J. Math. Biology (1982) 15: 107–127

Journal of
Mathematical
Biology

Weak Convergence of Discrete Time Non-Markovian Processes Related to Selection Models in Population Genetics

Masaru Iizuka and Hirotsugu Matsuda

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

Abstract. We consider discrete time stochastic processes defined by solutions to some non-linear difference equations whose coefficients are autocorrelated random sequences. It is proved that these processes converge weakly in $D[0, T]$ to diffusion processes, under the assumption that the random sequences satisfy some mixing condition. Diffusion approximation for stochastic selection models in population genetics is discussed, as the application of this limit theorem.

Key words: Non-linear stochastic difference equation – Weak convergence in $D[0, T]$ – Convergence of finite-dimensional distribution – ϕ-mixing process – Diffusion approximation – Stochastic selection – Population genetics

1. Introduction

The change of gene frequency is usually described by a discrete time model in population genetics. Stochastic selection and size (random sampling) effect are the main stochastic factors for gene frequency change. Generally, the stochastic selection is autocorrelated from generation to generation.

Experimentally observable quantities in population genetics are the distribution of gene frequency, moments of gene frequency, and so on. It is difficult, however, to obtain in the discrete time model the explicit expression for these quantities. To obtain it, so-called "diffusion model" (diffusion approximation) is frequently made use of for the original discrete time model. The explicit expression for the above quantities can then be obtained in the diffusion model with the help of the theory of diffusion processes; here, Kolmogorov forward and backward equations are extensively used [Kimura (1955, 1964); Jensen and Pollak (1969); Crow and Kimura (1970); Kimura and Ohta (1971); Gillespie (1973a, b; 1978); Hartl and Cook (1973, 1974); Cook and Hartl (1974); Karlin and Levikson (1974); Karlin and Lieberman (1974); Levikson and Karlin (1975); Takahata et al. (1975); Hartl (1977); Li (1977); Maruyama (1977); Ewens (1979)].

However, in these treatises the diffusion approximation has been introduced more or less heuristically. As to the adequacy of the diffusion approximation, Feller

0303–6812/82/0015/0107/$04.20

(1951) studied for the first time how a passage from a Markov chain of gene frequencies to a diffusion process could be effected. For such discrete time Markov processes the convergence to the corresponding diffusion processes has been proved for many cases where the stochastic factor is size effect [Trotter (1958); Watterson (1962); Guess (1973); Sato (1976a, b, c; 1978)] and/or stochastic selection without autocorrelation [Okada (1979)]. On the other hand, rigorous results are few for the case of autocorrelated stochastic selection, where gene frequencies of a discrete time model do not form a Markov process. For results of weak convergence of a sequence of continuous time non-Markovian processes to a diffusion process, see Papanicolaou and Kohler (1974) and Kurtz (1975). Norman (1975) studied the diffusion approximation of discrete time non-Markovian processes, but his result is not suitable to the autocorrelated stochastic selection. Guess and Gillespie (1977) proved weak convergence of the discrete time processes defined by solutions to linear first-order stochastic difference equations of the form

$$X_{k+1} - X_k = A(s_k) + B(s_k)X_k \qquad (k = 0, 1, 2, \ldots), \tag{1.1}$$

where A and B are functions of an autocorrelated stochastic variable s_k.

However, the difference equation of stochastic selection in population genetics is usually non-linear. For instance, consider a haploid model in which the genotypes A and a have fitness $1 + s_k$ and 1 in the kth generation. In this model the change of the gene frequency of A allele is given by

$$X_{k+1} - X_k = \frac{s_k X_k(1 - X_k)}{1 + s_k X_k}, \tag{1.2}$$

where X_k is the gene frequency in the kth generation. For such non-linear cases, Gillespie and Guess (1978) heuristically gave the form of the limiting diffusion processes by patching together of linear processes.

