J. Math. Biol. (1987) 25: 643-652

Journal of
Mathematical
Biology

Weak convergence of a sequence of stochastic difference equations to a stochastic ordinary differential equation

Masaru Iizuka*

Department of Biology, Faculty of Science, Kyushu University, Fukuoka 812, Japan

Abstract. We consider a sequence of discrete parameter stochastic processes defined by solutions to stochastic difference equations. A condition is given that this sequence converges weakly to a continuous parameter process defined by solutions to a stochastic ordinary differential equation. Applying this result, two limit theorems related to population biology are proved. Random parameters in stochastic difference equations are autocorrelated stationary Gaussian processes in the first case. They are jump-type Markov processes in the second case. We discuss a problem of continuous time approximations for discrete time models in random environments.

Key words: Stochastic difference equation — Stochastic ordinary differential equation — Weak convergence in $D_{\mathbb{R}^n}[0, \infty)$ — Autocorrelated random environment — Degree of autocorrelation — Continuous time approximation — Diffusion approximation — Population biology

1. Introduction

In population genetics fundamental models are usually discrete time models, and the evolution of the system is described by a difference equation. Random fluctuation of environments introduces stochastic effects to the system and makes this equation a stochastic difference equation. In general, random fluctuation of environments is not independent between two distinct generations, but has autocorrelation. This implies that the analysis of population genetical models in random environments is that of stochastic difference equations whose parameters are autocorrelated stochastic processes.

It is difficult, however, to obtain in these discrete time models the explicit expression for biologically important quantities such as the distribution function of gene frequencies, average heterozygosity, and so on. To obtain them, continuous time approximations are frequently used for the original discrete time

* *Present address*: National Institute of Environmental Health Sciences, Research Triangle Park, NC 27709, USA

models. In spite of common use of the continuous time approximations, they have been introduced more or less heuristically, and this causes some confusion concerning biological consequences. There are two ways to get the approximating continuous time models. The first method is proceeding to a continuous approximation of difference equations by ordinary differential equations before adding random effects on environments (for example, see Sasaki and Iwasa 1987). The continuous time models are described by stochastic ordinary differential equations in this case. The second method is that the effect of random environments is added before the continuous time approximation. As pointed out by Gillespie and Guess (1978), in general these two methods do not produce the continuous time models of the same type. The faithful method to the original models is the second one, while the first method is very easy to use. In this paper, we consider the problem to find a condition under which the second method results in the same model as the first method. In other words, we will find the property of randomly fluctuating environments of the original discrete time models assumed implicitly when the first method is used.

We consider a sequence of discrete parameter stochastic processes $\{X_k^\varepsilon\}$ governed by stochastic difference equation of the form $X_{k+1}^\varepsilon - X_k^\varepsilon = F_\varepsilon(X_k^\varepsilon, Y_k^\varepsilon)$, where $\{Y_k^\varepsilon\}$ are so called driving processes and $F_\varepsilon(x, y)$ are appropriate functions described later. Our objective is to observe the asymptotic behavior of $\{X_k^\varepsilon\}$ as $\varepsilon \to 0$, changing the scale of the time.

Guess and Gillespie (1977) studied the case where $\{Y_k^\varepsilon\}$ are bounded uniformly mixing processes with mean 0 such that the mixing rate $\phi_\varepsilon(k)$ satisfies $\lim_{\varepsilon\to 0} \varepsilon\nu_\varepsilon = 0$ where $\nu_\varepsilon = \sum_{k=0}^{\infty} \phi_\varepsilon^{1/2}(k)$. Assuming that $F_\varepsilon(x, y)$ is a linear function of x and $F_\varepsilon(x, y) = \varepsilon f(x, y) + \varepsilon^2 g(x, y) + o(\varepsilon^2)$ where $f(x, y)$ and $g(x, y)$ are some nice functions, they showed that the continuous parameter process $x^\varepsilon(t) = X_{[t/\varepsilon^2\nu_\varepsilon]}^\varepsilon$ converges weakly to a diffusion process. For general $F_\varepsilon(x, y)$, Iizuka and Matsuda (1982) proved that $x^\varepsilon(t)$ converges weakly to a diffusion process (also see Kushner and Huang 1981; Seno and Shiga 1984 and Watanabe 1984). However, if $\lim_{\varepsilon\to 0} \varepsilon\nu_\varepsilon = \nu > 0$, then $\{x^\varepsilon(t)\}$ does not converge weakly to a diffusion process. Here we are interested in the characterization of this limit process.

In this paper, we shall consider the problem in a more general setting. Consider the stochastic difference equation introduced above. Assuming that there are scalings $\tau_\varepsilon > 0$, $\tau_\varepsilon \to 0$ as $\varepsilon \to 0$, such that $y^\varepsilon(t) = Y_{[t/\tau_\varepsilon]}^\varepsilon$ converges weakly to a stochastic process $y(t)$ and $F_\varepsilon(x, y)/\tau_\varepsilon$ converges to $F(x, y)$, we will show that $x^\varepsilon(t) = X_{[t/\tau_\varepsilon], \varepsilon>0}^\varepsilon$, converges weakly to solutions to a stochastic ordinary differential equation $dx/dt(t) = F(x(t), y(t))$. See Theorem 1 in Sect. 2 for a more precise statement. As an example of the above, we shall consider in Sect. 5 the case of Gillespie and Guess (1978) where the driving processes $\{Y_k^\varepsilon\}$ are strongly autocorrelated, i.e. $\lim_{\varepsilon\to 0} \varepsilon\nu_\varepsilon = \nu > 0$. It will be shown that $x^\varepsilon(t) = X_{[t/\varepsilon^2\nu_\varepsilon]}^\varepsilon$ converges weakly to solutions to a stochastic ordinary differential equation.

