SOLUTIONS OF ORDINARY DIFFERENTIAL EQUATIONS AS LIMITS OF PURE JUMP MARKOV PROCESSES

THOMAS G. KURTZ, University of Wisconsin

1. Introduction

In a great variety of fields, e.g., biology, epidemic theory, physics, and chemistry, ordinary differential equations are used to give continuous deterministic models for dynamic processes which are actually discrete and random in their development. Perhaps the simplest example is the differential equation

$$\frac{d}{dt}M = \lambda M,$$

used to describe a number of processes including radioactive decay and population growth.

Most of these processes may also be described using a Markov chain model. For example, the usual Markov chain analog for (1.1) would be a branching process X(t) with

$$E(X(t)) = X(0)e^{\lambda t}$$
.

In this case, it is well known that for a sequence of initial values $X_n(0)$ satisfying

(1.2)
$$\lim_{n\to\infty}\frac{X_n(0)}{n}=M(0),$$

we have, for every $\varepsilon > 0$,

(1.3)
$$\lim_{n\to\infty} P\left\{\sup_{s\leq t} \left| \frac{X_n(s)}{n} - M(s) \right| > \varepsilon \right\} = 0,$$

where M(s) is the solution of (1.1) with initial value M(0), that is

$$M(s) = M(0)e^{\lambda s}$$
.

It is the purpose of this paper to extend (1.3) to a large class of differential equations and approximating pure jump Markov processes.

Received 22 July 1969. Research supported in part under contract NSF GP 4784 at Stanford University, Stanford, California, and the NIH at the University of Wisconsin, Madison, Wisconsin.

2. The limit theorems

Let $X_n(t)$ be a sequence of pure jump Markov processes with measurable state spaces (E_n, \mathcal{B}_n) , where $E_n \in \mathcal{B}^K$, the Borel sets in \mathbb{R}^K , and $\mathcal{B}_n = \{E_n \cap B \colon B \in \mathcal{B}^K\}$.

Suppose $X_n(t)$ has right continuous sample paths. Denote the parameter of the exponential waiting time distribution for $x \in E_n$ by $\lambda_n(x)$ and the exit distribution by $\mu_n(x,\Gamma)$, $\Gamma \in \mathcal{B}_n$, that is

$$P\{X_n(\tau_x^n) \in \Gamma \mid X(0) = x\} = \mu_n(x,\Gamma),$$

where τ_x^n is the first exit time from x, and

$$\lambda_n(x) = \left[E(\tau_x^n \mid X_n(0) = x) \right]^{-1}.$$

We assume that for each $\Gamma \in \mathcal{B}_n$, $\mu_n(x, \Gamma)$ and $\lambda_n(x)$ are \mathcal{B}_n -measurable functions of x, and $\lambda_n(x)$ is bounded on bounded subsets of E_n . Define

$$F_n(x) = \lambda_n(x) \int_E (z-x)\mu_n(x,dz).$$

Note that $F_n: E_n \to \mathbb{R}^K$.

Proposition (2.1). Suppose

(2.2)
$$\sup_{n} \sup_{x \in E_{n}} \lambda_{n}(x) \int_{E_{n}} |z - x| \mu_{n}(x, dz) < \infty$$

and there exists a sequence $\varepsilon_n > 0$ with

$$\lim_{n\to\infty}\,\varepsilon_n=0\,,$$

such that

(2.3)
$$\lim_{n\to\infty} \sup_{x\in E_{-}} \lambda_n(x) \int_{|z-x|>\varepsilon_n} |z-x| \mu_n(x,dz) = 0.$$

Then for every $\delta > 0$

$$\lim_{n \to \infty} \sup_{x \in E_n} P\{ \sup_{s \le t} |X_n(s) - X_n(0) - \int_0^s F_n(X_n(u)) du | > \delta |X_n(0) = x \} = 0.$$

Proof. Without loss of generality, we may assume

(2.4)
$$\mu_n(x, \{z: z_1 = \sup_{i} |z_i - x_i| + x_1\}) = 1,$$

where z_i denotes the *i*th component of the vector z, since we can always increase the dimension of the state space by one and define a new process $X_n(t) = (X_{n,1}(t), \tilde{X}_n(t))$ where $\tilde{X}_n(t)$ is the original process and $X_{n,1}(t)$ is determined by (2.4). Let

$$Y_n(t) = \int_0^t F_n(X_n(s))ds + Y_n(0).$$

The pair $(X_n(t), Y_n(t))$ is a Markov process with state space $E_n \times \mathbb{R}^K$.