In this paper, we consider discrete time stochastic processes defined by solutions to some non-linear first-order stochastic difference equations such as (1.2). We prove weak convergence of the processes assuming that the sequence of bounded random variables $\{s_k\}$ is a certain ϕ-mixing process with mean zero. The definition of the ϕ-mixing process and the related inequalities are given in §2. In §3 we consider discrete time processes $\{Z_k^{(n)}\}$ defined by solutions to

$$Z_{k+1}^{(n)} - Z_k^{(n)} = \varepsilon_n s_k^{(n)} + \varepsilon_n^2 G(Z_k^{(n)}, s_k^{(n)}, \varepsilon_n). \tag{1.3}$$

Here, $\{s_k^{(n)}\}$ $(n = 1, 2, \ldots)$ is a sequence of ϕ-mixing processes with mixing rate $\phi_n(k)$ such that $v_n = \sum_{k=0}^{\infty} \phi_n^{1/2}(k) < \infty$. The sequences of positive parameters ε_n and v_n satisfy the condition $\lim_{n\to\infty} \varepsilon_n v_n = 0$. We prove, under certain conditions on the function $G(Z_k^{(n)}, s_k^{(n)}, \varepsilon_n)$, weak convergence of $z^{(n)}(t) = Z_{[t/\varepsilon_n^2 v_n]}^{(n)}$ to some diffusion process, applying a theorem of Gikhman [Gikhman (1969)].

In §4, we consider the more general case of discrete time processes $\{X_k^{(n)}\}$, which are defined by solutions to

$$X_{k+1}^{(n)} - X_k^{(n)} = \varepsilon_n f(X_k^{(n)})\{s_k^{(n)} + \varepsilon_n g(X_k^{(n)}, s_k^{(n)}, \varepsilon_n)\}, \tag{1.4}$$

where $f(x)$ and $g(x, s, \varepsilon)$ are certain functions of x and (x, s, ε). Using the results of §3, we also prove weak convergence of $x^{(n)}(t) = X_{[t/\varepsilon_n^2 v_n]}^{(n)}$ to some diffusion process.

The limiting properties of $\{X_k\}$ of (1.2) can then be studied as a special case of that of $\{X_k^{(n)}\}$ of (1.4). In §5 we apply our result to stochastic selection models in population genetics.

2. Preliminaries

2.1 ϕ-Mixing Processes and Inequalities for Moments

Let $\{s_k(\omega)\}_{k=-\infty}^{\infty}$ be a strictly stationary sequence of random variables on a probability space $(\Omega, \mathscr{F}, P)$. Let $\mathscr{F}_{k_1}^{k_2}$ be a σ-algebra generated by $\{s_k\}_{k=k_1}^{k_2}$ $(-\infty \leqslant k_1 \leqslant k_2 \leqslant \infty)$. Consider a function ϕ: $\{0, 1, 2, \ldots\} \to [0, 1]$ such that $1 = \phi(0) \geqslant \phi(1) \geqslant \cdots$, $\lim_{m\to\infty} \phi(m) = 0$.

Definition. $\{s_k\}_{k=-\infty}^{\infty}$ is ϕ-mixing (uniformly mixing) if, for each k $(-\infty < k < \infty)$ and for each m $(m > 0)$

$$\sup\{|P(E_2|E_1) - P(E_2)|;\ E_1 \in \mathscr{F}_{-\infty}^{k},\ E_2 \in \mathscr{F}_{k+m}^{\infty}\} = \phi(m). \tag{2.1}$$

Here we regard $P(E_2|E_1) - P(E_2) = 0$, if $P(E_1) = 0$ [see Billingsley (1968) p. 166].

In this section, we assume that $\{s_k\}_{k=-\infty}^{\infty}$ is a strictly stationary ϕ-mixing process with $v = \sum_{k=0}^{\infty} \phi^{1/2}(k) < \infty$, $S = \sup_{k,\omega} |s_k(\omega)| < \infty$ and $E\{s_k\} = 0$. $E\{x\}$ means the expectation of x. For this process we get the following inequalities.

Lemma 2.1.

$$\sum_{k=0}^{\infty} \phi(k) \leqslant v, \tag{2.2}$$

$$\sum_{k=1}^{\infty} k\phi(k) \leqslant (v-1)^2. \tag{2.3}$$

Proof. The first inequality is obvious since $0 \leqslant \phi(k) \leqslant 1$. For each integer n $(n \geqslant 1)$

$$\sum_{k=1}^{n} k\phi(k) = \sum_{j=1}^{n} \sum_{k=j}^{n} \phi(k) \leqslant \sum_{j=1}^{n} \phi^{1/2}(j) \sum_{k=j}^{n} \phi^{1/2}(k) \leqslant (v-1)^2. \tag{2.4}$$

This leads to the second inequality.

