A RANDOM TROTTER PRODUCT FORMULA

THOMAS G. KURTZ

Abstract. Let $X(t)$ be a pure jump process with state space $S$ and let $\xi_0, \xi_1, \xi_2, \cdots$ be the succession of states visited by $X(t)$, $\Delta_0, \Delta_1, \cdots$ the sojourn times in each state, $N(t)$ the number of transitions before $t$ and $\Delta_t = \sum_{k=0}^{N(t)-1} \Delta_k$. For each $x \in S$ let $T_x(t)$ be an operator semigroup on a Banach space $L$. Define $T_x(t, w) = T_{\xi_0}(1/\lambda) \Delta_0 T_{\xi_1}(1/\lambda) \Delta_1 \cdots T_{\xi_{N(t)-1}}(1/\lambda) \Delta_{N(t)-1}$. Conditions are given under which $T_x(t, w)$ will converge almost surely (or in probability) to a semigroup of operators as $\lambda \to \infty$. With $S = \{1, 2\}$ and

$X(t) = 1, \quad 2n \leq t < 2n + 1,$

$= 2, \quad 2n + 1 \leq t < 2n + 2,$

$n = 0, 1, 2, \cdots$ the result is just the "Trotter product formula".

1. Introduction. Let $X(t)$ be a stochastic process with values in a separable, locally compact metric state space $S$. Of course $X(t)$ is a function from a sample space $\Omega$ into $S$. We will assume that $\Omega = D_S(0, \infty)$, the space of right continuous functions with left hand limits taking values in $S$ and $X(t, w) = w(t)$.

Furthermore we will assume that $X(t)$ is a pure jump process; that is, there is a set $N \subseteq \Omega$ with $P(N) = 0$ such that for every pair $(t, w)$, $w \notin N$, $X(t, w) = X(t+s, w)$ for all sufficiently small $s > 0$, and $X(t, w)$ has only a finite number of discontinuities in a finite time interval. Under this assumption it makes sense to talk about $\xi_0, \xi_1, \xi_2, \cdots$, the sequence of states visited, and $\Delta_0, \Delta_1, \cdots$, the sojourn times in these states. In addition we define $N(t)$ to be the number of transitions before time $t$ and

$\Delta_t = t - \sum_{k=0}^{N(t)-1} \Delta_k.$

For each $x \in S$ let $T_x(t)$ be a semigroup of linear operators on a Banach space $L$ with infinitesimal operator $A_x$, satisfying $\|T_x(t)\| \leq e^{\alpha t}$ for some
fixed \( x \). We define the random evolution governed by \( X(\lambda t) \) by

\[
(1.1) \quad T_{\lambda}(t, w) = T_{\delta_0}(\frac{1}{\lambda} \Delta_0) T_{\delta_1}(\frac{1}{\lambda} \Delta_1) \cdots T_{\delta_{n(t)}}(\frac{1}{\lambda} \Delta_{n(t)}).
\]

The definition of a random evolution, originally given by Griego and Hersh [1] for \( X(t) \) a Markov chain, can perhaps best be motivated in the following way:

For each \( x \in S \), let \( P_x(t, y, \Gamma) \) be a Markov transition function on a measurable state space \( (E, \mathcal{B}) \), and let \( T_x(t) \) be the corresponding semi-group on \( B(E, \mathcal{B}) \), the space of bounded measurable functions. Let \( Z \) be the process (assuming one exists) whose development is governed by \( P_x(t, y, \Gamma) \) on time intervals in which \( X \) is in state \( x \). Then, at least intuitively,

\[
E(f(Z(t)) \mid X(s), s \leq t, Z(0) = z) = (T_{\delta_0}(\Delta_0) T_{\delta_1}(\Delta_1) \cdots T_{\delta_{n(t)}}(\Delta_{n(t)}))f(z).
\]

We are interested in the behavior of \( T_{\lambda}(t, w) \) as \( \lambda \) tends to infinity, that is, in what happens if the mode of development of the random evolution (or the process \( Z \)) changes at a very rapid rate.

In §2, we will give conditions under which \( T_{\lambda}(t, w) \) converges almost surely to a semigroup whose infinitesimal operator is the closure of \( \int A_x f \mu(dx) \) where \( \mu \) satisfies

\[
\lim_{t \to \infty} \frac{1}{t} \int_0^t g(X(s)) \, ds = \int g(x) \mu(dx)
\]

for all bounded continuous \( g \).