2. The main theorem

Let $\{Y_k^\varepsilon, -\infty < k < \infty\}$ be a sequence of $\mathbb{R}^n$-valued discrete parameter stochastic processes on a probability space $(\Omega, \mathscr{F}, P)$. Let $\{F_\varepsilon(x, y), x \in \mathbb{R}^m, y \in \mathbb{R}^n\}$ be a sequence of $\mathbb{R}^m$-valued functions on $\mathbb{R}^m \times \mathbb{R}^n$. We define a sequence of discrete parameter processes $\{X_k^\varepsilon, k \geq 0\}$ by solutions to the stochastic difference equation

$$X_{k+1}^\varepsilon - X_k^\varepsilon = F_\varepsilon(X_k^\varepsilon, Y_k^\varepsilon), \qquad k \geq 0, \tag{2.1}$$

and a given random vector X_0^ε.

On the other hand, let $y(t)$, $-\infty < t < \infty$, be an $\mathbb{R}^n$-valued continuous parameter process on $(\Omega, \mathcal{F}, P)$, and $F(x, y)$, $x \in \mathbb{R}^m$, $y \in \mathbb{R}^n$, be an $\mathbb{R}^m$-valued function on $\mathbb{R}^m \times \mathbb{R}^n$. We define a continuous parameter process $x(t)$, $t \geq 0$, by solutions to the stochastic ordinary differential equation (differential equation with random parameter)

$$\frac{dx(t)}{dt} = F(x(t), y(t)), \qquad t \geq 0, \tag{2.2}$$

and a given random vector $x(0)$, under some conditions on $F(x, y)$ described later.

Let $\{\tau_\varepsilon\}$ be a sequence of positive numbers such that $\tau_\varepsilon \to 0$ as $\varepsilon \to 0$. We define two sequences of continuous parameter processes $\{y^\varepsilon(t), t \geq 0\}$ and $\{x^\varepsilon(t), t \geq 0\}$ by

$$y^\varepsilon(t) = Y_{[t/\tau_\varepsilon]}^\varepsilon, \tag{2.3}$$

$$x^\varepsilon(t) = X_{[t/\tau_\varepsilon]}^\varepsilon, \tag{2.4}$$

for $t \geq 0$, where $[t]$ is the integer part of t. Sample paths of $y^\varepsilon(t)$ and $x^\varepsilon(t)$ belong to $D_{\mathbb{R}^n}[0, \infty]$ and $D_{\mathbb{R}^m}[0, \infty]$, respectively. Here, $D_{\mathbb{R}^r}[0, \infty)$ is the space of $\mathbb{R}^r$-valued functions on $[0, \infty)$ that are right-continuous and have left-hand limits, with the Skorohod topology (see Lindvall (1973)).

We consider the following conditions.

[A1] For each compact subset K of $\mathbb{R}^n$, there exists a positive constant L_K such that

$$\sup_{y \in K} |F(x_2, y) - F(x_1, y)| \leq L_K |x_2 - x_1|, \tag{2.5}$$

for $x_1, x_2 \in \mathbb{R}^m$. There exists a positive constant L such that

$$\sup_{x \in \mathbb{R}^m} |F(x, y_2) - F(x, y_1)| \leq L |y_2 - y_1|, \tag{2.6}$$

for $y_1, y_2 \in \mathbb{R}^n$.

[A2] For each compact subset K of $\mathbb{R}^n$,

$$\sup_{x \in \mathbb{R}^m, y \in K} |F_\varepsilon(x, y)/\tau_\varepsilon - F(x, y)| \to 0 \quad \text{as } \varepsilon \to 0. \tag{2.7}$$

[A3] $\{x^\varepsilon(0)\}$ and $\{y^\varepsilon(t), t \geq 0\}$ converge weakly to $x(0)$ and $y(t)$, $t \geq 0$, jointly in $\mathbb{R}^m \times D_{\mathbb{R}^n}[0, \infty)$.

Remark. Under [A1], (2.2) has one and only one solution (Coddington and Levinson 1955).

Now, we state a limit theorem.

Theorem 1. Assume [A1]~[A3]. *Then* $\{x^\varepsilon(t), t \geq 0\}$ *converges weakly to* $x(t)$, $t \geq 0$, *in* $D_{\mathbb{R}^m}[0, \infty)$.

3. Proof of Theorem 1

Since (2.2) is equivalent to

$$x(t)=\int_0^t F(x(s), y(s))\, ds+x(0), \tag{3.1}$$

$\{x(s), 0\leq s\leq t\}$ is determined by $\{y(s), 0\leq s\leq t\}$ and $x(0)$. We can express $\{x(s), 0\leq s\leq t\}=\Phi_t(\{y(s), 0\leq s\leq t\}, x(0))$. Φ_t is a mapping from $D_{\mathbb{R}^n}[0, t]\times\mathbb{R}^m$ to $C_{\mathbb{R}^m}[0, t]$, where $C_{\mathbb{R}^m}[0, t]$ is the space of $\mathbb{R}^m$-valued continuous functions on $[0, t]$ with the uniform convergence topology.