Lemma 2.2. *If ξ is $\mathscr{F}_{-\infty}^{k}$-measurable and $|\xi| \leqslant C_1 < \infty$ and η is $\mathscr{F}_{k+m}^{\infty}$-measurable $(m \geqslant 0)$ and $|\eta| \leqslant C_2 < \infty$, then for each k_1 and k_2 $(-\infty \leqslant k_1 \leqslant k_2 \leqslant k)$*

$$|E\{\xi\eta|\mathscr{F}_{k_1}^{k_2}\} - E\{\xi|\mathscr{F}_{k_1}^{k_2}\}E\{\eta\}| \leqslant 2C_1C_2\phi(m). \tag{2.5}$$

Proof. Since we can approximate ξ and η by simple random variables, in treating the general case we may suppose that

$$\xi = \sum_i u_i \chi_{A_i}, \tag{2.6}$$

$$\eta = \sum_j v_j \chi_{B_j}, \tag{2.7}$$

where $\{A_i\}$ $(\{B_j\})$ is a finite decomposition of Ω into elements of $\mathscr{F}_{-\infty}^{k}$ $(\mathscr{F}_{k+m}^{\infty})$. χ_A is the indicator function of A. In this case

$$E\{\eta\} = \sum_j v_j P(B_j), \tag{2.8}$$

$$E\{\xi|\mathscr{F}_{k_1}^{k_2}\} = \sum_i u_i P(A_i|\mathscr{F}_{k_1}^{k_2}), \tag{2.9}$$

$$E\{\xi\eta|\mathscr{F}_{k_1}^{k_2}\} = \sum_{i,j} u_i v_j P(A_i B_j|\mathscr{F}_{k_1}^{k_2}). \tag{2.10}$$

For each $F \in \mathscr{F}_{k_1}^{k_2}$, $P(F) > 0$ we have

$$\begin{aligned}
&\left|\sum_{i,j} u_i v_j P(A_i B_j|F) - \left\{\sum_i u_i P(A_i|F)\right\}\left\{\sum_j v_j P(B_j)\right\}\right| \\
&= \left|\sum_{i,j} u_i v_j P(A_i|F)P(B_j|A_iF) - \sum_{i,j} u_i v_j P(A_i|F)P(B_j)\right| \\
&\leqslant C_1 C_2 \sum_i P(A_i|F) \sum_j |P(B_j|A_iF) - P(B_j)|.
\end{aligned} \tag{2.11}$$

As $A_iF \in \mathscr{F}_{-\infty}^{k}$ and $B_j \in \mathscr{F}_{k+m}^{\infty}$, the proof of Lemma 20.1 of Billingsley (1968) implies

$$\sum_j |P(B_j|A_iF) - P(B_j)| \leqslant 2\phi(m). \tag{2.12}$$

Therefore we have

$$\left|\sum_{i,j} u_i v_j P(A_i B_j|F) - \left\{\sum_i u_i P(A_i|F)\right\}\left\{\sum_j v_j P(B_j)\right\}\right| \leqslant 2C_1 C_2 \phi(m). \tag{2.13}$$

This leads to (2.5).

Lemma 2.3. *For each i, j and k ($k \leqslant i \leqslant j$)*

(i) $$|E\{s_i s_j\}| \leqslant 2S^2\phi(j-i), \tag{2.14}$$

(ii) $$|E\{s_i s_j|\mathscr{F}_{-\infty}^{k}\}| \leqslant 2S^2\phi(j-i), \tag{2.15}$$

(iii) $$|E\{s_i s_j|\mathscr{F}_{-\infty}^{k}\} - E\{s_i s_j\}| \leqslant 4S^2\phi((i-k)^v(j-i)), \tag{2.16}$$

where $m^v n = \max(m, n)$.

Proof. As $E\{s_j\} = 0$, Lemma 20.2 of Billingsley (1968) and Lemma 2.2 lead to (i) and (ii) immediately. We show (iii). If $j - i \geqslant i - k$, we have

$$\begin{aligned}
|E\{s_i s_j|\mathscr{F}_{-\infty}^{k}\} - E\{s_i s_j\}| &\leqslant |E\{s_i s_j|\mathscr{F}_{-\infty}^{k}\}| + |E\{s_i s_j\}| \\
&\leqslant 4S^2\phi(i-j),
\end{aligned} \tag{2.17}$$

using (i) and (ii). If $j - i < i - k$, we have

$$|E\{s_i s_j|\mathscr{F}_{-\infty}^{k}\} - E\{s_i s_j\}| \leqslant 2S^2\phi(i-k). \tag{2.18}$$

Here we used Lemma 2.2 for $1 \in \mathscr{F}_{-\infty}^{k}$ and $s_i s_j \in \mathscr{F}_i^{\infty}$.