We now apply the methods discussed in [3] to the sequence of processes $(X_n(t), Y_n(t))$.

Let L_n be the Banach space of bounded, measurable functions on $E_n \times \mathbb{R}^K$, and let L be the Banach space of bounded, continuous functions on \mathbb{R}^K that vanish at infinity. In all cases the norm is the sup norm. We define the semigroup

$$T_n(t) \colon L_n \to L_n$$

by

$$T_n(t)f(x,y) = E\{f(X_n(t), Y_n(t)) \mid X_n(0) = x, Y_n(0) = y\}.$$

The semigroup of operators on L in which we are interested is just $T(t) \equiv I$, the identity on L. For $f \in L$ define $P_n f(x, y) \in L_n$ by

$$P_n f(x, y) = f(x - y).$$

In particular, if f is differentiable and has compact support, $P_n f$ will be in the domain of the weak infinitesimal operator \tilde{A}_n of $T_n(t)$, as defined in Dynkin [1] with

$$(2.5) \quad \tilde{A}_n P_n f(x,y) = \int_E (f(z-y) - f(x-y) - \sum_i (z_i - x_i) \partial_i f(x-y)) \lambda_n(x) \mu_n(x,dz),$$

where $\partial_i f$ denotes the derivative of f with respect to the *i*th variable. Let

$$\partial f = \sup_{i} \sup_{u} \left| \partial_{i} f(u) \right|$$

and

$$w(\varepsilon) = \sup_{i} \sup_{u} \sup_{|s| \le \varepsilon} \left| \partial_{i} f(u+s) - \partial_{i} f(u) \right|.$$

We note that

$$\lim_{\varepsilon \to 0} w(\varepsilon) = 0.$$

Applying the mean value theorem to (2.5), we have

$$\begin{split} \left| \widetilde{A}_n P_n f(x, y) \right| &\leq 2 \partial f \lambda_n(x) \int_{|z-x| > \varepsilon_n} \left| z - x \right| \mu_n(x, dz) \\ &+ K w(\varepsilon_n) \lambda_n(x) \int_{|z-x| \leq \varepsilon_n} \left| z - x \right| \mu_n(x, dz). \end{split}$$

Therefore,

$$\lim_{n\to\infty} \sup_{x,y} \left| \tilde{A}_n P_n f(x,y) \right| = 0.$$

Theorems (2.1) and (3.5) of [3] imply for every $f \in L$

$$\lim_{n\to\infty} \sup_{s\le t} \sup_{x,y} \left| E_{(x,y)}(f(X_n(s)-Y_n(s))) - f(x-y) \right| = 0.$$

It follows immediately that for every $\varepsilon > 0$

$$(2.6) \lim_{n \to \infty} \sup_{x \in E_n, s \le t} P\left\{ \left| X_n(s) - X_n(0) - \int_0^s F_n(X_n(u)) du \right| > \varepsilon \left| X_n(0) = x \right\} = 0.$$

Let

$$\bar{F} = \sup_{x} \sup_{x} |F_n(x)|$$

and suppose

$$0 = t_0 < t_1 < \dots < t_k = t_k$$

with

$$\eta = \max_{i} (t_i - t_{i-1}).$$

Then

$$\sup_{s \leq t} |X_{n}(s) - X_{n}(0) - \int_{0}^{s} F_{n}(X_{n}(u)) du |$$

$$(2.7) \qquad \leq \sup_{1 \leq i \leq k} |X_{n}(t_{i}) - X_{n}(0) - \int_{0}^{t_{i}} F_{n}(X_{n}(u)) du | +$$

$$\sup_{1 \leq i \leq k} \sup_{t_{i-1} \leq s \leq t_{i}} |X_{n}(s) - X_{n}(t_{i-1}) - \int_{t_{i-1}}^{s} F_{n}(X_{n}(u)) du |.$$

