} > \delta \qquad &\text{for } i = 2, 3, \cdots, n \\ t_0 = 0& \\ t_n \geq t > t_{n-1}&. \end{aligned} \tag{3.3}$$

It is worth noting that w' is an increasing function both in t and in δ and that

$$w'(x, \delta, t) \leq w'(y, \delta, t) + 2 \sup_{0 \leq s \leq t+\delta} \| x(s) - y(s) \| \tag{3.4}$$

by the triangle inequality. The theorem that substitutes in $D_B[0, \infty)$ for the Arzela-Ascoli theorem in $C_B[0, \infty)$ is the following.

THEOREM 3.2. *A set $A \subseteq D_B[0, \infty)$ has compact closure in the Skorohod topology if and only if*

(i) *for each $s \geq 0$, there is a compact set $K^s \subseteq B$ such that*

$$x(s) \in K^s \quad \text{for all} \quad x \in A \tag{3.5}$$

(ii) *for each $t > 0$*

$$\lim_{\delta \to 0} \sup_{x \in A} w'(x, \delta, t) = 0. \tag{3.6}$$

This translates to give

THEOREM 3.3. *Let $\{Y_n\}$ be a sequence of processes with sample paths in $D_B[0, \infty)$. Then $\{Y_n\}$ is relatively compact if and only if*

(i) *for every $\varepsilon > 0$, and s in a dense subset of $[0, \infty)$, there is a compact set $K^s \subseteq B$ such that*

$$\liminf_{n \to \infty} P\{Y_n(s) \in K^s_\varepsilon\} \geq 1 - \varepsilon \tag{3.7}$$

and

(ii) *for every $\varepsilon > 0$ and $t > 0$, there exists $\delta > 0$ such that*

$$\limsup_{n \to \infty} P\{w'(Y_n, \delta, t) \geq \varepsilon\} \leq \varepsilon. \tag{3.8}$$

The balance of this section will link this characterization of relative compactness to the characterization in Theorem 3.1. Notice that Theorem 3.1 requires that the limit points be continuous processes. This may be verified by using the following criterion:

For $y \in D_B[0, \infty)$ and $u > 0$, define

$$J(y, u) = \sup_{0 \le s \le u} \| y(s) - y(s-) \| \tag{3.9}$$

and

$$J(y) = \int_0^\infty e^{-u}[J(y, u) \wedge 1]\, du. \tag{3.10}$$

Since J is a continuous function on $D_B[0, \infty)$, we have the following theorem:

THEOREM 3.4. *Let Y, Y_n, $n \ge 1$, be processes with sample paths in $D_B[0, \infty)$ and suppose that Y_n converges to Y. Then Y is continuous if and only if $J(Y_n)$ converges weakly to* 0.

We omit the proof. The next step involves a systematic cataloging of the oscillation of an element y in $D_B[0, \infty)$.

Let $\varepsilon > 0$ and let $\sigma_0 = \tau_0 = 0$, and define for each $k \ge 1$,

$$\tau_k = \inf\left\{ s > \tau_{k-1} : \| y(s) - y(\tau_{k-1}) \| \ge \frac{\varepsilon}{2} \right\} \tag{3.11}$$

and

$$\sigma_k = \sup\left\{ s < \tau_k : \| y(s) - y(\tau_k) \| \ge \frac{\varepsilon}{2} \right\} \tag{3.12}$$

with the convention that once τ_k reaches infinity, then both τ_ℓ and σ_ℓ stay at infinity for $\ell > k$. This choice yields the fact that (σ_k, τ_{k+1}) is the largest open interval whose closure contains τ_k and $\| y(s) - y(\tau_k) \| \le \varepsilon/2$ for all $s \in (\sigma_k, \tau_{k+1})$. In addition, if $\tau_{k+1} - \sigma_k > \delta$ for all k with $\tau_k \le t$, then $w'(y, \delta/2, t) \le \varepsilon$. The conclusion one can make is

LEMMA 3.5. *Let $\{Y_n\}$ be a sequence of processes with sample paths in $D_B[0, \infty)$. Let τ_k^n and σ_k^n be defined in the manner above. Then for every $\varepsilon > 0$ and $t > 0$ there exists $\delta > 0$ such that*

$$\limsup_{n \to \infty} P\{ w'(Y_n, \delta, t) \ge \varepsilon \} \le \varepsilon \tag{3.13}$$

if and only if for every $\varepsilon > 0$ and $t > 0$ there exists $\delta > 0$

$$\limsup_{n \to \infty} \sup_{k \ge 0} P\{ \tau_{k+1}^n - \sigma_k^n < \delta,\ \tau_k^n \le t \} \le \varepsilon. \tag{3.14}$$

The next link gives us a means of estimating this probability.

LEMMA 3.6. *Let Y be an adapted process with sample paths in $D_B[0, \infty)$. Let*

$M(t)$ be the set of all stopping times bounded by t. For $n > 0$, let

$$C(\eta) = \sup_{\tau\in M(t+2\eta)}\sup_{0\le u\le 2\eta}E\{\sup_{0\le v<3\eta\wedge\tau}\| Y(\tau + u) - Y(\tau)\| \times\| Y(\tau) - Y(\tau - v)\|\}. \tag{3.15}$$

Then for each $\varepsilon > 0$ and $\tau \in M(t + \eta)$

$$P\left\{\sup_{0\le u\le\eta}\| Y(\tau + u) - Y(\tau)\| \ge \frac{\varepsilon}{2}, \sup_{0\le v\le\eta\wedge\tau}\| Y(\tau) - Y(\tau - v)\| \ge \frac{\varepsilon}{2}\right\} \le \frac{100C(\eta)}{\varepsilon^2}. \tag{3.16}$$

The left hand side of equation (3.16) with $\tau = \tau_k^n \wedge t$ bounds

$$P\{\tau_{k+1}^n - \sigma_k^n < \eta, \tau_k^n \le t\} \tag{3.17}$$

uniformly in k. With this in mind one can establish:

THEOREM 3.7. *If for every $\varepsilon > 0$ and every s in a dense set of $[0, \infty)$ there exists a compact set $K^s \subseteq B$ such that*

$$\liminf_{n\to\infty}P\{Y_n(s) \in K_\varepsilon^s\} \ge 1 - \varepsilon, \tag{3.18}$$

then $\{Y_n\}$ is relatively compact if and only if there exists $C_n(\eta)$ such that

$$E\{C_n(\eta)\mid \mathscr{F}_n^s\} \ge E\{\| Y_n(s + u) - Y_n(s)\| \mid \mathscr{F}_n^s\} \times \| Y_n(s) - Y_n(s - v)\| \wedge 1 \tag{3.19}$$

for all

$$0 < \eta < 1, \quad 0 \le s \le t + 2\eta, \quad 0 < u \le 2\eta, \quad 0 \le v \le 3\eta \wedge s,$$

where $\mathscr{F}_n^s = \sigma\{Y_n(u): u \le s\}$,

$$\text{limit}_{\eta\to 0}\limsup_{n\to\infty}EC_n(\eta) = 0, \tag{3.20}$$

and

$$\text{limit}_{\eta\to 0}\limsup_{n\to\infty}E\| Y_n(\eta) - Y_n(0)\| = 0. \tag{3.21}$$

Now all the parts are in place to prove Theorem 3.1.

PROOF. The sufficiency of the relative compactness of $\{Y_n\}$ is easy to verify. If $F \in C(D_R[0, \infty))$, and $\theta \in \mathscr{D}'$, then the function $F_\theta(y) = F((\theta, y))$ belongs to $C(D_B[0, \infty))$. Consequently convergence in distribution of a subsequence of $\{Y_n\}$ implies convergence in distribution for the subsequence $\{(\theta, Y_n)\}$ with the same indexing. Also, if Y, a limit point for $\{Y_n\}$, is continuous, (θ, Y) is a limit point for $\{(\theta, Y_n)\}$ and (θ, Y) is continuous.

If $\{(\theta, Y_n)\}$ is relatively compact then by the elementary facts on weak convergence, $\{(\theta, Y_n)^k\}$ is relatively compact for $k = 1, 2, \cdots$. Also, because the limit points are continuous functions,

$$(3.22)\qquad (\theta_1, Y_n)(\theta_2, Y_n) = \tfrac{1}{2}\{(\theta_1 + \theta_2, Y_n)^2 - (\theta_1, Y_n)^2 - (\theta_2, Y_n)^2\}$$

is relatively compact. In addition, by equation (3.4), $\{f(Y_n)\}$ is relatively compact for any f in the closure of the algebra generated by $f(y) \equiv 1$ and the functions described in equation (3.22). This set is $C(B)$ by the Stone-Weierstrass theorem. In short, $\{(\theta, Y_n)\}$ is relatively compact for each $\theta \in \mathscr{D}'$ if and only if $\{f(Y_n)\}$ is relatively compact for each $f \in C(B)$. In addition, if Y is a candidate for a limit point of $\{Y_n\}$ and Y has a jump discontinuity, then since the linear functionals separate points, there exists $\theta \in \mathscr{D}'$ so that (θ, Y) has a jump discontinuity. Therefore, all possible limit points for $\{Y_n\}$ are processes having continuous sample paths.