Lemma 2.4. *For each k and m ($-\infty < k < \infty$, $m \geqslant 1$)*

$$\left|E\left\{\left(\sum_{j=1}^{m} s_{k+j}\right)^2 \middle| \mathscr{F}_{-\infty}^{k}\right\} - E\left\{\left(\sum_{j=1}^{m} s_{k+j}\right)^2\right\}\right| \leqslant 16S^2(v-1)^2. \tag{2.19}$$

Proof. By an inequality

$$\left(\sum_{j=1}^{m} s_{k+j}\right)^2 \leqslant 2\sum_{j=1}^{m}\sum_{i=0}^{m-j} s_{k+j}s_{k+j+i} \tag{2.20}$$

and (iii) of Lemma 2.3, we have

$$\left|E\left\{\left(\sum_{j=1}^{m} s_{k+j}\right)^2 \middle| \mathscr{F}_{-\infty}^{k}\right\} - E\left\{\left(\sum_{j=1}^{m} s_{k+j}\right)^2\right\}\right| \leqslant 8S^2 \sum_{j=1}^{m}\sum_{i=0}^{m-j} \phi(i^v j)$$

$$\leqslant 16S^2 \sum_{\substack{i,j=1\\ i\leqslant j}}^{m} \phi(j) = 16S^2 \sum_{j=1}^{m} j\phi(j) \leqslant 16S^2(v-1)^2. \tag{2.21}$$

Lemma 2.5. *For each k and m* $(-\infty < k < \infty,\ m \geqslant 1)$

$$\left|E\left\{\left(\sum_{j=1}^{m} s_{k+j}\right)^2\right\} - mV\right| \leqslant 4S^2(v-1)^2, \tag{2.22}$$

where

$$V = \sum_{j=-\infty}^{\infty} E\{s_0 s_j\}. \tag{2.23}$$

Proof. By stationarity, we have

$$\left|E\left\{\left(\sum_{j=1}^{m} s_{k+j}\right)^2\right\} - mV\right| = \left|\sum_{j=-m}^{m} (m-|j|)E\{s_0 s_j\} - m\sum_{j=-\infty}^{\infty} E\{s_0 s_j\}\right|$$

$$\leqslant 2m \sum_{j=m+1}^{\infty} |E\{s_0 s_j\}| + 2\sum_{j=1}^{m} j|E\{s_0 s_j\}|$$

$$\leqslant 4mS^2 \sum_{j=m+1}^{\infty} \phi(j) + 4S^2 \sum_{j=1}^{m} j\phi(j) \leqslant 4S^2(v-1)^2. \tag{2.24}$$

Corollary 2.6.

$$0 \leqslant V \leqslant 4S^2 v. \tag{2.25}$$

Proof. By Lemma 2.5, we have

$$\left|E\left\{\left(\sum_{j=1}^{m} s_{k+j}\right)^2\right\}\Big/ m - V\right| \leqslant 4S^2(v-1)^2/m. \tag{2.26}$$

As $v < \infty$, (2.26) means $V \geqslant 0$.

On the other hand, by stationarity, we have

$$|V| \leqslant 2\sum_{j=0}^{\infty} |E\{s_0 s_j\}| \leqslant 4S^2 \sum_{j=0}^{\infty} \phi(j) \leqslant 4S^2 v. \tag{2.27}$$

By Lemma 2.4 and Lemma 2.5, we have the following proposition.

Lemma 2.7. *For each k and m $(-\infty < k < \infty, m \geqslant 1)$*

$$\left| E\left\{\left(\sum_{j=1}^{m} s_{k+j}\right)^2 \middle| \mathscr{F}^k_{-\infty}\right\} - mV\right| \leqslant 20S^2(v-1)^2. \tag{2.28}$$

The following inequality is obvious from Lemma 20.4 of Billingsley (1968).

Lemma 2.8. *For each positive δ,*

$$P\left(\left|\sum_{k=1}^{m} s_k\right| > \delta\right) \leqslant (cS/\delta)^4 m^2 v^2 \qquad (m \geqslant 1), \tag{2.29}$$

where $c > 0$ and $c^4 = 768$.