Noting that

$$\sup_{t_{i-1} \le s \le t_i} |X_n(s) - X_n(t_{i-1})| \le \sqrt{K} |X_{n,1}(t_i) - X_{n,1}(t_{i-1})|,$$

the second term in (2.7) is bounded by

$$\eta \bar{F} + \sup_{i} \sqrt{K |X_{n,1}(t_i) - X_{n,1}(t_{i-1})} - \int_{t_{i-1}}^{t_i} F_{n,1}(X_n(u)) du + \sqrt{K \eta \bar{F}}.$$

It follows from (2.6) that for every $\varepsilon > 0$

$$\lim_{n\to\infty} \sup_{x\in E_n} P\Big\{\sup_{1\leq i\leq k} \Big| X_n(t_i) - X_n(0) - \int_0^{t_i} F_n(X_n(u)) du \Big| > \varepsilon \Big| X_n(0) = x\Big\} = 0,$$

and hence

$$\lim_{n\to\infty} \sup_{x\in E_n} P\left\{ \sup_{s\leq t} \left| X_n(s) - X_n(0) - \int_0^s F_n(X_n(u)) du \right| > \sqrt{K\varepsilon + (\sqrt{K} + 1)\eta F} \left| X_n(0) = x \right| \right\} = 0.$$

But η and ε may be selected so that

$$\sqrt{K\varepsilon} + (\sqrt{K} + 1)\eta \bar{F} < \delta$$

which completes the proof.

Corollary (2.8) In place of (2.2) and (2.3), suppose there exist $\Gamma_n \subset E_n$ such that

(2.9)
$$\sup_{n} \sup_{x \in \Gamma} \lambda_n(x) \int_{F} |z - x| \mu_n(x, dz) < \infty$$

and

(2.10)
$$\lim_{n\to\infty} \sup_{x\in\Gamma_n} \lambda_n(x) \int_{|z-x|>\varepsilon_n} |z-x| \mu_n(x,dz) = 0.$$

Let τ be the first exit time from Γ_n . Then for every $\delta > 0$

$$\lim_{n\to\infty} \sup_{x\in E_n} P\left\{\sup_{s\leq t} \left|X_n(s\wedge\tau) - X_n(0) - \int_0^{s\wedge\tau} F_n(X_n(u))du\right| > \delta \left|X_n(0) = x\right\} = 0.$$

Proof. Let $\tilde{X}_n(s) = X_n(s \wedge \tau)$. Then $\tilde{X}_n(s)$ is a pure jump Markov process with waiting time parameter

$$\tilde{\lambda}_n(x) = \begin{cases} \lambda_n(x), & x \in \Gamma_n \\ 0, & x \notin \Gamma_n \end{cases}$$

and exit distribution

$$\tilde{\mu}_n(x,\Gamma) = \begin{cases} \mu_n(x,\Gamma), & x \in \Gamma_n, \\ \text{arbitrary}, & x \notin \Gamma_n. \end{cases}$$

The sequence $\tilde{X}_n(s)$ satisfies (2.2) and (2.3) and the corollary follows.

Theorem (2.11). Suppose there exists $E \subset \mathbb{R}^K$, a function $F: E \to \mathbb{R}^K$ and a constant M such that

$$|F(x) - F(y)| \leq M|x-y|, \quad x, y \in E,$$

and

(2.12)
$$\lim_{n\to\infty} \sup_{x\in E_n\cap E} |F_n(x)-F(x)| = 0.$$

Let $X(s,x_0)$ $0 \le s \le t$, $x_0 \in E$, satisfy

$$X(0,x_0) = x_0,$$

$$\frac{\partial}{\partial s}X(s,x_0) = F(X(s,x_0)),$$

and suppose there exists $\eta > 0$ such that for every n

(2.13)
$$E_n \cap \{ y \in \mathbb{R}^K : \inf_{s \le t} |y - X(s, x_0)| \le \eta \} \subset E.$$

Suppose the sequence satisfies the conditions of Corollary (2.8) for $\Gamma_n = E_n \cap E$. Then $\lim_{n\to\infty} X_n(0) = x_0$ implies for every $\delta > 0$

(2.14)
$$\lim_{n \to \infty} P\{\sup_{s \le t} |X_n(s) - X(s, x_0)| > \delta\} = 0.$$