For each $x \in B$, $f_x(y) = \|x - y\|$ is a continuous function on B. Since K is compact, it is also totally bounded, and hence for each $\delta > 0$, it contains a finite set S_δ so that

$$(3.24)\qquad K \subseteq \cup_{x\in S_\delta} B(x, \delta).$$

By the triangle inequality, for any $y, y' \in K$, there exists $x \in S_\delta$ so that

$$(3.25)\qquad \|y - y'\| \le |\, \|y - x\| - \|x - y'\|\, | + 2\delta.$$

For $0 < \eta < 1$, $0 \le s \le t + 2\eta$, $0 \le u \le 2\eta$, $0 \le v \le 3\eta \wedge s$,

$$\begin{aligned}
&(\|Y_n(s+u) - Y_n(s)\| \wedge 1)(\|Y_n(s) - Y_n(s-v)\| \wedge 1)\\
&\quad\le \sup_{x\in S_\delta} |(\|Y_n(s+u) - x\| - \|x - Y_n(s)\|)\\
&\qquad\times (\|Y_n(s) - x\| - \|x - Y_n(s-v)\|)|\\
(3.26)\qquad &\qquad + 4(\delta + \delta^2) + I_{\{Y_n(s)\in K_\varepsilon^c \text{ for some } s\le t+2\eta\}}\\
&\quad\le \sup_{x\in S_\delta} w'(\|Y_n(\cdot) - x\|, 5\eta, t + 3\eta)\\
&\qquad + 4(\delta + \delta^2) + I_{\{Y_n(s)\in K_\varepsilon^c \text{ for some } s\le t+2\eta\}}
\end{aligned}$$

which we choose to be $C_n(\eta)$. By condition 3.1

$$(3.27)\qquad \operatorname{limit}_{\eta\to 0} \limsup_{n\to\infty} EC_n(\eta) = 0.$$

Also by the same condition

$$\begin{aligned}
(3.28)\qquad &\|Y_n(\eta) - Y_n(0)\|\\
&\quad\le \sup_{x\in S_\delta} |\, \|Y_n(\eta) - x\| - \|x - Y_n(0)\|\, | + 2\delta + I_{\{Y(0)\in K_\varepsilon^c\}}
\end{aligned}$$

for all positive η. Therefore

$$(3.29)\qquad \operatorname{limit}_{\eta\to 0} \limsup_{n\to\infty} E\|Y_n(\eta) - Y_n(0)\| = 0$$

and the relative compactness follows from Theorem 3.6. □

Thus we have reduced our work from the investigation of Banach valued processes to real valued processes. The theorem we shall use to establish tightness for processes on the line is the following.

THEOREM 3.8. *Let* $\{X_n\}$ *be a family of processes with sample paths in* $D_R[0, \infty)$ *such that for every* $\varepsilon > 0$ *and every* s *in a dense subset of* $[0, \infty)$ *there exists a compact set* K^s *such that*

$$\liminf_{n\to\infty} P\{X_n(s) \in K_\varepsilon^s\} \geq 1 - \varepsilon. \tag{3.30}$$

Then the following criterion guarantees the relative compactness of $\{X_n\}$:
For each $t > 0$, *and some (and hence all)* $r > 0$, *and for*

$$0 < \eta < 1, \quad 0 \leq s \leq t + 2\eta, \quad 0 \leq u \leq 2\eta \tag{3.31}$$

there exists $C_n(\eta)$ *such that*

$$E\{C_n(\eta) \mid \mathscr{F}_t^n\} \geq E\{|X_n(t+u) - X_n(t)|^r \mid \mathscr{F}_t^n\} \tag{3.32}$$

and

$$\text{limit}_{\eta\to 0} \limsup_{n\to\infty} EC_n(\eta) = 0. \tag{3.33}$$

This theorem follows easily from Theorem 3.7.

4. Tightness theorems. The goal of this section is to establish the existence of a limiting random evolution. We state precisely the assumptions in the following theorem.

THEOREM 4.1. *Given the random evolutions* $\{Y_n\}$ *(defined in* (2.11)) *satisfying*

(i) *conditions* (2.1)–(2.8),
(ii) *the compact containment criterion* (3.1),
(iii) $Y_n(0)$ *converges weakly to* $Y(0)$.

Then $\{Y_n\}$ *is relatively compact in the topology of weak convergence. Furthermore, any limit process* Y *lies in* $C_B[0, \infty)$ *and for each* $\theta \in \mathscr{D}'$ *(defined in* (2.7) *and* (2.8)),

$$M^\theta(t) = \left(\theta,\ Y(t) - Y(0) - \frac{1}{2}\int_0^t \int_\Xi A^2(\xi) Y(s)\mu(d\xi)\, ds\right) \tag{4.1}$$

is a martingale, and $M^\theta(t)^2 - V^\theta(t)$ *is a martingale where*

$$V^\theta(t) = \int_0^t \int_\Xi (\theta, A(\xi)Y(s))^2 \mu(d\xi)\, ds. \tag{4.2}$$

The result will follow from a sequence of four lemmas. The first, Lemma 4.2, asserts that any possible limit resides in $C_B[0, \infty)$. This places us in the setting for which we can apply Theorem 3.1. In Lemma 4.4, we prove that $\{(\theta, Y_n)\}$ is relatively compact for each $\theta \in \mathscr{D}'$. Now, we can conclude that $\{Y_n\}$ is relatively compact. Because M^θ is a martingale, the conditional moments are easier to estimate. This is the role of Lemma 4.3. Lemma 4.5 checks the validity of statements (4.1) and (4.2). Because $Y(s)$ is not necessarily in $\cap\ \mathscr{D}(A^2(\xi))$, (4.1) and (4.2) must be viewed in the weak sense.

LEMMA 4.2. *$J(Y_n)$ converges weakly to zero.*

PROOF. In this situation, Y_n jumps at time k/n^2, $k = 1, 2, \cdots$. Therefore,

$$\tag{4.3} \begin{aligned} J(Y_n, t) &= \sup_{k \le n^2 t} \left\| Y_n\left(\frac{k}{n^2}\right) - Y_n\left(\frac{k-1}{n^2}\right) \right\| \\ &= \sup_{k \le n^2 t} \left\| \left(\left(\exp \frac{1}{n} A_k\right) - I\right) Y_n\left(\frac{k-1}{n^2}\right) \right\|. \end{aligned}$$

Then for each $\varepsilon > 0$, there exists a compact set K so that (3.1) holds.

$$\tag{4.4} J(Y_n, t) I_{K_\varepsilon} \le \sup_{k \le n^2 t} \sup_{y \in K_\varepsilon} \left\| \left(\exp \frac{1}{n} A_k - I\right) y \right\|.$$

Because K is compact, we can choose a finite set of elements S_δ from $\cap\ \mathscr{D}(A(\xi))$ so that

$$\tag{4.5} \begin{aligned} J(Y_n, t) I_{K_\varepsilon} &\le \sup_{k \le n^2 t} \sup_{y \in S_\delta} \left\| \left(\exp \frac{1}{n} A_k - I\right) y \right\| + \delta \\ &\le \sup_{k \le n^2 t} \sup_{y \in S_\delta} \left\| \int_0^{1/n} A_k \exp s A_k y \, ds \right\| + \delta \\ &\le \left(\frac{M e^\gamma}{n}\right) \sup_{k \le n^2 t} \sup_{y \in S_\delta} \| A_k y \| + \delta. \end{aligned}$$

Let $X_k = \sup_{y \in S_\delta} \| A_k y \|$. Then $\{X_k\}$ is an independent and identically distributed sequence of positive random variables and

$$\tag{4.6} J(Y_n, t) I_{K_\varepsilon} \le \left(\frac{M e^\gamma}{n}\right) \sup_{k \le n^2 t} X_k + \delta.$$

Because S_δ is a finite set, X_1 has finite second moments. Therefore

$$\tag{4.7} \operatorname{limit}_{n \to \infty} P\left\{ M \frac{e^\gamma}{n} \sup_{k \le n^2 t} X_k + \delta > \varepsilon \right\} = 0.$$

Letting $\delta \to 0$ gives the theorem.