Lemma 1. *Assume* [A1]. *Then Φ_t is a continuous mapping for each fixed t.*

Proof. Let $\{y_k(s), 0\leq s\leq t\}$, $k=1, 2, \dots$, be a sequence from $D_{\mathbb{R}^n}[0, t]$ converging to $y(s)$, $0\leq s\leq t$, in the Skorohod topology. Then for any positive number γ and δ, there exist a positive N and a homeomorphism λ_k from $[0, t]$ onto itself such that $\sup_{0\leq s\leq t}|y(\lambda_k s)-y_k(s)|<\gamma$ and $\sup_{0\leq s\leq t}|\lambda_k s-s|<\delta$ hold for all $k\geq N$. Then we have

$$\sup_{k\geq N}\sup_{0\leq s\leq t}|y_k(s)|\leq\gamma+\sup_{0\leq s\leq t}|y(\lambda_k s)|<\infty. \tag{3.2}$$

Define $z_k(t)$ and $z(t)$ by

$$z_k(t)=\int_0^t F(z_k(s), y_k(s))\, ds+x_k, \tag{3.3}$$

$$z(t)=\int_0^t F(z(s), y(s))\, ds+x, \tag{3.4}$$

where $x_k\to x$ as $k\to\infty$ in $\mathbb{R}^m$. By [A1], if $k\geq N$,

$$\begin{aligned}|z_k(t)-z(t)|&<|x_k-x|+\int_0^t|F(z_k(s), y_k(s))-F(z(s), y(s))|\, ds\\ &\leq|x_k-x|+L\int_0^t|y_k(s)-y(s)|\, ds+L_{K_1}\int_0^t|z_k(s)-z(s)|\, ds,\end{aligned} \tag{3.5}$$

where $K_1=\{y\in\mathbb{R}^n;\ |y|\leq\gamma+\sup_{0\leq s\leq t}|y(s)|\}$. By Gronwall's lemma,

$$|z_k(t)-z(t)|\leq\left\{|x_k-x|+L\int_0^t|y_k(s)-y(s)|\, ds\right\}\exp(L_{K_1}t). \tag{3.6}$$

If s is a continuity point of y, then $y_k(s)$ converges to $y(s)$. Since the points of discontinuity of y are at most countable,

$$\int_0^t|y_k(s)-y(s)|\, ds\to 0\quad\text{as } k\to\infty. \tag{3.7}$$

Therefore z_k converges to z uniformly. The proof of the lemma is complete.

Proof of Theorem 1. For each fixed t, define $z^\varepsilon(t)$ by

$$z^\varepsilon(t)=\int_0^t F(z^\varepsilon(s), y^\varepsilon(s))\, ds+x^\varepsilon(0). \tag{3.8}$$

By Lemma 1 and [A3], $\{z^\varepsilon(s), 0\leq s\leq t\}$ converges weakly to $x(s)$, $0\leq s\leq t$, in $D_{\mathbb{R}^m}[0, t]$ (see Theorem 5.1 of Billingsley 1968).

On the other hand, note

$$x^{\varepsilon}(t) = X^{\varepsilon}_{[t/\tau_\varepsilon]} = \sum_{k=0}^{[t/\tau_\varepsilon]-1} F_\varepsilon(X^{\varepsilon}_k, Y^{\varepsilon}_k) + X^{\varepsilon}_0$$

$$= \frac{1}{\tau^{\varepsilon}} \int_0^{[t/\tau_\varepsilon]\tau_\varepsilon} F_\varepsilon(x^{\varepsilon}(s), y^{\varepsilon}(s))\, ds + x^{\varepsilon}(0). \tag{3.9}$$

By tightness of $\{y^{\varepsilon}(t)\}$ in $D_{\mathbb{R}^n}[0,\infty)$, for each η (>0), there exists a positive constant a such that $P(\sup_{0\leq s\leq t} |y^{\varepsilon}(s)| > a) \leq \eta$, $\varepsilon > 0$ (see Theorem 15.2 of Billingsley (1968)). On $\Omega^{\varepsilon}_\eta = \{\omega \in \Omega;\ \sup_{0\leq s\leq t} |y^{\varepsilon}(s)| \leq a\}$, we have

$$|z^{\varepsilon}(t) - x^{\varepsilon}(t)| \leq L_{K_2} \int_0^t |z^{\varepsilon}(s) - x^{\varepsilon}(s)|\, ds + t\Big\{ \sup_{x\in\mathbb{R}^m, y\in K_2} |F(x,y) - F_\varepsilon(x,y)/\tau_\varepsilon| \Big\}$$

$$+ \tau_\varepsilon \Big\{ \sup_{x\in\mathbb{R}^m, y\in K_2} |F(x,y)| \Big\}, \tag{3.10}$$

where $K_2 = \{y \in \mathbb{R}^n;\ |y| \leq a\}$. Using Gronwall's lemma, we get

$$|z^{\varepsilon}(t) - x^{\varepsilon}(t)| \leq \Big\{ t\Big(\sup_{x\in\mathbb{R}^m, y\in K_2} |F(x,y) - F_\varepsilon(x,y)/\tau_\varepsilon| \Big)$$