The following inequality estimates the maximum of the partial sums.

Lemma 2.9. *For each m $(m \geqslant 1)$,*

$$E\left\{\left[\max_{k \leqslant m}\left(\left|\sum_{j=1}^{k} s_j\right|\right)\right]^4\right\} \leqslant KS^4 m^2 v^2, \tag{2.30}$$

where K is a positive constant.

Proof. Let

$$t_j = s_j/Sv^{1/2} \qquad (1 \leqslant j \leqslant m), \tag{2.31}$$

then by Lemma 20.4 of Billingsley (1968), we have

$$E\left\{\left(\sum_{j=1}^{m} t_j\right)^4\right\} \leqslant c^4 m^2. \tag{2.32}$$

Therefore, by Corollary B1 of Serfling (1970),

$$E\left\{\left[\max_{k \leqslant m}\left(\left|\sum_{j=1}^{k} t_j\right|\right)\right]^4\right\} \leqslant Km^2, \tag{2.33}$$

where K is a constant. Substituting (2.31) into (2.33), we have (2.30).

Corollary 2.10. *For each positive δ,*

$$P\left(\max_{k \leqslant m}\left(\left|\sum_{j=1}^{k} s_j\right|\right) > \delta\right) \leqslant KS^4 m^2 v^2/\delta^4 \qquad (m \geqslant 1), \tag{2.34}$$

where K is a positive constant.

2.2 A Criterion for the Convergence of the Finite-Dimensional Distribution

Gikhman (1969) gave a criterion for the convergence of the finite-dimensional distribution of a sequence of $\mathbb{R}^d$-valued random processes to that of a diffusion process. We consider the simple case of $d = 1$ and time homogeneity.

Let $\{y_n(t), (t \in [0, T])\}$ be a sequence of real-valued random processes satisfying the following conditions.

(1) There exist a sequence of families of monotonically increasing σ-algebras $\{\mathscr{F}_n(t), (t \in [0, T])\}$, a sequence of decompositions $\{t_{nk}, k = 0, 1, 2, \ldots, m_n;$

$n = 1, 2, \ldots\}$ of the interval $[0, T]$, $0 = t_{n0} < t_{n1} < \cdots < t_{nm_n} = T$, $\max_k \Delta t_{nk} \to 0$ $(n \to \infty)$, and real-valued continuous functions $a(z)$ and $b(z)$, $(z \in \mathbb{R})$ such that

a) $y_n(t)$ is $\mathscr{F}_n(t)$-measurable for any $t \in [0, T]$, $n = 1, 2, \ldots$,

b) $$P(|\Delta y_{nk}| > \delta | \mathscr{F}_n(t_{nk})) = \rho'_{nk} \Delta t_{nk}, \tag{2.35}$$

c) $$E\{\chi_\delta(\Delta y_{nk}) \Delta y_{nk} | \mathscr{F}_n(t_{nk})\} = b(y_n(t_{nk})) \Delta t_{nk} + \rho''_{nk} \Delta t_{nk}, \tag{2.36}$$

d) $$E\{\chi_\delta(\Delta y_{nk})(\Delta y_{nk})^2 | \mathscr{F}_n(t_{nk})\} = a(y_n(t_{nk})) \Delta t_{nk} + \rho'''_{nk} \Delta t_{nk}, \tag{2.37}$$

where $\Delta t_{nk} = t_{n(k+1)} - t_{nk}$, $\Delta y_{nk} = y_n(t_{n(k+1)}) - y_n(t_{nk})$, $\chi_\delta(x) = 1$, if $|x| < \delta$, and $\chi_\delta(x) = 0$ otherwise, δ is an arbitrarily fixed positive constant.

(2) The functions $b(z)$ and $c(z) = a^{1/2}(z)$, together with their derivatives up to fourth order inclusive, are continuous and bounded.

(3) $$\sum_{k=0}^{m_n - 1} E\{\rho'_{nk} + |\rho''_{nk}| + |\rho'''_{nk}|\} \Delta t_{nk} \to 0 \quad \text{as} \quad n \to \infty. \tag{2.38}$$

(4) The distribution of $y_n(0)$ converges weakly to that of a random variable y_0.