LEMMA 4.3. *$\{M_n^\theta\}$ is relatively compact for all $\theta \in \mathscr{D}'$.*

PROOF. Let $\eta > 0$, $0 \leq t + 2\eta$, $0 \leq k \leq 2\, n^2\eta$ and $\theta \in \mathscr{D}'$. Then

$$(4.8)\quad \begin{aligned} &E\left\{\left(M_n^\theta\left(s + \frac{k}{n^2}\right) - M_n^\theta(s)\right)^2 \,\Big|\, \mathscr{F}_s^n\right\} \\ &= E\left\{\sum_{i=1}^{k-1} E\left\{\left(M_n^\theta\left(s + \frac{i}{n^2}\right) - M_n\left(s + \frac{i-1}{n^2}\right)\right)^2 \,\Big|\, \mathscr{F}^n_{s+((i-1)/n^2)}\right\} \,\Big|\, \mathscr{F}_s^n\right\}\end{aligned}$$

by the martingale property and basic facts on conditional expectations. Each term in the summand is precisely the type of expression we simplified in equation (2.20). The result of that calculation was

$$(4.9)\quad \begin{aligned} &E\left\{\left(M_n^\theta\left(s + \frac{i}{n^2}\right) - M_n^\theta\left(s + \frac{i-1}{n^2}\right)\right)^2 \,\Big|\, \mathscr{F}^n_{s+((i-1)/n^2)}\right\} \\ &= \int_\Xi \left(\theta, A(\xi)\, Y_n\left(s + \frac{i-1}{n^2}\right)\right)^2 \mu(d\xi)\, \frac{1}{n^2} + \text{extra terms.}\end{aligned}$$

In Section 2, we glibly collected these terms as $O(1/n)$. Now we must be precise. If we write Taylor's theorem in integral form and let $j = [n^2 s]$, then the extra terms are

$$(4.10)\quad \begin{aligned} E\Bigg\{&\left(\int_0^{1/n}\left(\frac{1}{n} - u\right)\left(\theta, A_{i+j}\exp(uA_{i+j})\, Y_n\left(s + \frac{i-1}{n^2}\right)\right) du\right. \\ &\left. + \int_0^{1/n}\left(\frac{1}{n} - u\right) E\left(\theta, A\, \exp(uA)\, Y_n\left(s + \frac{i-1}{n^2}\right)\right) du\right) \\ &\times \left(\theta, \left(\exp\left(\frac{1}{n} A_{i+j}\right) - E\left(\exp \frac{1}{n} A\right)\right) Y_n\left(s + \frac{i-1}{n^2}\right)\right) \,\Big|\, \mathscr{F}^n_{s+((i-1)/n^2)}\Bigg\}\end{aligned}$$

plus a second term which results from switching the plus and minus signs in the terms above. These two terms are handled in the same fashion. We now consider the first term.

Let $\varepsilon > 0$. Choose a compact K whose existence is guaranteed by the compact containment criterion so that (3.1) holds. The absolute value of (4.10) is bounded above by

$$(4.11)\quad \begin{aligned} E\Bigg\{\sup_{y\in K_\varepsilon}\Bigg|&\left(\int_0^{1/n}\left(\frac{1}{n} - u\right)\left(\theta, A\, \exp(uA)y\right) du\right. \\ &\left. + \int_0^{1/n}\left(\frac{1}{n} - u\right) E(\theta, A^2\exp(uA)y)\, du\right) \\ &\times \left(\theta, \left(\exp\left(\frac{1}{n} A(\xi)\right) - E\left(\exp \frac{1}{n} A\right)\right) y\right)\Bigg| \,\Big|\, \mathscr{F}_{s+(i/n^2)}\Bigg\}\end{aligned}$$

on the set K_ε. Choose a finite set S_δ so that this supremum may be taken over the set S_δ at the expense of an additional δ. Thus, we may bound these terms by

$$\begin{aligned} E\Bigg\{\sup_{y\in S_\delta}\Bigg(&\int_0^{1/n}\left(\frac{1}{n}-u\right)Me^{\gamma}(\|\theta\|\ \|Ay\|+\delta)\,du \\ &+\int_0^{1/n}\left(\frac{1}{n}-u\right)Me^{\gamma}(E\ \|\theta\|\ \|A^2y\|+\delta)\,du\Bigg) \\ &\times 2Me^{\gamma}(\|\theta\|\ \ \|y\|+\delta)\mid \mathscr{F}^{\,n}_{s+((i-1)/n^2)}\Bigg\} \\ \le E\Bigg\{&\sup_{y\in S_\delta}\frac{Me^{\gamma}}{n^2}\|\theta\|(\|Ay\|+E\|A^2y\|+\delta)(\|\theta\|\ \|y\|+\delta)\Bigg\} \end{aligned} \tag{4.12}$$

for each i. By estimating in a similar way, and possibly enlarging S_δ, the listed term on the right hand side of equation (4.9) is bounded by

$$\frac{1}{n^2}E\left\{\sup_{y\in S_\delta}(\|\theta\|\ \|Ay\|+\delta)\right\}. \tag{4.13}$$

Collect terms and return to equation 4.8 in order to see that

$$\begin{aligned} &E\left\{\left(M_n^{\theta}\left(s+\frac{k}{n^2}\right)-M_n^{\theta}(s)\right)^2\mid \mathscr{F}^{\,n}_{s}\right\} I_{K_\varepsilon} \\ &\le \eta t E\{\sup_{y\in S_\delta}\|\theta\|(\|Ay\|+E\|\ A^2y\|+\delta)(\|\theta\|\ \ \|y\|+\delta)\}Me^{\gamma} \\ &\quad + 8\eta t E\{\sup_{y\in S_\delta}\|\theta\|\ \ \|Ay\|+\delta\}\equiv EC_n(\eta). \end{aligned} \tag{4.14}$$

Thus, we can choose C_n independent of n, and

$$\text{limit}_{\eta\to 0}\text{limsup}_{n\to\infty}EC_n(\eta)=0. \tag{4.15}$$

After noting that $\{E\{M_n^{\theta}\}^2\}$ is bounded, we see that this theorem follows from Theorem 3.8. □

LEMMA 4.4. *$\{(\theta, Y_n)\}$ is relatively compact for all $\theta \in \mathscr{D}'$.*

PROOF. Choose K so that (3.1) holds, and let $\eta > 0$, $0 \le s \le t + 2\eta$, $r = 1$, and $0 \le k \le 2n^2\eta$. Then

$$
\begin{aligned}
&\left| \left(\theta,\ Y_n\left(s + \frac{k}{n^2}\right) - Y_n(s)\right) \right| \\
&\quad = \left| M_n^\theta\left(s + \frac{k}{n^2}\right) - M_n^\theta(s) \right. \\
&\qquad \left. + \left(\theta,\ \textstyle\sum_{i=1}^k E\left(\exp\left(\frac{1}{n} A\right) - I\right) Y_n\left(s + \frac{i-1}{n^2}\right)\right) \right| \\
&\quad \le \left| M_n^\theta\left(s + \frac{k}{n^2}\right) - M_n^\theta(s) \right| \\
&\qquad + \textstyle\sum_{i=1}^k \left| \left(\theta,\ E\left(\exp\left(\frac{1}{n} A\right) - I\right) Y_n\left(s + \frac{i-1}{n^2}\right)\right) \right| .
\end{aligned} \tag{4.16}
$$

In the next step, we plan to take the expectation and evaluate the limit superior as $n \to \infty$. By the previous proposition, the first term will vanish as $\eta \to 0$. In other words,

$$
\begin{aligned}
&\operatorname{limsup}_{n\to\infty} E \left| \left(\theta,\ Y_n\left(s + \frac{k}{n^2}\right) - Y_n(s)\right) \right| I_{K_\epsilon} \\
&\quad \le \operatorname{limsup}_{n\to\infty} E I_{K_\epsilon}\ \textstyle\sum_{i=1}^k \left| \left(\theta,\ E\left(\exp\left(\frac{1}{n} A\right) - I\right) Y_n\left(s + \frac{i-1}{n^2}\right)\right) \right| \\
&\quad + O(\eta).
\end{aligned} \tag{4.17}
$$

Since K is compact, we can choose S_δ as outlined in the proposition above. If $u \le 2\,\eta \wedge t$,

$$
\begin{aligned}
&\operatorname{limsup}_{n\to\infty} E \left| (\theta,\ Y_n(s + u) - Y_n(s)) \right| I_{K_\epsilon} \\
&\quad \le \operatorname{limsup}_{n\to\infty} \\
&\qquad \cdot E I_K \textstyle\sum_{i=1}^k \left| \displaystyle\int_0^{1/n} \left(\frac{1}{n} - r\right)\left(\theta,\ E(A^2 \exp(rA)) Y_n\left(s + \frac{i-1}{n^2}\right)\right) dr \right| \\
&\qquad + O(\eta) \\
&\quad \le \operatorname{limsup}_{n\to\infty} \\
&\qquad \cdot E \left\{ \sup_{y \in S_\delta} \textstyle\sum_{i=1}^k \displaystyle\int_0^{1/n} \left(\frac{1}{n} - r\right) \left| (\theta,\ E(A^2 \exp(rA)) y) + \delta \right| dr \right\} \\
&\quad + O(\eta) \\
&\quad \le M e^{\gamma} \eta \operatorname{limsup}_{n\to\infty} \sup_{y \in S_\delta} \left(\| \theta \| E \| A^2 y \| + \delta \right) + O(\eta)
\end{aligned} \tag{4.18}
$$

because $k/n^2 < \eta$. Hence,

$$\text{limit}_{\eta\to 0}\text{limsup}_{n\to\infty} E\,|(\theta,\ Y_n(s+u) - Y_n(s))|\, I_{K_\varepsilon} = 0. \tag{4.19}$$

Therefore, since this inequality holds for all ε, we have the theorem. □

LEMMA 4.5. *For each $\theta \in \mathscr{D}'$, $M^\theta(t)$ and $M^\theta(t)^2 - V^\theta(t)$ are martingales.*

PROOF. Recall that

$$\begin{aligned} M_n^\theta(t) &= \left(\theta,\ Y_n(t) - Y_n(0) - \textstyle\sum_{i=1}^{[n^2 t]} E\left(\exp\left(\frac{1}{n}A\right) - I\right) Y_n\left(\frac{i-1}{n^2}\right)\right) \\ &= (\theta,\ Y_n(t) - Y_n(0)) - \textstyle\sum_{i=1}^{[n^2 t]} E\left(\theta,\ \int_0^{1/n} A\exp(sA)\, Y_n\left(\frac{i-1}{n^2}\right)\right). \end{aligned} \tag{4.20}$$

Each of these terms is bounded in expectation, independent of n, by (4.14) and each term converges to a continuous process. We know that $\{Y_n\}$ is a tight sequence. Choose a subsequence $\{Y_{n'}\}$ that converges to Y. Then $M_{n'}^\theta$ converges weakly to M^θ and M^θ is a martingale. In a similar manner, one may argue that, along the same subsequence, $V_{n'}^\theta$ converges to V^θ and $M^\theta(t)^2 - V^\theta(t)$ is a martingale. □

A limiting process $Y(t)$ exists, M describes the average behavior of the limiting evolution, and V describes the variance about that average behavior. There may be one, several, or many limiting processes. We would like to incorporate M and V into a characterization of Y and use this characterization of Y to guarantee uniqueness. If a similar process exists on the Banach space dual to B, then we can use M and V alone to secure uniqueness via a duality argument.