$$+ \tau_\varepsilon \Big(\sup_{x\in\mathbb{R}^m, y\in K_2} |F(x,y)| \Big) \Big\} \exp(L_{K_2} t), \tag{3.11}$$

on $\Omega^{\varepsilon}_\eta$. For simplicity, we assume that $\sup_{x\in\mathbb{R}^m; y\in K_2} |F(x,y)| < +\infty$. By [A2] and (3.11), for each γ (>0), there exists ε_0 (>0) such that if $\varepsilon < \varepsilon_0$, then

$$P\Big(\sup_{0\leq s\leq t} |z^{\varepsilon}(s) - x^{\varepsilon}(s)| > \gamma \Big) \leq P\Big(\sup_{0\leq s\leq t} |z^{\varepsilon}(s) - x^{\varepsilon}(s)| > \gamma;\ \Omega^{\varepsilon}_\eta \Big)$$

$$+ P(\Omega \backslash \Omega^{\varepsilon}_\eta) \leq 2\eta. \tag{3.12}$$

This implies that

$$\lim_{\varepsilon\to 0} P\Big(\sup_{0\leq s\leq t} |z^{\varepsilon}(s) - x^{\varepsilon}(s)| > \gamma \Big) = 0. \tag{3.13}$$

For general $F(x, y)$ satisfying [A1] and [A2], we can get the same estimation using [A1], tightness of $\{y^{\varepsilon}(s), 0 \leq s \leq t\}$, that of $\{x^{\varepsilon}(0)\}$, and Gronwall's lemma. By Theorem 4.1 of Billingsley (1968), we have proved the theorem.

4. Applications of Theorem 1

In this section, we consider two examples for Theorem 1.

Example 1: Stationary Gaussian processes

Gillespie and Guess (1978) considered the following problem. Let Y^{ε}_k, $-\infty < k < \infty$, be a stationary Gaussian process on $\mathbb{R}^1$ such that $E[Y^{\varepsilon}_k] = 0$ and $E[Y^{\varepsilon}_k Y^{\varepsilon}_{k+j}] = (1 - a\varepsilon)^j$, $j \geq 0$, where a is a positive constant. Put $s^{\varepsilon}_k = \sigma\varepsilon Y^{\varepsilon}_k + \mu\varepsilon$, where μ and σ (>0) are constants. Let $G(x, s)$ be a real valued smooth function on $\mathbb{R}^1 \times \mathbb{R}^1$ such that $G(x, 0) = 0$. Let X^{ε}_k, $k \geq 0$, be a process defined by solutions to

$$X^{\varepsilon}_{k+1} - X^{\varepsilon}_k = G(X^{\varepsilon}_k, s^{\varepsilon}_k), \qquad k \geq 0. \tag{4.1}$$

On the other hand, let $x(t)$, $t \geq 0$, be a real valued continuous parameter process defined by solutions to

$$\frac{dx(t)}{dt} = \{\mu + \sigma y(t)\} G_s(x(t), 0), \tag{4.2}$$

where $G_s = \partial G / \partial s$ and $y(t)$ is an Ornstein-Uhlenbeck process such that $E[y(t)] = 0$, $E[y^2(t)] = 1$ and autocorrelation function $\exp(-a|t|)$, that is, a diffusion process with infinitesimal generator $A = a(d^2/dy^2) - ay(d/dy)$.

Guess and Gillespie (1977) proved weak convergence of $x^\varepsilon(t) = X^\varepsilon_{[t/\varepsilon]}$ to $x(t)$ in the case where $G(x, y)$ is a linear function of x. Gillespie and Guess (1978) proposed a conjecture that $x^\varepsilon(t)$ converges weakly to $x(t)$ for general $G(x, s)$. We prove this conjecture as the first application of Theorem 1.

We consider the following conditions.

[A4] $G_s(x, 0)$ is Lipschitz continuous in x, and $\sup_{x \in \mathbb{R}^1} |G_s(x, 0)| < \infty$.

[A5] For each compact subset K of $\mathbb{R}^1$,

$$\sup_{x \in \mathbb{R}^1, s \in K} \left| \frac{\partial^2 G}{\partial s^2}(x, s) \right| < \infty.$$

Theorem 2. *Let X^ε_k and $x(t)$ be processes defined by solutions to* (4.1) *and* (4.2), *respectively. Assume* [A4], [A5], *and that X^ε_0 converges in probability to a constant $x(0)$. Then $x^\varepsilon(t) = X^\varepsilon_{[t/\varepsilon]}$ converges weakly to $x(t)$ in $D_{\mathbb{R}^1}[0, \infty)$.*

Proof. By (4.1), $G(x, 0) = 0$, and $s^\varepsilon_k = \sigma\varepsilon Y^\varepsilon_k + \mu\varepsilon$, we have (2.1), where

$$F_\varepsilon(x, y) = \varepsilon(\mu + \sigma y) G_s(x, 0) + \varepsilon^2 (\mu + \sigma y)^2 \frac{\partial^2 G}{\partial s^2}(x, \varepsilon h(\mu + \sigma y))/2, \qquad 0 \leq h \leq 1. \tag{4.3}$$

Let $F(x, y) = (\mu + \sigma y) G_s(x, 0)$ and $\tau_\varepsilon = \varepsilon$. If we prove [A1]~[A3], then we have the conclusion by Theorem 1. [A1] and [A2] is easily shown by [A4] and [A5]. Thus, it is enough to prove $y^\varepsilon(t) = Y^\varepsilon_{[t/\varepsilon]}$ converges weakly to $y(t)$ in $D_{\mathbb{R}^1}[0, \infty)$.