Griego and Hersh [1] and Hersh and Pinsky [2] consider the case where \( \int A_x f \mu(dx) = 0 \) (i.e. the limiting semigroup is the identity) and \( X(t) \) is a finite Markov chain. They give limit theorems for \( E(T_{\lambda}(\lambda t, w)) \) under the assumption that the \( T_x(t) \) commute. In a subsequent paper we will show that many of their results hold without the assumption of commutativity.

In what follows we will use a number of different Banach spaces. We will use subscripts on the norm notation only when there is a possibility of confusion (e.g. the norm on \( L \) will be denoted by \( \| \cdot \|_L \)).

2. The limit theorem.

Theorem (2.1). Let \( X(t) \) be a pure jump process with state space \( S \). Suppose \( S \) is a separable, locally compact metric space and there is a measure \( \mu \) on the Borel subsets of \( S \) such that \( \mu(S) = 1 \) and

\[
(2.2) \quad P\left( \lim_{t \to \infty} \frac{1}{t} \int_0^t g(X(s)) \, ds = \int g(x) \mu(dx) \right) = 1
\]

for every real, bounded, continuous function \( g \).
For each \( x \in S \) let \( T_x(t) \) be a semigroup of linear operators on a Banach space \( L \) with infinitesimal operator \( A_x \) satisfying \( \| T_x(t) \| \leq e^{\alpha t} \), for some \( \alpha \) independent of \( x \).

Let \( D \) be the set of \( f \in L \) such that \( A_x f : S \to L \) is a bounded continuous function of \( x \). Define \( A f = \int A_x f \mu(dx) \) for \( f \in D \).

If \( D \) is dense in \( L \) and \( \mathcal{R}(\mu - A) \) is dense in \( L \) for some \( \mu > \alpha \), then the closure of \( A \) is the infinitesimal operator for a strongly continuous semigroup \( T(t) \) defined on \( L \) and

\[
P \left( \lim_{\lambda \to \infty} T_{\lambda}(t, w)f = T(t)f \right) = 1
\]

for every \( f \in L \).

To prove Theorem (2.1) we will use the following which is a consequence of the results in [3].

**Theorem (2.4).** For \( 0 < \lambda < \infty \), let \( M_{\lambda} \) be a Banach space and \( \mathcal{M} \) the Banach space of bounded functions \( \lambda \to f(\lambda) \in M_{\lambda} \) with \( \| f(\cdot) \| = \sup_{\lambda} \| f(\lambda) \| \).

Let \( \lim_{\lambda \to \infty} \) denote any notion of limit (e.g. strong convergence, weak convergence) such that \( P f(\cdot) = \lim_{\lambda \to \infty} f(\lambda) \) defines a bounded linear operator from the subspace of convergent functions into another Banach space \( M \).

For each \( \lambda \) let \( S_{\lambda}(t) \) be a semigroup of linear operators on \( M_{\lambda} \) with infinitesimal operator \( B_{\lambda} \) satisfying \( \| S_{\lambda}(t) \| \leq e^{\alpha t} \) for some \( \alpha \) independent of \( \lambda \).

Suppose \( \lim_{\lambda \to \infty} f(\lambda) = 0 \) implies

\[
\lim_{\lambda \to \infty} S_{\lambda}(t)f(\lambda) = 0, \quad \text{all } t,
\]

and

\[
\lim_{\lambda \to \infty} (\lambda - B_{\lambda})^{-1}f(\lambda) = \lim_{\lambda \to \infty} \int_{0}^{\infty} e^{-\mu t}S_{\lambda}(t)f(\lambda) dt = 0. \quad \text{all } \mu > \alpha.
\]

Let

\[
\mathcal{Q}(A) = \left\{ g \in M : \exists f(\cdot) \in \mathcal{M} \ni \lim_{\lambda \to \infty} f(\lambda) = g \text{ and } \lim_{\lambda \to \infty} B_{\lambda}f(\lambda) = Ag \text{ exists} \right\}.
\]

(\( A \) may be multivalued.)

If \( \mathcal{Q}(A) \) is dense in \( M \) and \( \mathcal{R}(\mu - A) \) is dense in \( M \) for some \( \mu > \alpha \), then the closure of \( A \) is the infinitesimal operator of a strongly continuous semigroup \( T(t) \) on \( M \) and \( \lim_{\lambda \to \infty} S_{\lambda}(t)f(\lambda) = g \in M \) implies \( \lim_{\lambda \to \infty} S_{\lambda}(t)f(\lambda) = T(t)g \).