5. Duality.

The concept of duality is embodied in the following discussion. We begin with a lemma. (See [5]).

LEMMA 5.1. *Suppose that $f: \mathbb{R}^2 \to \mathbb{R}$ is absolutely continuous on lines. Denoting $\nabla f = (f_1, f_2)$, suppose*

$$\int_0^t \int_0^t |f_i(r, s)|\, dr\, ds < \infty, \quad i = 1, 2, \quad t > 0. \tag{5.1}$$

Then for almost every s

$$f(s, 0) - f(0, s) = \int_0^s (f_1(r, s-r) - f_2(r, s-r))\, dr. \tag{5.2}$$

PROOF.

$$\begin{aligned} &\int_0^t \int_0^s (f_1(r, s-r) - f_2(r, s-r))\, dr\, ds \\ &\quad = \int_0^t \int_r^t f_1(s-r, r)\, ds\, dr - \int_0^t \int_r^t f_2(r, s-r)\, ds\, dr \\ &\quad = \int_0^t [f(t-r, r) - f(0, r) - f(r, t-r) + f(r, 0)]\, dr \\ &\quad = \int_0^t [f(r, 0) - f(0, r)]\, dr. \end{aligned} \tag{5.3}$$

Differentiating with respect to t gives the result. □

THEOREM 5.2. *Let Θ and Y be independent measurable processes, Y on a Banach space, Θ on its dual space. Let f, g, and h be measurable on the product space. If*

$$f(\theta, Y(t)) - \int_0^t h(\theta, Y(s))\, ds \tag{5.4}$$

is a martingale with respect to $\mathscr{F}_t = \sigma\{Y(s): s \le t\}$ for each θ, and

$$f(\Theta(t), y) - \int_0^t g(\Theta(s), y)\, ds \tag{5.5}$$

is a martingale with respect to $\mathscr{G}_t = \sigma\{\Theta(s): s \le t\}$ for each y, then for almost every t

$$\begin{aligned} &E[f(\Theta(t), Y(0))] - E[f(\Theta(0), Y(t))] \\ &\quad = E\left[\int_0^t g(\Theta(s), Y(t-s)) - h(\Theta(s), Y(t-s))\, ds\right]. \end{aligned} \tag{5.6}$$

PROOF. Set

$$F(s, t) = E[f(\Theta(s), Y(t))]. \tag{5.7}$$

Then by the martingale property

$$F(s, t) - F(0, t) = \int_0^s E[g(\Theta(r), Y(t))]\, dr \tag{5.8}$$

and

$$F_1(s, t) = E[g(\Theta(s), Y(t))]. \tag{5.9}$$

Similarly,

$$F_2(s, t) = E[h(\Theta(s), Y(t))]. \tag{5.10}$$

Now apply Lemma 5.1 to F.

This theorem holds for Θ and Y on any metric space. We now list three immediate corollaries. The first is the duality identity.

COROLLARY 5.3. *If* $g(\theta, y) = h(\theta, y)$ *then*

$$E[f(\Theta(t), Y(0))] = E[f(\Theta(\theta), Y(t))]. \tag{5.11}$$

COROLLARY 5.4. *Let* Y *be measurable and adapted to* $\mathscr{F}_t$ *and* τ *be an* $\mathscr{F}_t$-*stopping time. Let* Θ *be measurable and adapted to* $\mathscr{G}_t$ *and* σ *be an* $\mathscr{G}_t$-*stopping time. Suppose that* Y *and* τ *are independent of* $\mathscr{G}_t$ *and* Θ *and* σ *are independent of* $\mathscr{F}_t$, *and let*

$$f(\Theta(t \wedge \tau), y) - \int_0^{t\wedge\tau} h(\Theta(s), y)\, ds \tag{5.12}$$

be an $\mathscr{G}_t$-*martingale for each* y, *and*

$$f(\theta, Y(t \wedge \sigma)) - \int_0^{t\wedge\sigma} g(\theta, Y(s))\, ds \tag{5.13}$$

be an $\mathscr{F}_t$-*martingale for each* θ. *Then*

$$\begin{aligned} &E[f(\Theta(t \wedge \tau), Y(0))] - E[f(\Theta(0), Y(t \wedge \sigma))] \\ &\quad = \int_0^t E[I_{\{s\le\tau\}} g(\Theta(s), Y((t-s)\wedge\sigma)) \\ &\qquad\qquad - I_{\{t-s\le\sigma\}} h(\Theta(s\wedge\tau), Y(t-s))]\, ds. \end{aligned} \tag{5.14}$$

COROLLARY 5.5. *Under the conditions of Corollary* 5.4 *if in addition*, $g(\theta, y) = h(\theta, y)$ *then*

$$\begin{aligned} &E[f(\Theta(t \wedge \tau), Y(0))] - E[f(\Theta(0), Y(t \wedge \sigma))] \\ &\quad = \int_0^t E[(I_{\{s\le\tau\}} - I_{\{t-s\le\sigma\}}) h(\Theta(s\wedge\tau), Y((t-s)\wedge\sigma))]\, ds. \end{aligned} \tag{5.15}$$

The duality identity can be used in several ways. Usually one must hunt for a second process so that the identity holds for a large class. The argument continues as follows:

Let Y be any process for which

$$f(\theta, Y(t)) - \int_0^t h(\theta, Y(s))\, ds \tag{5.16}$$

is a martingale. If the duality identity holds, then

$$E[f(\Theta(t), Y(0))] = E[f(\Theta(0), Y(t))] \tag{5.17}$$

for all Y which are dual to Θ. If the class of f is sufficiently rich, then there may be only one process that could satisfy the prescription above. In other words, the discovery of a dual process allows us to assert the uniqueness of the original process.

Focus on any situation in which the sequence $\{Y_n\}$ is tight. Then a limit process exists and

$$M^\theta(t) = \left(\theta,\ Y(t) - Y(0) - \frac{1}{2}\int_0^t \int_\Xi A^2(\xi)Y(s)\mu(d\xi)\ ds\right) \tag{5.18}$$

is a martingale with quadratic variation process

$$V^\theta(t) = \int_0^t \int_\Xi (\theta, A(\xi)Y(s))^2\mu(d\xi)\ ds. \tag{5.19}$$

No canonical method is known for finding the dual process. However in this situation, we are extraordinary fortunate, since the dual process arises naturally on the dual space. Starting again, we recall that

$$\begin{aligned}(\theta,\ Y_n(t)) &= \left(\theta,\ \Pi_{i=1}^{[n^2t]} \exp \frac{1}{n} A_i Y_n(0)\right)\\ &= \left(\Pi_{i=1}^{[n^2t]} \exp \frac{1}{n} A^*_{[n^2t]-i}\theta,\ Y_n(0)\right).\end{aligned} \tag{5.20}$$

We can recreate the entire procedure, reversing the roles of the Banach space and its dual. As we do this, we would like the resulting limit process on the dual to be independent of any limit of $\{Y_n\}$. We shall do this with the following definition:

Choose an independent sequence $A'_1, A'_2, \cdots$ each according to μ and each independent of $\{A_j\}$. Let $\theta \in \bigcap_{\xi\in\Xi} \mathscr{D}(A^{*2}(\xi))$, and define

$$\Theta_n(t) = \Pi_{i=1}^{[n^2t]} \exp \frac{1}{n} A_i^*\theta. \tag{5.21}$$

If a limiting process $\Theta(t)$ exists, then we have martingales

$$M_y(t) = \left(\Theta(t) - \Theta(0) - \frac{1}{2}\int_0^t \int_\Xi A^{*2}(\xi)\Theta(s)\mu(d\xi)\ ds,\ y\right) \tag{5.22}$$

for each $y \in \bigcap_{\xi\in\Xi} \mathscr{D}(A^2(\xi))$. The quadratic variation process associated to M_y is

$$V_y(t) = \int_0^t \int_\Xi (A^*(\xi)\Theta(s),\ y)^2\mu(d\xi)\ ds. \tag{5.23}$$

THEOREM 5.6. *Under the hypotheses* (2.1)–(2.10), *a process* Θ *exists for each starting point* $\theta \in \mathscr{D}'$ *and satisfies* (5.22) *and* (5.23), *and hence the process* Y *is unique in law.*

PROOF. After only a moment of reflection, one realizes that if (2.1)–(2.10) are satisfied, then, by the symmetry of the assumptions a process Θ does exist and satisfies (5.22) and (5.23). The only cause for hesitation would be in deciding the applicability of Prohorov's theorem. However, hypothesis (2.8) asserts that $\{\Theta_n(t)\} \subseteq \mathscr{D}'$ and that $\mathscr{D}'$ is a subset of a complete and separable metric space. Therefore, the theory in Section 3 applies with trivial adaptation.