Since $y^\varepsilon(t)$ is a Gaussian process with $E[y^\varepsilon(t)] = 0$ and $E[y^\varepsilon(t) y^\varepsilon(s)] = (1 - a\varepsilon)^{[t/\varepsilon] - [s/\varepsilon]}$, $t \geq s$, we have $\lim_{\varepsilon \to 0} E[y^\varepsilon(t)] = 0$ and $\lim_{\varepsilon \to 0} E[y^\varepsilon(t) y^\varepsilon(s)] = \exp(-a(t - s))$, $t \geq s$, and the finite dimensional distribution of $y^\varepsilon(t)$ converges to that of $y(t)$. Using Schwarz's inequality and

$$\int_{-\infty}^{\infty} x^{2n} \exp(-cx^2)\, dx = C^1_n (\pi / c^{2n+1})^{1/2},$$

we have $E[|y^\varepsilon(t)|] \leq 1$ and

$$E[|y^\varepsilon(t) - y^\varepsilon(t_1)|^n |y^\varepsilon(t_2) - y^\varepsilon(t)|^n] < C^2_n (t_2 - t_1)^n,$$

for fixed $t_1 < t < t_2$, where C^l_n and C^2_n are constants. This estimation implies that $\{y^\varepsilon(t)\}$ is tight in $D_{\mathbb{R}^1}[0, \infty)$ by means of the tightness criterion by Kolmogorov-Chentsov (see Kunita 1986).

Example 2: *Jump-type Markov processes*

We consider jump-type Markov processes for Y^ε_k and $y(t)$. For each $\varepsilon > 0$, let $\lambda_\varepsilon(y)$ be a real valued measurable function on $\mathbb{R}^n$ such that $0 < \lambda_\varepsilon(y) < 1$. We denote the Borel field of $\mathbb{R}^n$ by $\mathcal{B}(\mathbb{R}^n)$. Let $\pi_\varepsilon(y, E)$ (resp. $\pi(y, E)$) be a real valued function on $\mathbb{R}^n \times \mathcal{B}(\mathbb{R}^n)$ such that $\pi_\varepsilon(y, \cdot)$ (resp. $\pi(y, \cdot)$) is a probability measure on $\mathscr{B}(\mathbb{R}^n)$ for fixed y, and $\pi_\varepsilon(\cdot, E)$ (resp. $\pi(\cdot, E)$) is a measurable function for fixed E such that $\pi_\varepsilon(y, \{y\}) = 0$ (resp. $\pi(y, \{y\}) = 0$). Let Y_k^ε, $-\infty < k < \infty$, be a Markov process on $\mathbb{R}^n$ with transition function

$$P(Y_{k+1}^\varepsilon \in E \mid Y_k^\varepsilon = y) = \begin{cases} 1 - \lambda_\varepsilon(y) & \text{if } E = \{y\} \\ \lambda_\varepsilon(y)\pi_\varepsilon(y, E) & \text{if } y \notin E \in \mathscr{B}(\mathbb{R}^n). \end{cases} \tag{4.4}$$

We define X_k^ε, $k \geq 0$, by (2.1) and a given random vector X_0^ε.

On the other hand, let $\lambda(y)$ be a real valued bounded measurable function on $\mathbb{R}^n$. Let $y(t)$, $-\infty < t < \infty$, be a continuous parameter jump-type Markov process on $\mathbb{R}^n$ with infinitesimal generator

$$Af(y) = \lambda(y)\left\{\int_{\mathbb{R}^n} f(z)\pi(y, dz) - f(y)\right\}. \tag{4.5}$$

We define $x(t)$, $t \geq 0$, by (2.2) and a given constant vector $x(0)$.

For simplicity, we consider the following case of compact state space.

[A6] There exists a compact subset K of $\mathbb{R}^n$ such that

$$\pi(y, K) = 1 \quad \text{for } y \in K, \text{ and } P(y(0) \in K) = 1.$$

By the same discussion as Stroock and Varadhan (1979), we have

Lemma 2. *Assume* [A6]. *Then the martingale problem for*

$$M_f(t) = f(y(t), t) - f(y(0), 0) - \int_0^t \left(\frac{\partial}{\partial s} + A\right) f(y(s), s)\, ds, \tag{4.6}$$

$f(y, s) \in C_b^{2,1}(\mathbb{R}^n \times [0, \infty))$, *has a unique solution, where* $C_b^{2,1}(\mathbb{R}^n \times [0, \infty))$ *is the space of* $C^{2,1}$*-class bounded functions on* $\mathbb{R}^n \times [0, \infty)$ *whose derivatives are bounded.*

To apply Theorem 1, we consider the following conditions.