In our application of Theorem (2.4), \( M_{\lambda} \) will be the space of bounded continuous functions from \( D_{S}(0, \infty) \) into \( L \) with \( \| g \| = \sup_{w \in D_{s}} \| g(w) \| \) for all \( \lambda > 0 \), and \( M \) will be \( L \). Let \( \theta_{t} \) be the shift operator on \( D_{S}(0, \infty) \),
that is \( \theta_t w(s) = w(s + t) \). We will say \( LIM_{A \to \infty} f(\lambda, w) = g \in L \equiv M \) if

\[
P\left( \lim_{A \to \infty} \sup_{s \leq t} \| f(\lambda, \theta_A w) - g \|_L = 0 \right) = 1
\]

for all \( t > 0 \). This notion of convergence is stronger than almost sure convergence and weaker than convergence uniform in \( w \). Although we are only interested in almost sure convergence we need the extra strength in order to insure that (2.5) and (2.6) hold.

Finally, the semigroups \( S_A(t) \) are given by

\[
S_A(t)f(\lambda, w) \equiv T_A(t, w)f(\lambda, \theta_A w).
\]

If \( f(\lambda, w) \equiv f \in L \) we will write \( S_A(t)f \).

To complete the proof of Theorem (2.1) we prove the following series of lemmas, all under the assumptions of the theorem.

**Lemma (2.9).** For \( f \in D \),

\[
\| T_A(t, w)f - f \|_L \leq e^{\alpha t} \sup_x A_x f \|_L
\]

and hence, since \( D \) is dense in \( L \),

\[
\lim_{t \to 0} \sup_{w, \lambda} \| T_A(t, w)g - g \| = 0
\]

for all \( g \in L \).

**Proof.**

\[
\| T_A(t, w)f - f \|_L \\
\leq e^{\alpha T} \sum_{k=0}^{N_A(t) - 1} \left\| T_{\xi_k} \left( \frac{1}{\lambda} \Delta_0 \right) \cdots T_{\xi_{k-1}} \left( \frac{1}{\lambda} \Delta_{k-1} \right) \left( T_{\xi_k} \left( \frac{1}{\lambda} \Delta_k \right) - I \right) f \right\|_L \\
\leq e^{\alpha T} \sum_{k=0}^{N_A(t) - 1} \frac{1}{\lambda} \Delta_k \| A_x f \|_L + \frac{1}{\lambda} \Delta_k \| A_x f \|_L \\
\leq e^{\alpha T} \sup_x \| A_x f \|_L.
\]

**Lemma (2.13).** There is a function \( \varepsilon(\lambda) \) satisfying \( \lim_{\lambda \to \infty} \varepsilon(\lambda) = 0 \) and

\[
P\left( \lim_{\lambda \to \infty} \sup_{t \leq T} \left| \frac{1}{\varepsilon(\lambda)} \int_t^{t + \varepsilon(\lambda)} g(X(\lambda s, w)) ds - \int g(x) \mu(dx) \right| = 0 \right) = 1
\]

for every real, bounded continuous function \( g \) and every \( T > 0 \).

**Proof.** Since the claim is that certain linear functionals of norm one on the space of bounded continuous functions converge to a bounded linear
functional of norm one that is given by a measure, it will suffice to prove the result for continuous functions vanishing at infinity. Since the space of continuous functions on \( S \) vanishing at infinity is separable, we need only consider a countable dense subset, say \( g_1 g_2 g_3 \cdots \).

Note that

\[
(2.15) \quad \frac{1}{\varepsilon} \int_t^{t+\varepsilon} g(X(\lambda s, w)) \, ds = \frac{1}{\lambda \varepsilon} \int_t^{\lambda(t+\varepsilon)} g(X(s, w)) \, ds
\]

is uniformly continuous in \( t \).