Because the increments of the Y_n processes are independent, we need only to prove that the one dimensional distributions converge. Let $f \in C_b^2(\mathbb{R})$, the set of continuous and bounded functions, having two continuous and bounded derivatives. Then by the Itô Lemma:

$$(5.24)\quad \begin{aligned} f((\theta, Y(t))) &- \frac{1}{2}\int_0^t \int_\Xi f'((\theta, Y(s)))(\theta, A^2(\xi)Y(s))\mu(d\xi)\, ds \\ &- \frac{1}{2}\int_0^t \int_\Xi f''((\theta, Y(s)))(\theta, A(\xi)Y(s))^2\mu(d\xi)\, ds \end{aligned}$$

and

$$(5.25)\quad \begin{aligned} f((\Theta(t), y)) &- \frac{1}{2}\int_0^t \int_\Xi f'((\Theta(s), y))(A^{*2}(\xi)\Theta(s), y)\mu(d\xi)\, ds \\ &- \frac{1}{2}\int_0^t \int_\Xi f''((\Theta(s), y))(A^*(\xi)\Theta(s), y)^2\mu(d\xi)\, ds \end{aligned}$$

are martingales for each $\theta \in \mathscr{D}'$. If we set

$$(5.26)\quad \begin{aligned} g(\theta, y) = \frac{1}{2}\int_\Xi f'((\theta, y))(\theta, A^2(\xi)y)\mu(d\xi) \\ - \frac{1}{2}\int_\Xi f''((\theta, y))(\theta, A(\xi)y)^2\mu(d\xi), \end{aligned}$$

then the pair f, g satisfies the conditions of Corollary 5.3 and the duality identity (5.11) holds. This determines $E[f(\theta, Y(t))]$ for each candidate Y for the limiting random evolution. Finally, by varying the starting point θ, and applying the Stone-Weierstrass theorem to the collection of functions

$$(5.27)\quad \{\textstyle\sum_{j=1}^n f_j((\theta_j, \cdot)) : f_j \in C_b^2(\mathbb{R}),\ \theta_j \in \mathscr{D}'\}$$

we can determine $E[F(Y(t))]$ for all $F \in C_b^2(B)$, and Y is unique in law. □

6. Examples. Finally, we must turn to the task of fitting the examples described in Section 1 into the scheme. As we run down the checklist of hypotheses, we soon learn that the compact containment criterion is both the most stringent requirement and the most difficult requirement to verify. In order to move off square zero, we need to have a ready supply of compact subsets of B. We shall accomplish this by introducing a Hilbert space H which compactly imbeds in the Banach space. This notion is not new [2]. For example, a mean zero Gaussian measure γ on a Banach space is uniquely determined by its reproducing kernel Hilbert space. Because the Hilbert space carries all of the information about the covariance structure of γ, it is capable of assuming a primary role in convergence theorems on Banach spaces.

Throughout this section, it will be more convenient to write E for $\int_{\Xi} \cdot\, \mu(d\xi)$ and $T(\xi, t)$ for $\exp tA(\xi)$. Usually we will not need to refer to ξ explicitly, and so its specific mention will be suppressed by simply writing $T(t) = \exp(tA)$. The notation $(\cdot, \cdot)_H$ and $\|\cdot\|_H$ will denote, respectively, the inner product and norm on H. In each of the three examples, we shall prove a lemma which concludes that for sufficiently small s

$$E\,\|T(s)f\|_H \le (1 + Ls^2)\,\|f\|_H. \tag{6.1}$$

Once we have this, compact containment follows quickly. To be more precise,

Proposition 6.1. *For the family of processes $\{Y_n\}$, and for every $\varepsilon > 0$ and $t > 0$, if inequality 6.1 holds, then there exists a compact K^t satisfying inequality (3.1), provided that $\sup_n\{E\,\|Y_n(0)\|_H\}$ is finite.*

Proof.

$$\begin{aligned} E\left\{\left\|Y_n\left(\frac{k}{n^2}\right)\right\|_H \,\middle|\, \mathscr{F}_{(k-1)/n^2}\right\} &= E\left\{\sup_{\|\theta\|_H=1}\left(\theta, Y_n\left(\frac{k}{n^2}\right)\right)_H \,\middle|\, \mathscr{F}_{(k-1)/n^2}\right\} \\ &\le \sup_{\|\theta\|_H=1} E\left\{\left(\theta, T\left(\frac{1}{n}\right)Y_n\left(\frac{k-1}{n^2}\right)\right)_H \,\middle|\, \mathscr{F}_{(k-1)/n^2}\right\} \qquad (6.2) \\ &\le \sup_{\|\theta\|_H=1}\left(\theta, Y_n\left(\frac{k-1}{n^2}\right)\right)_H = \left\|Y_n\left(\frac{k-1}{n^2}\right)\right\|_H. \end{aligned}$$

To obtain the final inequality, note that the mapping

$$s \longrightarrow (\theta, T(\xi, s)y) \tag{6.3}$$

determines a continuous semigroup on $\mathbb{R}$, i.e. multiplication by $e^{sa(\xi)}$ for some $a(\xi) \in \mathbb{R}$. Use the convexity of the exponential, Jensen's inequality, the independence of $T(1/n)$ and $\mathscr{F}_{(k-1)/n^2}$ and the fact that $Ea(\xi) = 0$. The conclusion is that $\|Y_n(k/n^2)\|_H$ is a submartingale, hence

$$\begin{aligned} P\{\sup_{0\le s\le t}\|Y_n(s)\|_H \ge M\} &\le \frac{1}{M}E\,\|Y_n(t)\| \\ &\le \frac{1}{M}\exp\frac{[n^2t]L}{n^2}E\,\|Y_n(0)\|_H \qquad (6.4) \\ &\le \frac{\exp tL}{M}E\,\|Y_n(0)\|_H < \varepsilon \end{aligned}$$

for M and n sufficiently large. Let K^t be the ball in H centered at zero with radius M. Because the imbedding is compact, K^t is compact in B. In addition, K^t satisfies (3.1). □

Example 6.1. *Bounded operators.* In the bounded operator case, we have required that a Hilbert space, as mentioned above, exist and that the set $\{A(\xi)\}$

remain a bounded set of bounded operators in H. In this setting, we can achieve the estimate we need.

LEMMA 6.2. *In the bounded operator case,* (6.1) *holds.*

PROOF. The proof is an exercise in the calculus for operators and semigroups.

$$T(t)f = f + \int_0^t T(s)Af\,ds. \tag{6.5}$$

Therefore,

$$\begin{aligned} \|T(t)f\|_H^2 &= \|f\|_H^2 + \int_0^t \int_0^t (T(s)Af, T(s')Af)_H\,ds'\,ds \\ &\quad + 2Re\int_0^t (f, T(s)Af)_H\,ds \\ &= \|f\|_H^2 + 2Re\int_0^t \int_0^s (T(s)Af, T(s')Af)_H\,ds'\,ds \\ &\quad + 2Re\int_0^t (f, T(s)Af)_H\,ds. \end{aligned} \tag{6.6}$$

Taking expectations on both sides yields

$$\begin{aligned} E\|T(t)f\|_H^2 &= \|f\|_H^2 + 2ReE\int_0^t \int_0^s (T(s)Af, T(s')Af)_H\,ds'\,ds \\ &\quad + 2ReE\int_0^t (f, T(s)Af)_H\,ds - 2ReE\int_0^t (f, Af)_H\,ds \\ &= \|f\|_H^2 + 2ReE\int_0^t \int_0^s (T(s)Af, T(s')Af)_H\,ds'\,ds \\ &\quad + 2ReE\int_0^t \int_0^s (f, T(s')A^2f)_H\,ds'\,ds. \end{aligned} \tag{6.7}$$

We have used the fact that $EA = 0$. Therefore,

$$\text{limit}_{t\to 0} \frac{E\|T(t)f\|_H^2 - \|f\|_H^2}{t^2} = E(Af, Af)_H + ReE(f, A^2f)_H. \tag{6.8}$$

In this case, A is a bounded operator, and so

$$E\|T(t)f\|_H^2 \le \|f\|_H^2 + 2t^2E(\|A^2\|_H + \|A\|_H^2)\|f\|_H^2 \tag{6.9}$$

for sufficiently small t. Taking $L = \frac{1}{4}E(\|A^2\|_H + \|A\|_H^2)$ and applying Jensen's inequality gives the lemma. □

If the process resides in a Hilbert space, then we may use the following argument to establish tightness:

By the lemma above, $E\| Y_n(s)\|$, $0 \le s \le t$, is uniformly bounded. If we also know that there is a finite dimensional subspace that contains almost all of the process up to time t, then we can conclude that there exists a compact set K so that (3.1) holds. Let y be one of the possible initial points for the evolution, and write

$$T_n(t) = \Pi_{i=1}^{[n^2t]} \exp \frac{1}{n} A_i. \tag{6.10}$$

In mathematical terms, what we seek is a projection Q with finite dimensional range so that

$$\sup_{0\le s\le t} \limsup_{n\to\infty} E\| T_n(s)y - QT_n(s)y\|^2 < \varepsilon. \tag{6.11}$$

Let R be a projection, then

$$E\| RT_n(s)y\|^2 = (y, ET_n^*(s)RT_n(s)y). \tag{6.12}$$

This suggests that we make the following definition. For L any bounded operator on a Hilbert space, let

$$\mathscr{T}_n(s)L = ET_n^*(s)LT_n(s). \tag{6.13}$$

A routine check will verify that $\mathscr{T}_n(s)$ is a semigroup. The discrete analog to the generator is

$$\begin{aligned} \mathscr{A}_n L &= n^2 E\left(\left(\exp \frac{1}{n} A^*\right) L \left(\exp \frac{1}{n} A\right) - L\right) \\ &= n^2 E \sum_{k=2}^{\infty} \frac{1}{n^k} \sum_{j=0}^{k} \frac{1}{j!} \frac{1}{(k-j)!} A^{*j} L A^{k-j}. \end{aligned} \tag{6.14}$$

Let $\mathscr{A}_L = \frac{1}{2} E(A^{*2}L + LA^2) + EALA$. Then

$$\mathscr{A}_n L - \mathscr{A} L = n^2 E \sum_{k=3}^{\infty} \frac{1}{n^k} \sum_{j=0}^{k} \frac{1}{j!} \frac{1}{(k-j)!} A^{*j} L A^{k-j}. \tag{6.15}$$

$\| \mathscr{A}_n L - \mathscr{A} L\| \to 0$ for every bounded operator L. Since $\mathscr{A}_n$ and $\mathscr{A}$ are bounded operators $\| \exp(s\mathscr{A}_n)L\text{-}\exp(s\mathscr{A})L\| \to 0$ for all bounded operators L and all

$s > 0$. Letting $\mathcal{T}(s) = \exp(s\mathcal{A})$, we would like to show that $\mathcal{T}_n(s) \to \mathcal{T}(s)$. If this were the case we could return to (6.13) to see that

$$\sup_{0\le s\le t}\limsup_{n\to\infty} E\,\| T_n(s)y - QT_n(s)y \|^2 = \sup_{0\le s\le t}(y,\ \mathcal{T}(s)(I-Q)y). \tag{6.16}$$

In addition,

$$(y,\ \mathcal{T}(s)(I-Q)y) = \sum_{n=0}^{\infty} \frac{s^n}{n!}\,(y,\ \mathcal{A}^n(I-Q)y)$$

$$= \sum_{n=0}^{\infty} \frac{s^n}{n!} \sum_{j=0}^{n} 2^{-j}\binom{n}{j} E(y,\ A^{*j}(I-Q)A^{n-j}y).$$

Because $\| I - Q \| = 1$, the tail of this sum is bounded in norm by the tail of the sum for $\| y \|^2 \exp(3/2)t\,\| A \|$. Choose N so that the sum above for $n > N$ is bounded in norm by $\varepsilon/2$. Then choose a projection Q so that each of the terms

$$E(y,\ A^{*j}(I-Q)A^{n-j}y) < \frac{\varepsilon}{(2t)^n}$$

for $n = 1, 2, \cdots, N$, and $j = 1, 2, \cdots, n$. With this choice for Q,

$$\sup_{0\le s\le t}(y,\ \mathcal{T}(s)(I-Q)y) < \varepsilon \tag{6.17}$$

and now we can choose K to be a bounded set in the range of Q.

The following proposition will close the argument.

PROPOSITION 6.3. *For all $s > 0$ and all bounded operators L,*

$$\text{limit}_{n\to\infty} \| \mathcal{T}_n(s)L - \mathcal{T}(s)L \| = 0. \tag{6.18}$$

PROOF. Because $\exp(s\mathcal{A}_n) \to \exp(s\mathcal{A})$, we only need to show that $\| \mathcal{T}_n(s)L - \exp(s\mathcal{A}_n)L \| \to 0$. This will justify calling $\mathcal{A}_n$ an analog to a generator. Also

$$\left\| \mathcal{T}_n\!\left(\frac{\ell}{n^2}\right)L - L \right\| = \left\| \sum_{n=0}^{\ell-1} \mathcal{T}_n\!\left(\frac{m}{n^2}\right)\left(\mathcal{T}_n\!\left(\frac{1}{n^2}\right)L - L\right) \right\| \le \ell e^{\omega\ell/n^2} \left\| \mathcal{T}_n\!\left(\frac{1}{n^2}\right)L - L \right\| \tag{6.19}$$

where $\omega = 2E \| A \|$. Let $k = [n^2 s]$.

$$\| \mathcal{T}_n(s)L - \exp(s\mathcal{A}_n)L \|$$
$$= \left\| \mathcal{T}_n\left(\frac{k}{n^2}\right) L - e^{-k} \sum_{j=0}^{\infty} \frac{k^j}{j!} \mathcal{T}_n\left(\frac{j}{n^2}\right) L \right\|$$
$$= e^{-k} \left\| \sum_{j=0}^{\infty} \frac{k^j}{j!} \left(\mathcal{T}_n\left(\frac{k}{n^2}\right) - \mathcal{T}_n\left(\frac{j}{n^2}\right)\right) L \right\|$$
$$\le e^{-k} \sum_{j=0}^{k} \frac{k^j}{j!} e^{j\omega/n^2} \left\| \mathcal{T}_n\left(\frac{k-j}{n^2}\right) L - L \right\|$$
$$+ e^{-k} \sum_{j=k+1}^{\infty} \frac{k^j}{j!} e^{k\omega/n^2} \left\| \mathcal{T}_n\left(\frac{j-k}{n^2}\right) L - L \right\|$$
$$\le \left(e^{-k} \sum_{j=0}^{k} \frac{k^j}{j!} e^{j\omega/n^2}(k-j)e^{(k-j)\omega/n^2} \right.$$
$$\left. + e^{-k} \sum_{j=k+1}^{\infty} \frac{k^j}{j!} e^{k\omega/n^2}(j-k)e^{(j-k)\omega/n^2} \right) \left\| \mathcal{T}_n\left(\frac{1}{n^2}\right) L - L \right\|$$
$$= \left(e^{-k} \sum_{j=0}^{k} \frac{k^j}{j!} (k-j)e^{k\omega/n^2} \right.$$
$$\left. + e^{-k} \sum_{j=k+1}^{\infty} \frac{k^j}{j!} (j-k)e^{j\omega/n^2} \right) \left\| \mathcal{T}_n\left(\frac{1}{n^2}\right) L - L \right\|.$$

For the left sum,

$$\sum_{j=0}^{k} \frac{k^j}{j!} (k-j) = \frac{k^{k+1}}{k!}.$$

For the right sum,

$$\sum_{j=k+1}^{\infty} \frac{k^j}{j!} e^{j\omega/n^2}(j-k) = \frac{k^{k+1}}{k!} e^{(k+1)\omega/n^2} + k(e^{\omega/n^2}-1)\exp(ke^{\omega/n^2}).$$

Return to the estimate, and use the Stirling formula.

$$\| \mathcal{T}_n(s)L - \exp(s\mathcal{A}_n)L \|$$
$$\le e^{-k}\left(e^{k\omega/n^2}(1+e^{\omega/n^2}) \frac{k^{k+1}}{k!} + k(e^{\omega/n^2}-1)\exp(ke^{\omega/n^2})\right)$$
$$\times \left\| \mathcal{T}_n\left(\frac{1}{n^2}\right) L - L \right\| \tag{6.20}$$
$$= \left(e^{k\omega/n^2}(1+e^{\omega/n^2})e^{-k} \frac{k^{k+1}}{k!} + k(e^{\omega/n^2}-1)\exp(k(e^{\omega/n^2}-1))\right)$$
$$\times \left\| \mathcal{T}_n\left(\frac{1}{n^2}\right) L - L \right\|$$

$$\leq \left(e^{s\omega}(1 + e^{\omega/n^2})\left(\frac{s}{2\pi}\right)^{1/2} + ns(e^{\omega/n^2} - 1)\exp(n^2 s(e^{\omega/n^2} - 1))\right)n$$

$$\times \left\| \mathscr{T}_n\left(\frac{1}{n^2}\right) L - L \right\|$$

$$\to \left(e^{s\omega}2\left(\frac{s}{2\pi}\right)^{1/2} + 0\cdot\exp(s\omega)\right) \cdot 0 = 0$$

as $n \to \infty$. □

In the case of bounded operators, $\mathscr{D}$, as defined in (2.2) is equal to B. Therefore, the hypotheses of Theorem 4.1 are satisfied, and a limit process $Y(t)$ exists with martingales (4.1) and (4.2). For the uniqueness, we must have a similar situation available for the dual process before we can apply Theorem 5.6. If B is a separable Hilbert space, we have shown that such a dual process can always be constructed. If B is finite dimensional, then all norms are equivalent, and we can always locate H, and always concoct a dual Θ.