[A7] $\lambda(y)$ is a bounded continuous function, and there exists a sequence of positive numbers $\{\tau_\varepsilon\}$ such that $\tau_\varepsilon \to 0$ and

$$\sup_y |\lambda(y) - \lambda_\varepsilon(y)/\tau_\varepsilon| \to 0 \quad \text{as } \varepsilon \to 0.$$

[A8] For each bounded continuous function $f(y)$ on $\mathbb{R}^n$ vanishing at infinity, $\int_{\mathbb{R}^n} f(z)\pi(y, dz)$ is continuous in y. For each C^∞-class function $g(y)$ on $\mathbb{R}^n$ with compact support,

$$\sup_y \left| \int_{\mathbb{R}^n} g(z)\pi(y, dz) - \int_{\mathbb{R}^n} g(z)\pi_\varepsilon(y, dz) \right| \to 0, \quad \text{as } \varepsilon \to 0.$$

Theorem 3. *Assume* [A1], [A2], [A6]~[A8]. *Further, assume that* Y_0^ε *converges weakly to* $y(0)$ *and* X_0^ε *converges in probability to a constant vector* $x(0)$. *Then,* $x^\varepsilon(t) = X_{[t/\tau_\varepsilon]}^\varepsilon$ *converges weakly to* $x(t)$ *in* $D_{\mathbb{R}^m}[0, \infty)$.

Proof. If we show that $y^\varepsilon(t) = Y_{[t/\tau_\varepsilon]}^\varepsilon$ converges weakly to $y(t)$ in $D_{\mathbb{R}^n}[0, \infty)$, we have the conclusion by Theorem 1. Indeed, we can prove this by the result of Kushner (1980). We give a brief outline of this.

For each C^∞-class function $f(y, t)$ on $\mathbb{R}^n \times [0, \infty)$ vanishing at infinity, let $f_\varepsilon(t) = f(y^\varepsilon(t), \tau_\varepsilon[t/\tau_\varepsilon])$. We define $\hat{A}_\varepsilon$ by

$$\hat{A}_\varepsilon f_\varepsilon(t) = \{E[f_\varepsilon(t + \tau_\varepsilon) \mid \mathscr{F}_t^\varepsilon] - f_\varepsilon(t)\}/\tau_\varepsilon, \tag{4.7}$$

where $\mathscr{F}_t^\varepsilon$ is the sub σ-field of $\mathscr{F}$ generated by $y^\varepsilon(s)$, $s \leq t$. By (4.6) and (4.7), we have the following inequalities after some calculation.

$$|f_\varepsilon(t) - f(y^\varepsilon(t), t)| \leq (t - \tau_\varepsilon[t/\tau_\varepsilon])\left\{\sup_{y,t}\left|\frac{\partial f}{\partial t}(y, t)\right|\right\}, \tag{4.8}$$

$$\begin{aligned}
&\left|\hat{A}_\varepsilon f_\varepsilon(t) - \left(\frac{\partial}{\partial t} + A\right) f(y^\varepsilon(t), t)\right| \\
&\leq \left\{\sup_y |\lambda(y)|\right\}\left\{\sup_y \left|\int_{\mathbb{R}^n} f(z, t)\pi_\varepsilon(y, dz) - \int_{\mathbb{R}^n} f(z, t)\pi(y, dz)\right|\right\} \\
&\quad + 2\left\{\sup_{y,t}|f(y, t)|\right\}\left\{\sup_y |\lambda(y) - \lambda_\varepsilon(y)/\tau_\varepsilon|\right\} \\
&\quad + 2(1 + [t/\tau_\varepsilon] - t/\tau_\varepsilon)\left\{\sup_y |\lambda_\varepsilon(y)|\right\}\left\{\sup_{y,t}\left|\frac{\partial f}{\partial t}(y, t)\right|\right\} \\
&\quad + (\tau_\varepsilon + t - \tau_\varepsilon[t/\tau_\varepsilon])\left\{\sup_{y,t}\left|\frac{\partial^2 f}{\partial t^2}(y, t)\right|\right\}\Big/ 2.
\end{aligned} \tag{4.9}$$

The right-hand sides of (4.8) and (4.9) tend to 0 as $\varepsilon \to 0$. By these inequalities, (3.1) and (3.2) of Kushner (1980) hold. Since $\{y^\varepsilon(t)\}$ is tight in $D_{\mathbb{R}}{}^n[0, \infty)$ by the following lemma, we have the conclusion by Theorem 3 of Kushner (1980).

Lemma 3. *Assume* [A6]. *Then* $\{y^\varepsilon(t)\}$ *is tight in* $D_{\mathbb{R}^n}[0, \infty)$.

Proof. For each C^∞-class function $f(y)$ on $\mathbb{R}^n$ vanishing at infinity, let $f^\varepsilon(t) = f(y^\varepsilon(t))$. Then, we can show that

$$\begin{aligned}
|\hat{A}_\varepsilon f^\varepsilon(t)| &= |E[f^\varepsilon(t + \tau_\varepsilon)|\mathscr{F}_t^\varepsilon] - f^\varepsilon(t)|/\tau_\varepsilon \\
&\leq 4\left\{\sup_y |\lambda(y)|\right\}\left\{\sup_y |f(y)|\right\}.
\end{aligned} \tag{4.10}$$

Therefore, the discrete parameter version of Theorem 2 of Kushner (1980) (see Sect. 4), we have the conclusion.