Proof. Corollary (2.8) and Condition (2.12) imply for every $\varepsilon > 0$

$$(2.15) \qquad \lim_{n\to\infty} P\left\{\sup_{s\leq t} \left|X_n(s\wedge\tau)-X_n(0)-\int_0^{s\wedge\tau} F(X_n(u))du\right|>\varepsilon\right\} = 0.$$

If

$$\sup_{s \le t} \left| X_n(s \wedge \tau) - X_n(0) - \int_0^{s \wedge \tau} F(X_n(u)) du \right| \le \varepsilon,$$

a standard argument implies, for n sufficiently large and ε sufficiently small,

(2.16)
$$\begin{aligned} \sup_{s \leq t} & \left| X_n(s \wedge \tau) - X(s \wedge \tau, x_0) \right| \\ & \leq \left(\left| X_n(0) - x_0 \right| + \varepsilon \right) e^{Mt} \\ & \leq \delta \wedge \eta \,. \end{aligned}$$

But if $|X_n(t \wedge \tau) - X(t \wedge \tau, x_0)| < \eta$, then $X_n(t \wedge \tau) \in \Gamma_n$ and $\tau > t$. Therefore (2.16) can be written

$$\sup_{s \le t} |X_n(s) - X(s, x_0)| \le \delta \wedge n,$$

and (2.14) follows from (2.15).

3. Density dependent families

We will now consider a somewhat more intuitive special case of Theorem (2.11). Denote by \mathbb{Z}^K , the set of K-vectors with integer components.

A one parameter family of Markov chains, $X_v(t)$, v positive with state spaces $E_v \subset \mathbb{Z}^K$, will be called density dependent if and only if there exist continuous functions f(x, l), $x \in \mathbb{R}^K$, $l \in \mathbb{Z}^K$, such that the infinitesimal parameters corresponding to $X_v(t)$ are given by

$$q_{k,k+l} = vf\left(\frac{1}{v}k,l\right), \qquad l \neq 0.$$

Define

$$F(x) = \sum_{l} lf(x, l).$$

Theorem (3.1). Suppose there exists an open set $E \subset \mathbb{R}^K$ and a constant M_E such that

$$(3.2) |F(x) - F(y)| < M_E |x - y| x, y \in E,$$

(3.3)
$$\sup_{x \in E} \sum_{l} |l| f(x, l) < \infty,$$

and

(3.4)
$$\lim_{d\to\infty} \sup_{x\in E} \sum_{|l|>d} |l| f(x,l) = 0.$$

Then for every trajectory $X(s, x_0)$ satisfying

$$X(0,x_0) = x_0,$$

$$X(s,x_0) \in E$$
 $0 \le s \le t$,

and

$$\frac{\partial}{\partial s}X(s,x_0) = F(X(s,x_0)),$$

 $\lim_{v\to\infty} v^{-1}X_v(0) = x_0 \text{ implies for every } \delta > 0,$

$$\lim_{v\to\infty} P\left\{\sup_{s\leq t} \left| \frac{1}{v}X_v(s) - X(s,x_0) \right| > \delta \right\} = 0.$$

Proof. In order to apply Theorem (2.11), we define

$$\tilde{X}_{v}(s) = \frac{1}{v}X_{v}(s),$$

and

$$\tilde{E}_v = \left\{ \frac{1}{v}k \colon k \in E_v \right\}.$$

For $v^{-1}k \in \tilde{E}_v$, the waiting time parameter and the exit distribution are given by

$$\lambda_v \left(\frac{1}{v} k\right) = v \sum_{l} f\left(\frac{1}{v} k, l\right)$$

and

$$\mu_{v} \ \left(\frac{1}{v}k, \Gamma\right) \ = \ v \sum_{\{l, \, (1/v)l \, \in \, \Gamma\}} f\!\left(\frac{1}{v}k, l\right) \middle/ \lambda_{v}\!\left(\frac{1}{v}\,k\right).$$

We observe that

$$F_{v}\left(\frac{1}{v}k\right) = \sum_{l} \frac{l}{v} v f\left(\frac{1}{v}k, l\right)$$
$$= F\left(\frac{1}{v}k\right),$$

so that (2.12) is trivially satisfied.