Consequently (2.2) implies

\[
P\left( \lim_{\lambda \to \infty} \sup_{t \leq T} \frac{1}{\varepsilon} \int_t^{t+\varepsilon} g(X(\lambda s, w)) \, ds - \int g(x) \mu(dx) \, \right) = 0 = 1,
\]

for every \( g, \varepsilon > 0 \) and \( T > 0 \). Let \( \varepsilon_n \to 0, T_n \to \infty \) and \( \delta_n \to 0 \). Then there exists a \( \lambda_n \) such that

\[
\sup_{\varepsilon_n} P\left( \left| \frac{1}{\varepsilon_n} \int_t^{t+\varepsilon_n} g(X(\lambda s, w)) \, ds - \int g(x) \mu(dx) \right| > \delta_n \right) \leq \delta_n.
\]

The lemma follows by setting \( \varepsilon(\lambda) = \varepsilon_n \) for \( \lambda_n \leq \lambda < \lambda_{n+1} \).

**Lemma (2.16).** Let \( f \in D \). Then \( g(w) = \frac{1}{\varepsilon} \int_0^\varepsilon T_\lambda(s, w) f \, ds \) is in \( \mathcal{D}(B_\lambda) \) and

\[
(2.17) \quad B_\lambda g(w) = \frac{1}{\varepsilon} \int_0^\varepsilon T_\lambda(s, w) A_X(\lambda s, w) f \, ds.
\]

**Proof.** The fact that \( g(w) \in \mathcal{D}(B_\lambda) \) is a standard result of semigroup theory. The form of \( B_\lambda g(w) \) is obtained from the following inequality.

\[
(2.18) \quad \left\| \frac{1}{\varepsilon} \int_0^\varepsilon T_\lambda(s, w) A_X(\lambda s, w) f \, ds - \frac{S_\lambda(t) g(w) - g(w)}{t} \right\|_L
\]

\[
= \left\| \frac{1}{\varepsilon} \int_0^\varepsilon T_\lambda(s, w) \left( A_X(\lambda s, w) f - \frac{T_\lambda(t, \theta_\lambda w) f - f}{t} \right) \, ds \right\|_L
\]

\[
\leq \frac{1}{\varepsilon} \int_0^\varepsilon \left\| A_X(\lambda s, w) f - \frac{T_\lambda(t, \theta_\lambda w) f - f}{t} \right\|_L \, ds.
\]

Noting that \( X(\lambda s, w) = X(0, \theta_\lambda w) \), we observe

\[
\left\| A_X(\lambda s, w) f - \frac{T_\lambda(t, \theta_\lambda w) f - f}{t} \right\|_L
\]

is bounded by (2.10) and goes to zero as \( t \) goes to zero for all \( s \). The lemma then follows by the dominated convergence theorem.
Lemma (2.19). Let \( f \in D \). Define
\[
f(\lambda, w) = \frac{1}{\epsilon(\lambda)} \int_0^{\epsilon(\lambda)} T_\lambda(s, w) f \, ds.
\]
Then
\[
\lim_{\lambda \to \infty} f(\lambda, w) = f
\]
and
\[
\lim_{\lambda \to \infty} B_\lambda(\lambda, w) = \int A_x f \mu(dx).
\]

Proof. We must show that
\[
limit_{\lambda \to \infty} \sup_{t \leq T} \left\| B_\lambda f(\lambda, \theta_{\lambda t} w) - \int A_x f \mu(dx) \right\|_L = 0
\]
almost surely.

\[
\sup_{t \leq T} \left\| B_\lambda f(\lambda, \theta_{\lambda t} w) - \int A_x f \mu(dx) \right\|_L
\]
\[
= \sup_{t \leq T} \left\| \frac{1}{\epsilon(\lambda)} \int_0^{\epsilon(\lambda)} T_\lambda(s, \theta_{\lambda t} w) A_{X(\lambda s, \theta_{\lambda t} w)} f - \int A_x f \mu(dx) \right\|_L
\]
\[
\leq \sup_{t \leq T} \left\| \frac{1}{\epsilon(\lambda)} \int_0^{\epsilon(\lambda)} (T_\lambda(s, \theta_{\lambda t} w) - I) A_{X(\lambda s, \theta_{\lambda t} w)} f \right\|
\]
\[
+ \sup_{t \leq T} \left\| \frac{1}{\epsilon(\lambda)} \int_0^{\epsilon(\lambda)} A_{X(\lambda s, \theta_{\lambda t} w)} f \right\|.
\]

The second term on the right can be rewritten as
\[
\sup_{t \leq T} \left\| \frac{1}{\epsilon(\lambda)} \int_t^{t+\epsilon(\lambda)} A_{X(\lambda s, w)} f ds - \int A_x f \mu(dx) \right\|
\]
and goes to zero almost surely by (2.14), the boundedness and continuity of \( A_x f \) as a function of \( x \), and the separability and local compactness of \( S \).