5. Relation to diffusion limits

Let $\{Y_k^\varepsilon, -\infty < k < \infty\}$ be a sequence of real valued bounded uniformly mixing (ϕ-mixing) processes with $E[Y_k^\varepsilon] = 0$ and mixing rate $\phi_\varepsilon(k)$ (see Sect. 20 of Billingsley (1968)). We assume that $\nu_\varepsilon = \sum_{k=0}^\infty \phi_\varepsilon^{1/2}(k)$ is finite for each $\varepsilon > 0$. ν_ε denotes the degree of autocorrelation of Y_k^ε. According to Guess and Gillespie (1977), we say $\{Y_k^\varepsilon\}$ has weak, moderate, or strong autocorrelation if $\sup_\varepsilon \nu_\varepsilon < \infty$, $\sup_\varepsilon \nu_\varepsilon = \infty$ and $\varepsilon\nu_\varepsilon \to 0$ as $\varepsilon \to 0$, or $\varepsilon\nu_\varepsilon \to \nu$, $0 < \nu < \infty$, as $\varepsilon \to 0$, respectively. Let $F_\varepsilon(x, y)$ be a real valued bounded function on $\mathbb{R}^1 \times \mathbb{R}^1$ such that

$$F_\varepsilon(x, y) = \varepsilon f(x, y) + \varepsilon^2 g(x, y) + o(\varepsilon^2), \tag{5.1}$$

where $f(x, y)$, $g(x, y)$ and $F_\varepsilon(x, y) - \varepsilon f(x, y) - \varepsilon^2 g(x, y)$ are bounded functions in x. We define $\{X_k^\varepsilon, k \geq 0\}$ by solutions to (2.1). Under additional conditions on $f(x, y)$ and $g(x, y)$, $x^\varepsilon(t) = X_{[t/\varepsilon^2\nu_\varepsilon]}$ converges weakly to a diffusion process in

$D_{\mathbb{R}^1}[0, \infty)$, if $\varepsilon\nu_\varepsilon \to 0$ as $\varepsilon \to 0$ (see Iizuka and Matsuda (1982), also see Guess and Gillespie (1977), Kushner and Huang (1981), Seno and Shiga (1984) and Watanabe (1984)), which implies that $\{x^\varepsilon(t)\}$ converges weakly to a diffusion process if $\{Y_k^\varepsilon\}$ has weak or moderate autocorrelation.

In this section, we construct an example that $\{Y_k^\varepsilon\}$ has strong autocorrelation and $\{x^\varepsilon(t)\}$ does not converge to a diffusion process, but converges to solutions to a stochastic ordinary differential equation.

Let Y_k^ε, $-\infty < k < \infty$, be a jump-type Markov process on $\{-1, 1\}$ with jump probability λ_ε $(0 < \lambda_\varepsilon < 1)$, that is, a Markov process on $\{-1, 1\}$ with $P(Y_{k+1}^\varepsilon = -y \mid Y_k^\varepsilon = y) = \lambda_\varepsilon$, $y = \pm 1$. We assume that

$$\lambda_\varepsilon / \varepsilon \to \alpha, \qquad 0 < \alpha < \infty \quad \text{as } \varepsilon \to 0. \tag{5.2}$$

Since Y_k^ε is a Markov process on a finite state space, Y_k^ε is a uniformly mixing process. Further, we can estimate ν_ε explicitly for this process. This is the reason for considering $\{-1, 1\}$ as state space. We show later in this section (see Lemma 4) that, for this Y_k^ε,

$$(1 - \lambda_\varepsilon)/2\lambda_\varepsilon \leq \nu_\varepsilon \leq \{1 + (1 - 2\lambda_\varepsilon)^{1/2}\}/2\lambda_\varepsilon. \tag{5.3}$$

From (5.2) and (5.3), we have $\varepsilon^2 \nu_\varepsilon \to 0$ as $\varepsilon \to 0$, and there exists $\nu = \lim_{\varepsilon \to 0} \varepsilon \nu_\varepsilon$, $0 < \nu < \infty$ (take a subsequence if necessary). This implies that $\{Y_k^\varepsilon\}$ has strong autocorrelation. We define X_k^ε, $k \geq 0$ by (2.1), (5.1), and a given random variable X_0^ε. We assume that $f(x, y)$ is Lipschitz continuous in x.

On the other hand, let $y(t)$, $-\infty < t < \infty$, be a continuous parameter jump-type Markov process on $\{-1, 1\}$ with jump rate $\lambda = \alpha/\nu$, that is, a Markov process on $\{-1, 1\}$ with infinitesimal generator $Af(y) = \lambda\{f(-y) - f(y)\}$, $y = \pm 1$. Let $F(x, y) = f(x, y)/\nu$. We define $x(t)$, $t \geq 0$, by solutions to stochastic ordinary differential equation (2.2) and a constant $x(0)$. By Theorem 3, we have the following result.

Corollary. *Assume that Y_0^ε converges weakly to $y(0)$ and X_0^ε converges in probability to a constant $x(0)$. Then $x^\varepsilon(t) = X_{[t/\varepsilon^2\nu_\varepsilon]}^\varepsilon$ converges weakly to $x(t)$ in $D_{\mathbb{R}^1}[0, \infty)$.*

This result provides an example that $\{Y_k^\varepsilon\}$ has strong autocorrelation and $\{X_{[t/\varepsilon^2\nu_\varepsilon]}^\varepsilon\}$ does not converge to a diffusion process, but converges weakly to solutions to a stochastic ordinary differential equation. This result implies that the strong autocorrelation of randomly fluctuating environments is implicitly assumed for the original discrete time models when the first method (mentioned in Sect. 1) to obtain the approximating continuous time models is used. For the analysis using stochastic ordinary differential equations, see Slatkin (1978), Matsuda and Ishii (1981), Ishii and Kitahara (1982), and Sasaki and Iwasa (1987).