Since E is open, for every closed trajectory $X(s, x_0)$ contained in E there exists $\eta > 0$ such that (2.13) is satisfied.

With $\Gamma_v = \tilde{E}_v \cap E$,

$$\sup_{v} \sup_{x \in \Gamma_{u}} \lambda_{v}(x) \int_{\widetilde{E}_{u}} \left| z - x \right| \mu_{v}(x, dz) \equiv \sup_{v} \sup_{(1/v)k \in \Gamma_{u}} \sum_{l} \frac{\left| l \right|}{v} v f\left(\frac{1}{v}k, l\right) < \infty,$$

by (3.3), and with $\varepsilon_v = v^{-\frac{1}{2}}$

$$\lim_{v \to \infty} \sup_{v^{-1}k \in \Gamma_{u}} \sum_{l/v > \varepsilon_{u}} \frac{\left| l \right|}{v} v f\left(\frac{1}{v}k, l\right) = 0$$

by (3.4), and the theorem follows.

As an example of a density dependent family, let us consider Markov chains analogous to a deterministic ecological model of Volterra which is discussed in [2]. This is a model for the development of two species, one of which is a food source for the other, say a population of rabbits and foxes. Implicit in a deterministic model of this nature is an assumption of an infinite population existing in an infinite region. The variables (X(t), Y(t)) can then be interpreted as the average population densities of the rabbits and foxes over the infinite region.

The Volterra model assumes that in the absence of foxes the rate of increase in the population density X(t) would be proportional to X(t), and that the rate at which foxes kill rabbits is proportional to X(t)Y(t). Applying the same sort of argument to the fox population leads to the following pair of differential equations for X(t) and Y(t)

$$\frac{d}{dt}X(t) = aX(t) - bX(t)Y(t),$$

$$\frac{d}{dt}Y(t) = cX(t)Y(t) - dY(t),$$

where a, b, c and d are positive.

It is shown in [2] that the trajectories are closed curves.

For the analogous family of Markov chains let V denote the area of the region in which the populations exist, let $X_{V}(t)$ be the number of rabbits and $Y_{V}(t)$ the number of foxes. Therefore $(V^{-1}X_{V}(t), V^{-1}Y_{V}(t))$ would be the average population densities. We assume that the probability of a particular rabbit being eaten by a fox in a short period of time is proportional to the density of the foxes and that the birthrate of the foxes is dependent upon the ease with which a fox catches a rabbit which is proportional to the density of the rabbits.

This leads to the following assumptions about the infinitesimal parameters of the Markov chains.

Let
$$\theta=(1,0)$$
 and $\phi=(0,1)$, then for $k\in\{(k_1,k_2)\colon k_1,k_2 \text{ non-negative integers}\}$
$$q_{k,k+\theta}=\lambda_1k_1\text{, birth of a rabbit,}$$

$$q_{k,k-\theta}=\left(\mu_1+b\frac{1}{V}k_2\right)k_1\text{, death of a rabbit}$$

$$q_{k,k+\phi}=\left(\lambda_2+c\frac{1}{V}k_1\right)k_2\text{, birth of a fox,}$$

 $q_{k,k-\phi} = \mu_2 k_2$, death of a fox,

where $\lambda_1 > \mu_1$ and $\lambda_2 < \mu_2$. We observe that these parameters correspond to a density dependent family with

$$f(x,\theta) = \lambda x_1$$

$$f(x,-\theta) = (\mu_1 + bx_2)x_1,$$

$$f(x,\phi) = (\lambda_2 + cx_1)x_2,$$

$$f(x,-\phi) = \mu_2 x_2$$

and

$$F_1(x) = (\lambda_1 - \mu_1)x_1 - bx_1x_2$$

$$F_2(x) = cx_1x_2 - (\mu_2 - \lambda_2)x_2.$$

Setting $a = \lambda_1 - \mu_1$ and $d = \mu_2 - \lambda_2$, we see that (3.5) is the system of differential equations that corresponds to the family. The conditions of Theorem (2.17) are easily verified for any bounded open set. Consequently,

$$\lim_{V \to \infty} \left(\frac{1}{V} X_V(0), \frac{1}{V} Y_V(0) \right) = (X(0), Y(0))$$

implies for every $\delta > 0$

$$\lim_{V\to\infty} P\left\{\sup_{s\leq t} \left| \frac{1}{V} X_V(s) - X(s) \right| + \left| \frac{1}{V} Y_V(s) - Y(s) \right| > \delta \right\} = 0.$$

4. Discrete time case

We state the following discrete time analogs of Proposition (2.1) and Theorem (2.11). The proofs are substantially the same using Theorem (2.13) of [3], in place of Theorem (2.1).