The first term on the right of (2.20) can be bounded by
\[
\sup_{t \leq T} \sup_{x \in K} \left\| \frac{1}{\epsilon(\lambda)} \int_0^{\epsilon(\lambda)} (T_\lambda(s, \theta_{\lambda t} w) - I) A_x f \right\|
\]
\[
+ (2 \sup \| A_x f \| \left( \sup_{t \leq T} \frac{1}{\epsilon(\lambda)} \int_0^{\epsilon(\lambda)} \chi_K(X(\lambda s, \theta_{\lambda t} w)) ds \right).
\]

A compact set \( K \) can be selected so that the \( \lim \sup \) of the second term on the right of (2.21) can be made arbitrarily small. Given a compact \( K \) the first goes to zero by the continuity of \( A_x f \), the compactness of \( K \) and (2.11).
Proof of Theorem (2.1). Lemma (2.19) implies that the operator \( A \) in Theorem (2.4) is an extension of \( Af = \int A_x f \mu(dx) \). Consequently, under the hypotheses of Theorem (2.1), Theorem (2.4) implies, for all \( f \in L \) and \( f(w) = f \),

\[
\lim_{\lambda \to \infty} S_\lambda(t)f(w) = T(t)f.
\]

This implies (2.3).

Remark. Since the probability measure in (2.2) is arbitrary, we have in fact proved convergence for every \( w \in D_\alpha(0, \infty) \) that is constant except for a discrete set of jumps and satisfies

\[
\lim_{t \to \infty} \frac{1}{t} \int_0^t g(w(t)) \, dt = \int g(x) \mu(dx)
\]

for all continuous \( g \). Consequently Theorem (2.1) gives a generalization of the "Trotter product formula", that is

Theorem (2.22) (Trotter [4]). Suppose \( T(t) \) and \( S(t) \) are semigroups of linear operators on \( L \), with infinitesimal operators \( A \) and \( B \), satisfying \( \|T(t)\| \leq e^{\alpha t} \) and \( \|S(t)\| \leq e^{\beta t} \). If \( \mathcal{D}(A) \cap \mathcal{D}(B) \) is dense in \( L \) and \( \mathcal{D}(\mu - \frac{1}{2}(A+B)) \) is dense in \( L \) for some \( \mu > \alpha \) then the closure of \( \frac{1}{2}(A+B) \) is the infinitesimal operator of a semigroup \( U(t) \) on \( L \) and

\[
\lim_{h \to 0} (T(h/2)S(h/2))^{[t/h]}f = U(t)f
\]

for all \( f \in L \).

We observe that almost sure convergence in (2.2) and (2.3) can be replaced by convergence in probability with only minor alteration in the proof. In particular, the notion of convergence becomes: \( \lim_{\lambda \to \infty} f(\lambda, w) = g \in L \) if

\[
\lim_{\lambda \to \infty} \mathbb{P}\left( \sup_{x \leq t} \|f(\lambda, \theta_{\lambda}w) - g\| > \varepsilon \right) = 0
\]

for every \( \varepsilon > 0 \) and every \( t > 0 \).

Example. Let \( X(t) \) satisfy the conditions of Theorem (2.1). Let

\[
F(x, z) : S \times \mathbb{R}^n \to \mathbb{R}^n
\]

be bounded and satisfy

\[
\lim_{r \to r_0} \sup_{z} |F(x, z) - F(x_0, z)| = 0
\]

for all \( x_0 \in S \), and

\[
\sup_x |F(x, z_1) - F(x, z_2)| < M|z_1 - z_2|
\]
for all \( z_1, z_2 \in \mathbb{R}^n \) and some fixed \( M \). Let \( Z_\lambda(t, z) \) be the solution of

\[
Z_\lambda(t, z) = z + \int_0^t F(X(\lambda s), Z_\lambda(s, z)) \, ds.
\]

Theorem (2.1) implies

\[
\lim_{\lambda \to \infty} \sup_{z} \left| Z_\lambda(t, z) - Z(t, z) \right| = 0
\]

where \( Z(t, z) \) is the solution of

\[
Z(t, z) = z + \int_0^t \int_S F(x, Z(s, z)) \mu(dx) \, ds.
\]

REFERENCES


DEPARTMENT OF MATHEMATICS, UNIVERSITY OF WISCONSIN, MADISON, WISCONSIN 53706