Finally, we prove the following lemma.

Lemma 4. *Let Y_k^ε, $-\infty < k < \infty$, be a jump-type Markov process on $\{-1, 1\}$ with jump probability λ_ε. Then we have* (5.3).

Proof. Let $p_k^\varepsilon = P(Y_k^\varepsilon = y \mid Y_0^\varepsilon = y)$, $y = \pm 1$, $k \geq 0$. Since $p_1^\varepsilon = 1 - \lambda_\varepsilon$, and $p_{k+1}^\varepsilon = (1 - \lambda_\varepsilon) p_k^\varepsilon + \lambda_\varepsilon (1 - p_k^\varepsilon)$, $k \geq 0$, we have

$$p_k^\varepsilon = \{1 + (1 - 2\lambda_\varepsilon)^k\}/2, \qquad k \geq 0. \tag{5.4}$$

Therefore, $E[Y_0^\varepsilon Y_k^\varepsilon]=p_k^\varepsilon-(1-p_k^\varepsilon)=(1-2\lambda_\varepsilon)^k$, $k\geqslant 1$, and

$$\sum_{k=-\infty}^{\infty} E[Y_0^\varepsilon Y_k^\varepsilon]=E[(Y_0^\varepsilon)^2]+2\sum_{k=1}^{\infty} E[Y_0^\varepsilon Y_k^\varepsilon]=(1-\lambda_\varepsilon)/\lambda_\varepsilon. \tag{5.5}$$

By (20.35) of Billingsley (1968),

$$\sum_{k=-\infty}^{\infty} E[Y_0^\varepsilon Y_k^\varepsilon]\leqslant 2\nu_\varepsilon. \tag{5.6}$$

From (5.5) and (5.6), we have $(1-\lambda_\varepsilon)/2\lambda_\varepsilon\leqslant\nu_\varepsilon$.

On the other hand, the stationary probability p of Y_k^ε is $p(\{1\})=p(\{-1\})=1/2$. By (5.4) and (20.11) of Billingsley (1968), we have $\phi_\varepsilon(k)\leqslant(1-2\lambda_\varepsilon)^k$. Then, we get $\nu_\varepsilon\leqslant\{1+(1-2\lambda_\varepsilon)^{1/2}\}/2\lambda_\varepsilon$.

Acknowledgments. The author would like to express his thanks to Professor H. Kunita for valuable discussions. He thanks Professor K. Sato and the referees for helpful suggestions.

References

1. Billingsley, P.: Convergence of probability measures. New York: Wiley 1968
2. Coddington, E. A., Levinson, N.: Theory of ordinary differential equations. New York: McGraw-Hill 1955
3. Gillespie, J. H., Guess, H. A.: The effects of environmental autocorrelations on the progress of selection in a random environment. Am. Natur. **112**, 897-909 (1978)
4. Guess, H. A., Gillespie, J. H.: Diffusion approximations to linear stochastic difference equations with stationary coefficients. J. Appl. Probab. **14**, 58-74 (1977)
5. Iizuka, M., Matsuda, H.: Weak convergence of discrete time non-Markovian processes related to selection models in population genetics. J. Math. Biol. **15**, 107-127 (1982)
6. Ishii, K., Kitahara, K.: Relaxation of systems under the effect of a colored noise. Prog. Theor. Phys. **68**, 665-668 (1982)
7. Kunita, H.: Tightness of probability measures in $D([0, T]; C)$ and $D([0, T]; D)$. J. Math. Soc. Japan **38**, 309-334 (1986)
8. Kushner, H. J.: A martingale method for the convergence of a sequence of processes to a jump-diffusion process. Z. Wahrschlenlichtentheor. Verw. Geb. **53**, 207-219 (1980)
9. Kushner, H. J., Huang, H.: On the weak convergence of a sequence of general stochastic difference equations to a diffusion. SIAM J. Appl. Math. **40**, 528-541 (1981)
10. Lindvall, T.: Weak convergence of probability measures and random functions in the function space $D[0,\infty)$. J. Appl. Probab. **10**, 109-121 (1973)
11. Matsuda, H., Ishii, K.: Stationary gene frequency distribution in the environment fluctuating between two distinct states. J. Math. Biol. **11**, 119-141 (1981)
12. Sasaki, A., Iwasa, Y.: Optimal recombination rate in fluctuating environments. Genetics **115**, 377-388 (1987)
13. Seno, S., Shiga, T.: Diffusion models of temporally varying selection in population genetics. Adv. Appl. Probab. **16**, 260-280 (1984)
14. Slatkin, M.: The dynamics of a population in a Markovian environment. Ecology **59**, 249-256 (1978)
15. Stroock, D. W., Varadhan, S. R. S.: Multidimensional Diffusion Processes. Berlin Heidelberg New York: Springer 1979
16. Watanabe, H.: Diffusion approximations of some stochastic difference equation II. Hiroshima Math. J. **14**, 15-34 (1984)

Received April 16/Revised August 19, 1987