Theorem 2.11 (Gikhman). *Assume* (1), (2), (3) *and* (4), *then the finite-dimensional distribution of* $y_n(t)$, $(t \in [0, T])$ *converges weakly to that of a diffusion process with diffusion term* $a(z)$ *and drift term* $b(z)$. *Its initial distribution is the distribution of* y_0.

3. The Case Corresponding to Diffusion Processes with Constant Diffusion Term

In this section we consider weak convergence of a sequence of real-valued discrete time stochastic processes $\{Z_k^{(n)}, k = 0, 1, 2, \ldots\}_{n=1}^\infty$, defined by solutions to (1.3). Considering (1.3), we introduce the following conditions:

[S] $\{s_k^{(n)}\}_{k=-\infty}^\infty$ $(n = 1, 2, \ldots)$ is a sequence of strictly stationary ϕ-mixing processes, on a probability space $(\Omega, \mathscr{F}, P)$, with $v_n = \sum_{k=0}^\infty \phi_n^{1/2}(k) < \infty$, $S = \sup_{n,k,\omega} |s_k^{(n)}(\omega)| < \infty$ and $E\{s_k^{(n)}\} = 0$ $(k = 0, 1, 2, \ldots, n = 1, 2, \ldots)$. Here, $\phi_n(k)$ is mixing rate of the process $\{s_k^{(n)}\}$. $\{v_n\}_{n=1}^\infty$ is a non-decreasing sequence. $s_0^{(n)}$ $(n = 1, 2, \ldots)$ have a common distribution.

[E] ε_n and $\varepsilon_n v_n$ tend to 0 as n tends to ∞. Let $V_n = \sum_{k=-\infty}^\infty E\{s_0^{(n)} s_k^{(n)}\}$ $(n = 1, 2, \ldots)$, then there exists a positive constant $\tilde{V} = \lim_{n\to\infty} V_n / v_n$.

Remark. If $\tilde{V}$ exists, then $0 \leqslant \tilde{V} \leqslant 4S^2$ is obvious by Corollary 2.6.

[I] $G(z, s, \varepsilon)$ is a function on $\mathbb{R} \times [-S, S] \times (0, 1)$ satisfying

(i) $$G(z, s, \varepsilon) = H(z, s) + R(z, s, \varepsilon). \tag{3.1}$$

Here $R(z, s, \varepsilon)$ is a uniformly bounded function, and $\lim_{\varepsilon \to 0} \tilde{R}(\varepsilon) = 0$, where $\tilde{R}(\varepsilon) = \sup_{z,s} |R(z, s, \varepsilon)|$.

(ii) $H(z, s)$ is fourth continuously differentiable with respect to z, and together with its derivatives up to fourth order inclusive are continuous and uniformly bounded.

(iii) $H(z, s)$ is continuous and uniformly bounded on $\mathbb{R} \times [-S, S]$.

[II] $Z_0^{(n)} \in \mathbb{R}$ $(n = 1, 2, \ldots)$ are random variables satisfying

(i) $Z_0^{(n)}$ is independent of $\{s_k^{(n)}\}_{k=-\infty}^{\infty}$ $(n = 1, 2, \ldots)$.
(ii) $Z_0^{(n)}$ converges weakly to a random variable $z_0 \in \mathbb{R}$ as $n \to \infty$.

First, we prove the convergence of the finite-dimensional distribution applying Theorem 2.11. From here on, let T be an arbitrarily fixed positive constant.

Theorem 3.1. *Assume* [S], [E], [I] *and* [II]. *Define a sequence of continuous time stochastic processes* $\{z^{(n)}(t),\ t \in [0, T]\}_{n=1}^{\infty}$, *by*

$$z^{(n)}(t) = Z^{(n)}_{[t/\varepsilon_n^2 v_n]}, \tag{3.2}$$

where $[x]$ *is the largest integer that does not exceed* x.

Then the finite-dimensional distribution of $z^{(n)}(t)$, $(t \in [0, T])$ *converges weakly to that of a diffusion process* $z(t)$, $(t \in [0, T])$ *whose infinitesimal generator* $\mathscr{G}$ *is*

$$\mathscr{G} = \frac{\tilde{V}}{2}\frac{\partial^2}{\partial z^2} + \mu H_s(z)\frac{\partial}{\partial z}, \tag{3.3}$$

and the distribution of $z(0)$ *is that of* z_0. *Here we put*

$$\mu = \lim_{n\to\infty} \frac{1}{v_n}, \tag{3.4}$$

and

$$H_s(z) = E\{H(z, s_0^{(n)}(\omega))\}, \qquad (z \in \mathbb{R}). \tag{3.5}$$

Remark. $H_s(z)$ is fourth continuously differentiable, and together with its derivatives up to fourth order inclusive are uniformly bounded, since $H(z, s)$ satisfies (ii) of [I].