EXAMPLE 6.2. *Transport phenomena.* In the search for an appropriate Hilbert space, consider the following line of reasoning for the transport phenomena. $A(\xi)$ is a first order differential operator. The simplest possible case is in one dimension with $A(\xi) = +(d/dx)$ or $-(d/dx)$, each with probability ½. By Donsker's invariance principle, the limiting random evolution is standard Brownian motion. Choosing $B = C_0(R)$ and choosing the Gaussian measure to be Wiener measure, one obtains the reproducing kernel Hilbert space

$$H = \left\{f \in B: f(x) = \int_0^x g(y)\,dy \text{ and } \int g^2(x)\,dx < \infty\right\}. \tag{6.21}$$

In other words, H is the Sobolev space $W^{1,2}(\mathbb{R})$. This points to a choice for $H = W^{\ell,2}(\mathbb{R}^d)$. Letting

$$\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_d) \text{ be a multi-index, } \quad |\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_d,$$

$$D^\alpha = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \frac{\partial^{\alpha_2}}{\partial x_1^{\alpha_2}} \cdots \frac{\partial^{\alpha_d}}{\partial x_d^{\alpha_d}},$$

and $\|\cdot\|_2$ be the L^2-norm, we can develop a definition of $W^{\ell,2}(\mathbb{R}^d)$ in the following way:

DEFINITION 6.4. For $|\alpha| \leq \ell$, let $D^\alpha f$ exist as an element in $L^2(\mathbb{R}^d)$ and denote

$$\|f\|_H = (\textstyle\sum_{|\alpha|\leq\ell} \|D^\alpha f\|_2^2)^{1/2}. \tag{6.22}$$

Then $\|f\|_H$ is a norm. $W^{\ell,2}(\mathbb{R}^d)$ is the completion of

$$\{f: D^\alpha f \text{ exists, } \textstyle\sum_{|\alpha|\leq\ell} \|D^\alpha f\|_2^2 < \infty\}$$

in the $\|\cdot\|_H$ norm.

This norm induces an inner product

$$(f, g)_H = \sum_{|\alpha| \le \ell} (D^\alpha f, D^\alpha g)_2. \tag{6.23}$$

By the standard theorems for Sobolev spaces [1], if $2\ell \ge d$, then the identity mapping

$$W^{\ell,2}(\mathbb{R}^d) \to C_0(\mathbb{R}^d)$$

is a compact imbedding with dense range. This will be our choice for H.

Unfortunately we cannot return to the proof of Lemma 6.2 to secure the estimate for this situation. The proof breaks down because, in the last step, we used the fact that $A(\xi)$ was bounded. For the transport phenomena, we take advantage of the identity

$$[T(t)f(x)]^2 = [f(v(x, t))]^2 = f^2(v(x, t)) = T(t)f^2(x) \tag{6.24}$$

where $v(t, x)$ is the flow beginning at x and generated by A. In order to make sense of all that will be written, we must have that $T(\xi, t)$ is a semigroup on H. In general, $T(\xi, t)$ will not be a contraction semigroup on H. The following hypotheses will guarantee that $\{T(\xi, t)\}$ has the desired properties:

$$A(\xi)f(x) = a(\xi, x)\nabla f(x). \tag{6.25}$$

Assume that $D^\alpha a(\xi, x)$ exists as a Lipschitz function with a common Lipschitz constant for all $|\alpha| \le \ell$.

Now it is time to move on to the details.

LEMMA 6.5. *For the transport phenomena, and* $f \in \bigcap_{\xi\in\Xi} \mathscr{D}(A^2(\xi))$, (6.1) *holds.*

PROOF. The proof procedes by induction on the length of α. The induction hypothesis is

$$E\| D^\beta T(t)f\|_2^2 \le (\| D^\beta f\|_2^2 + C_\beta t^2 \sum_{|\gamma|\le k} \| D^\gamma f\|_2^2) \tag{6.26}$$

for all multi-indices β having length k, and for t sufficiently small. C_β and the range for which the inequality holds depends only upon the Lipschitz constant. We shall use $\beta \le \alpha$ to mean that $\beta_i \le \alpha_i$ for each $i = 1, 2, \cdots d$. Let

$$v(x, t) = (v_1(x, t), v_2(x, t), \cdots, v_d(x, t))$$

be the flow starting at x with velocity

$$a(x) = (a_1(x), a_2(x), \cdots, a_d(x)). \tag{6.27}$$

Since $v(x, 0) = x$ and $\dfrac{\partial v}{\partial t}(x, 0) = a(x)$,

$$v(x, t) = x + ta(x) + t^2 b(x, t). \tag{6.28}$$

The function $b(x, t)$ is bounded and has bounded first derivatives for all x and for sufficiently small t by the Lipschitz condition on a and its derivatives. Even though the details are long, the central idea is not difficult. For $k = 0$, we expand

$T(t)f^2(x)$ in t by Taylor's theorem. The term of order t disappears since $EA = 0$, and the derivatives on f for the t^2 term may be placed on v, and we have control over the size of v. Therefore, $\| T(t)f \|_H = (1 + O(t^2)) \| f \|_H$. For the higher order terms, we follow the same plan. In these cases, we find that the term of order t^2 involves derivatives of f of lower order, and by induction, we have estimated this term.

For $k = 0$, there is only the zero index. Apply Taylor's theorem for $f \in C^2$. After integrating and taking the expectation we have

$$E \int [T(t)f(x)]^2 \, dx = \int f^2(x) \, dx + tE \int \nabla[f^2(x)] \cdot a(x) \, dx + E \int \int_0^t (t - s) \frac{\partial^2}{\partial s^2} f^2(v(x, s)) \, ds \, dx. \tag{6.29}$$

The middle term on the right side of equation (6.29) vanishes since $Ea = 0$. For the rightmost term, let D^2 denote the Hessian operator, and reverse the order of integration to obtain

$$E \int_0^t (t - s) \int \frac{\partial}{\partial s} v^T(x, s) \cdot D^2 f^2(v(x, s)) \cdot \frac{\partial}{\partial s} v(x, s) + \nabla f^2(v(x, s)) \cdot \frac{\partial^2}{\partial s^2} v(x, s) \, dx \, ds. \tag{6.30}$$

Focusing on the second term, and isolating one of the summands, we integrate by parts to obtain

$$-\int \left[\frac{\partial^2}{\partial x_i \partial s^2} v_i(x, s) \right] f^2(v(x, s)) \, dx \le C_0' \int f^2(v(x, s)) \, dx = C_0' \| T(s)f \|_2^2. \tag{6.31}$$

The term in brackets is one of the terms that the Lipschitz hypothesis guarantees us will be bounded uniformly. A similar estimate that uses integration by parts two times yields the same result for the first term in (6.30). Returning to equation (6.29) we have,

$$E \| T(t)f \|_2^2 \le \| f \|_2^2 + 2C_0' \int_0^t (t - s) E \| T(s)f \|_2^2 \, ds \le \| f \|_2^2 + 2C_0' \int_0^t (t' - s) E \| T(s)f \|_2^2 \, ds$$

for all $t' > t$. Apply Gronwall's inequality and let $t' \to t$, then

$$E \| T(t)f \|^2 \le \| f \|_2^2 (1 + C_0' t^2) = \| f \|_2^2 + C_0' t^2 \| f \|_2^2. \tag{6.32}$$

That handles the case $k = 0$. To continue, let β be a multi-index obtained from

α by reducing the value from one of the indices by one. For definiteness in notation during this estimate, assume that $\beta_1 = \alpha_1 - 1$. Let ∂_i denote the partial derivative with respect to x_i. From equation (6.28), we see that

$$\frac{\partial}{\partial x_1} v_1(x, t) = 1 + t \frac{\partial}{\partial x_1} a_1(x) + t^2 \partial_1 b_i(x, t) \tag{6.33}$$

and

$$\frac{\partial}{\partial x_1} v_i(x, t) = t \frac{\partial}{\partial x_1} a_i(x) + t^2 \partial_1 b_i(x, t) \tag{6.34}$$

for $i = 2, \cdots, d$. Hence,

$$D^\alpha[T(t)f(x)] = D^\alpha[f(v(x, t))] = \sum_{i=1}^d D^\beta\left[(\partial_i f)(v(x, t)) \left(\frac{\partial}{\partial x_1} v_i(x, t)\right)\right].$$

Therefore,

$$\begin{aligned} E \int & D^\alpha[T(t)f(x)]^2 \, dx \\ &= E \int \left[\sum_{i=1}^d D^\beta[T(t)(\partial_i f)(x)] \left(t \frac{\partial}{\partial x_1} a_i(x) + t^2 \partial_1 b_i(x, t)\right)\right]^2 dx \\ &\quad + E \int D^\beta[T(t)(\partial_1 f)(x)]^2 \, dx \\ &\quad + 2E \int \sum_{i=1}^d \left[D^\beta[T(t)(\partial_i f)(x)] \left(t \frac{\partial}{\partial x_1} a_i(x) + t^2 \partial_1 b_i(x, t)\right)\right] \\ &\qquad \times D^\beta[T(t)(\partial_1 f)(x)] \, dx \\ &= A + B + C. \end{aligned} \tag{6.35}$$

Let's estimate each term in turn.