Proposition (4.1). Let $X_n(k)$ be a sequence of discrete time Markov processes, with measurable state spaces (E_n, \mathcal{B}_n) , $E_n \in \mathcal{B}^K$, and one step transition functions denoted by

$$\mu_n(x,\Gamma) = P\{X_n(k+1) \in \Gamma \mid X_n(k) = x\}.$$

Suppose there exist sequences of positive numbers α_n and ε_n

$$\lim_{n\to\infty} \alpha_n = \infty \quad \text{and} \quad \lim_{n\to\infty} \varepsilon_n = 0$$

such that

(4.2)
$$\sup_{n} \sup_{x \in E_{n}} \alpha_{n} \int_{E_{n}} |z-x| \mu_{n}(x,dx) < \infty$$

and

(4.3)
$$\lim_{n\to\infty} \sup_{x\in E_n} \alpha_n \int_{|z-x|>\varepsilon_n} |z-x| \mu_n(x,dz) = 0.$$

Let

$$F_n(x) = \alpha_n \int_{E_n} (z-x) \mu_n(x,dz).$$

Then for every $\delta > 0$, t > 0

$$\lim_{n\to\infty} \sup_{x\in E_n} P\left\{ \sup_{k\leq a_n t} \left| X_n(k) - X_n(0) - \sum_{l=0}^k \frac{1}{\alpha_n} F_n(X_n(k)) \right| > \delta \left| X_n(0) = x \right\} \right\} = 0.$$

Corollary (4.4). Suppose there exist $\Gamma_n \subset E_n$ such that

(4.5)
$$\sup_{n} \sup_{x \in \Gamma_{n}} \alpha_{n} \int_{E_{n}} |z - x| \, \mu_{n}(x, dz) < \infty$$

and

(4.6)
$$\lim_{n\to\infty} \sup_{x\in\Gamma_n} \alpha_n \int_{|z-x|>\varepsilon_n} |z-x| \mu_n(x,dz) = 0.$$

Let τ_{Γ_n} be the first exit time from Γ_n . Then for every $\delta>0$, t>0

$$\lim_{n\to\infty} \sup_{x\in E_n} P\left\{\sup_{k\leq \alpha_n t} \left| X_n(k\wedge \tau) - X_n(0) - \sum_{l=0}^{k-\tau-1} \frac{1}{\alpha_n} F_n(X_n(l)) \right| > \delta \left| X_n(0) = x \right\} \right.$$

Theorem (4.7). Suppose there exists $E \subset \mathbb{R}^K$, a function $F: E \to \mathbb{R}^K$ and a constant M such that

$$|F(x)-F(y)| \leq M|x-y|, \quad x,y \in E,$$

and

$$\lim_{n\to\infty} \sup_{x\in E_n\cap E} |F_n(x)-F(x)| = 0.$$

Let $X(s,x_0)$ be as in Theorem (2.11) and suppose the sequence satisfies the conditions of Corollary (4.4) for $\Gamma_n = E_n \cap E$. Then $\lim_{n\to\infty} X_n(0) = x_0$ implies for every $\delta > 0$, t > 0

$$\lim_{n\to\infty} P\{\sup_{s\leq t} |X_n([\alpha_n s]) - X(s,x_0)| > \delta\} = 0.$$

References

- [1] DYNKIN, E. B. (1965) Markov Processes I. Academic Press, New York, and Springer-Verlag, Berlin-Göttingen-Heidelberg.
- [2] KEMENY, J. G. AND SNELL, J. L. (1962) Mathematical Models in the Social Sciences. Ginn and Company, Boston.
- [3] Kurtz, T. G. (1969) Extensions of Trotter's operator semigroup approximation theorems. J. Functional Anal. 3, 111-132.