Proof. Let $\mathscr{F}_m^{(n)}$ $(m \geqslant 0, n \geqslant 1)$ be a σ-algebra generated by $\{s_k^{(n)}\}_{k=0}^m$ and $Z_0^{(n)}$, and $\mathscr{F}_{-1}^{(n)}$ be a σ-algebra generated by $Z_0^{(n)}$. $Z_k^{(n)}$ is $\mathscr{F}_{k-1}^{(n)}$-measurable by the definition of $Z_k^{(n)}$. Let

$$\mathscr{F}_n(t) = \mathscr{F}^{(n)}_{[t/\varepsilon_n^2 v_n]-1} \qquad (n \geqslant 1). \tag{3.6}$$

$\{\mathscr{F}_n(t),\ t \in [0, T]\}_{n=1}^{\infty}$ is a sequence of monotonically increasing σ-algebras, and $z^{(n)}(t)$ is $\mathscr{F}_n(t)$-measurable.

Next, we define a sequence of decompositions $\{t_{nk}, k = 0, 1, 2, \ldots, m_n; n \geqslant 1\}$ of the interval $[0, T]$ as follows. Let

$$t_{nk} = k\gamma_n \qquad (0 \leqslant k \leqslant m_n - 1), \qquad t_{nm_n} = T; \tag{3.7}$$

where

$$\gamma_n = \varepsilon_n^2 v_n [1/\varepsilon_n(\varepsilon_n v_n)^a], \qquad (0 < a < \tfrac{1}{3}), \tag{3.8}$$

and

$$m_n = [T/\gamma_n] + 1. \tag{3.9}$$

γ_n and $\Delta t_{nk} = t_{n(k+1)} - t_{nk}$ $(0 \leqslant k \leqslant m_n - 1)$ tend to 0 as n tends to ∞, because of [E].

Let

$$a(z) = \tilde{V}, \tag{3.10}$$

$$b(z) = \mu H_s(z). \tag{3.11}$$

We prove that the condition (3) of Theorem 2.11 is satisfied for $\{z^{(n)}(t), t \in [0, T]\}_{n=1}^{\infty}$. For $k = 0, 1, 2, \dots, m_n - 2$,

$$\Delta z_{nk}^{(n)} = z^{(n)}(t_{n(k+1)}) - z^{(n)}(t_{nk}) = \Delta_m Z_{[t_{nk}/\varepsilon_n^2 v_n]}^{(n)}, \tag{3.12}$$

where

$$m = [1/\varepsilon_n(\varepsilon_n v_n)^a], \tag{3.13}$$

$$\Delta_j Z_k^{(n)} = Z_{k+j}^{(n)} - Z_k^{(n)} \qquad (k, j \geqslant 0). \tag{3.14}$$

By (3.6) and (3.12), we have

$$\begin{aligned} \rho'_{nk} &= P(|\Delta z_{nk}^{(n)}| > \delta | \mathscr{F}_n(t_{nk}))/\Delta t_{nk} \\ &= P(|\Delta_m Z_{[t_{nk}/\varepsilon_n^2 v_n]}^{(n)}| > \delta | \mathscr{F}_{[t_{nk}/\varepsilon_n^2 v_n]-1}^{(n)})/m\varepsilon_n^2 v_n, \end{aligned} \tag{3.15}$$

$$\begin{aligned} \rho''_{nk} &= E\{\Delta z_{nk}^{(n)} \chi_\delta(\Delta z_{nk}^{(n)}) | \mathscr{F}_n(t_{nk})\}/\Delta t_{nk} - b(z^{(n)}(t_{nk})) \\ &= E\{\Delta_m Z_{[t_{nk}/\varepsilon_n^2 v_n]}^{(n)} \chi_\delta(\Delta_m Z_{[t_{nk}/\varepsilon_n^2 v_n]}^{(n)}) | \mathscr{F}_{[t_{nk}/\varepsilon_n^2 v_n]-1}^{(n)}\}/m\varepsilon_n^2 v_n \\ &\quad - \mu H_s(Z_{[t_{nk}/\varepsilon_n^2 v_n]}^{(n)}), \end{aligned} \tag{3.16}$$