$$\begin{aligned} A &\le E \int d \sum_{i=1}^d \left[D^\beta[T(t)(\partial_i f)(x)] \left(t \frac{\partial}{\partial x_1} a_i(x) + t^2 \partial_1 b_i(x, t)\right)\right]^2 dx \\ &\le C'_\beta t^2 d \sum_{i=1}^d \left(\| D^\beta(\partial_i f) \|_2^2 + C_\beta t^2 \sum_{|\gamma| \le k} \| D^\gamma(\partial_i f) \|_2^2\right) \end{aligned}$$

by the induction hypothesis. C'_β arise from the bounds on $(\partial a_i/\partial x_1)$ and $\partial_1 b_i$. By possibly increasing C'_β, we can absorb the terms of order t^4, and conclude that

$$A \le C''_\beta t^2 \sum_{|\gamma| \le k+1} E \| D^\gamma f \|_2^2. \tag{6.36}$$

Also by the induction hypothesis

$$\begin{aligned} B &\le \| D^\beta(\partial_1 f) \|_2^2 + C_\beta \sum_{|\gamma| \le k} E \| D^\gamma(\partial_1 f) \|_2^2 \\ &\le \| D^\alpha f \|_2^2 + C_\beta \sum_{|\gamma| \le k+1} E \| D^\gamma f \|_2^2. \end{aligned} \tag{6.37}$$

Lastly, we estimate C using the induction hypothesis and Taylor's theorem.

$$C \leq 2tE \int \sum_{i=1}^{d} \left[D^{\beta}T(t)(\partial_i f)(x)\left(\frac{\partial}{\partial x_1} a_i(x)\right)\right] D^{\beta}T(t)(\partial_1 f)(x)\, dx$$
$$+ C''_{\beta} C_{\beta} t^2 \sum_{|\gamma| \leq k} E \| D^{\gamma}(\partial_1 f) \|_2^2$$
$$= 2tE \int \sum_{i=1}^{d} D^{\beta}[\partial_i f(x) + tc(x, t) \cdot \nabla(\partial_i f)(x)] \frac{\partial}{\partial x_1} a_i(x)$$
$$\times D^{\beta}[\partial_1 f(x) + t\hat{c}(x, t) \cdot \nabla(\partial_1 f)(x)]\, dx$$
$$+ C''_{\beta} C_{\beta} t^2 \sum_{|\gamma| \leq k} E \| D^{\gamma}(\partial_1 f) \|_2^2$$

where c and $\hat{c}$ arising from Taylor's theorem are bounded in x and in t. Thus

$$C \leq 2tE \int \sum_{i=1}^{d} D^{\beta}(\partial_i f)(x) \frac{\partial}{\partial x_1} a_i(x) D^{\beta}(\partial_1 f)(x)\, dx$$
$$(6.38) \qquad + C'_{\beta} t^2 \int \sum_{i=1}^{d} (| D^{\beta}(\partial_i f)(x) \| \nabla(\partial_1 f)(x) |$$
$$+ | D^{\beta}(\partial_1 f)(x) | | \nabla(\partial_i f)(x) | + | \nabla(\partial_i f)(x) \| \nabla(\partial_1 f(x)) |)$$
$$+ C'_{\beta} C_{\beta} t^2 \sum_{|\gamma| \leq k+1} E \| D^{\gamma}(\partial_1 f) \|_2^2.$$

The first term vanishes since $Ea_i = 0$ for all $i = 1, 2, \cdots, d$. Sum the terms A, B, and C to see that

$$(6.39) \qquad E \| D^{\alpha} T(t) f \|_2^2 \leq (\| D^{\alpha} f \|_2^2 + C_{\alpha} t^2 \sum_{|\gamma| \leq |\alpha|} \| D^{\gamma} f \|_2^2)$$

for all multi-indices α.

$$E \| T(t) f \|_H^2 = \sum_{|\alpha| \leq \ell} E \| D^{\alpha} T(t) f \|_2^2$$
$$(6.40) \qquad \leq \sum_{|\alpha| \leq \ell} (\| D^{\alpha} f \|_2^2 + C_{\alpha} t^2 \sum_{|\gamma| \leq |\alpha|} \| D^{\gamma} f \|_2^2)$$
$$= (1 + 2Lt^2) \| f \|_H^2$$

for some constant L. The proof concludes on an application of Jensen's inequality. □

For the transport phenomena $A(\xi) = a(\xi, x) \cdot \nabla$ and the set up takes the form

$$(6.41) \qquad \text{(i)} \int_{\Xi} a(\xi, x) \mu(d\xi) = 0.$$

$$(6.42) \qquad \text{(ii)}\ \mathscr{D} \supseteq C_0^2(\mathbb{R}^d) \text{ is dense in } B.$$

For each ξ, $\exp sA(\xi)$: $\mathscr{D} \to C_0^2(\mathbb{R}^d)$.

$$(6.43) \qquad \text{(iii)} \int_{\Xi} \| a(\xi, x) \cdot \nabla f \| \mu(d\xi) < \infty, \text{ and}$$

$$(6.44)\qquad \text{(iv)} \int_{\Xi} \|(a(\xi, x)\cdot\nabla)^2 f\|\,\mu(d\xi) < \infty, \text{ for all } f \in \mathscr{D}.$$

In fact, more was assumed by the Lipschitz hypothesis.

$$(6.45)\qquad \text{(v) } \xi \to a(\xi, x) \text{ is a measurable function.}$$

$$(6.46)\qquad \text{(vi) } \|\exp sa(\xi, x)\cdot\nabla\| = 1.$$

If $\mathscr{D}' = W^{1,1}(\mathbb{R}^d)$, then $\mathscr{D}'$ has the desired properties. (See [12]). In checking compact containment for $\{\Theta_n\}$, note that $A^* = \nabla\cdot a$ does not generate a flow since $\nabla\cdot af = a\cdot\nabla f + (\nabla\cdot a)f$. The estimates above can be modified to include the potential term $\nabla\cdot a$ by approximating the semigroup as suggested by the Trotter product formula.

$$(6.47)\qquad \exp\left(\frac{1}{n}\nabla\cdot a\right)f \approx \exp\frac{1}{n}(\nabla\cdot a)\exp\frac{1}{n}(a\cdot\nabla)f.$$

EXAMPLE 6.3. *Random classical mechanics.* The notions and estimates related to the random classical mechanics are quite similar to the transport phenomena. The compact imbedding is the identity mapping

$$(6.48)\qquad W^{\ell,2}(\mathbb{R}^d) \to L^2(\mathbb{R}^d),\quad 2\ell \geq d.$$

Here, $A(\xi)f(x) = \{f(x), H(\xi, x)\}$ and the regularity hypothesis is that $D^\alpha H(\xi, x)$ exists as a Lipschitz function with a common Lipschitz constant for all $|\alpha| \leq \ell + 1$. As for the dual process, $A^* = -A$, i.e., the action of a semigroup on the dual process is the time reversal of the action of the original process. Because A appears in statements (4.1) and (4.2) of Theorem 4.1 only through a squaring, the laws of the limit processes Θ and Y are identical. In other words, Y is its own dual, and therefore its existence implies its uniqueness!

Acknowledgement. I would like to express my gratitude to the referee whose comments on a careful reading of the manuscript have considerably strengthened the presentation of this material.

REFERENCES

[1] ADAMS, ROBERT A. (1979) *Sobolev Spaces.* Academic, New York.

[2] ARONZAJN, N. (1950). Theory of reproducing kernels. *Trans. Amer. Math. Soc.* **68** 337–404.

[3] BERGER, MARC A. (1983). Central limit theorem of random matrices. Preprint, The Weizmann Institute of Science.

[4] BILLINGSLEY, PATRICK (1968). *Convergence of Probability Measures.* Academic, New York.

[5] DAWSON, DONALD A. and KURTZ, THOMAS G. (1983). Applications of duality to measure-valued diffusion processes. Preprint.

[6] ETHIER, STEWART N. and KURTZ, THOMAS G. (1983). *Markov Processes: Characterization and Convergence.* New York (to appear).

[7] GRIEGO, R. J. and HERSH, R. (1969). Random evolutions, Markov chains, and systems of partial differential equations. *Proc. National Acad. Sci.* **62** 305–308.

[8] HERSH, REUBEN (1972). Random evolutions: A survey of results and problems. *Rocky Mountain J. Math.* **4** 443–475.

[9] HERSH, R. and PINSKY, M. (1972). Random evolutions are asymptotically Gaussian. *Comm. Pure Appl. Math.* XXV 33–44.

[10] FURSTENBERG, HARRY and KESTEN, HARRY (1960). Products of random matrices. *Ann. Math. Statist.* **31** 457–469.

[11] KURTZ, THOMAS G. (1975) Semigroups of conditioned shifts and approximation of Markov processes. *Ann. Probab.* **3** 618–642.

[12] PHILLIPS, RALPH S. (1955). The adjoint semi-group. *Pacific J. Math.* **5** 269–283.

[13] PINSKY, MARK A. (editor) (1975). Probabilistic methods in differential equations. *Lecture Notes in Math.* **451**. Springer, Berlin.

[14] SIMON, BARRY AND REED, MICHAEL (1977). *Methods of Mathematical Physics, Volume II: Fourier Analysis, Self-Adjointness*. Academic, New York.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF BRITISH COLUMBIA
VANCOUVER, BRITISH COLUMBIA V6T 1Y4
CANADA.