$$\begin{aligned} \rho'''_{nk} &= E\{(\Delta z_{nk}^{(n)})^2 \chi_\delta(\Delta z_{nk}^{(n)}) | \mathscr{F}_n(t_{nk})\}/\Delta t_{nk} - a(z^{(n)}(t_{nk})) \\ &= E\{(\Delta_m Z_{[t_{nk}/\varepsilon_n^2 v_n]}^{(n)})^2 \cdot \chi_\delta(\Delta_m Z_{[t_{nk}/\varepsilon_n^2 v_n]}^{(n)}) | \mathscr{F}_{[t_{nk}/\varepsilon_n^2 v_n]-1}^{(n)}\}/m\varepsilon_n^2 v_n - \tilde{V}, \end{aligned} \tag{3.17}$$

for $k = 0, 1, 2, \dots, m_n - 2$, where δ is an arbitrarily fixed positive constant.

In the following, it is enough to assume that $H(z, s)$ is a linear combination of the product of uniformly bounded functions on $\mathbb{R}$ and $[-S, S]$, because we can obtain the essentially same estimations in Lemma A.4 and Lemma A.5 under the condition (iii) of [I] using Stone-Weierstrass theorem and the uniform tightness of the distribution of $Z_k^{(n)}$.

The right-hand sides of the inequalities in Lemma A.1, Lemma A.3 and Lemma A.5 tend to 0 as $n \to \infty$, since $m\varepsilon_n \to \infty$ and $m^3\varepsilon_n^4 v_n \to 0$ as $n \to \infty$ in this case. If we consider sufficiently large n, the condition $m\varepsilon_n^2 < \delta/M$ in these lemmas is satisfied for all m since $m\varepsilon_n^2 \to 0$ as $n \to \infty$ in this case. By (3.15), (3.16), (3.17) and these lemmas, we have

$$\sum_{k=0}^{m_n-2} E\{\rho'_{nk} + |\rho''_{nk}| + |\rho'''_{nk}|\} \Delta t_{nk} \to 0 \qquad \text{as} \qquad n \to \infty. \tag{3.18}$$

It is easy to see that $E\{\rho'_{n(m_n-1)}\}$, $E\{|\rho''_{n(m_n-1)}|\}$ and $E\{|\rho'''_{n(m_n-1)}|\}$ tend to 0 as $n \to \infty$, by the same way.

Therefore, condition (3) of Theorem 2.11 is satisfied. Since all the other conditions of this theorem are clearly satisfied, we get the conclusion.

Next, we consider weak convergence in $D[0, T]$. $D[0, T]$ is the space of functions on $[0, T]$ that are right-continuous and have left-hand limits, and with the Skorohod topology, $D[0, T]$ is a separable metric space (see Billingsley (1968)).

Let P_n be a probability measure on $D[0, T]$ induced by $z^{(n)}(t)$, $(t \in [0, T])$. The convergence of the finite-dimensional distribution of $z^{(n)}(t)$ to $z(t)$ leads weak convergence of $z^{(n)}(t)$ to $z(t)$ in $D[0, T]$, if $\{P_n\}_{n=1}^{\infty}$ is uniformly tight (see Theorem 15.1 of Billingsley (1968)). By Theorem 8.2, Theorem 8.3 and Theorem 15.5 of Billingsley (1968), a sufficient condition for uniform tightness of $\{P_n\}$ in $D[0, T]$ is that following [T-1] and [T-2] are satisfied.

[T-1] For each positive η, there exists a K such that

$$P(|z^{(n)}(0)| > K) \leqslant \eta \qquad (n \geqslant 1). \tag{3.19}$$

[T-2] For each positive ε and η, there exists a δ, with $0 < \delta < 1$, and an integer n_0 such that

$$P\left(\sup_{t \leqslant s \leqslant t+\delta} |z^{(n)}(s) - z^{(n)}(t)| \geqslant \varepsilon \right) \Big/ \delta \leqslant \eta \qquad (n \geqslant n_0), \tag{3.20}$$

for all $t \in [0, T - \delta]$.

We can really show that [T-1] and [T-2] are satisfied, as follows.

Theorem 3.2. *Assume all the conditions of Theorem* 3.1, *then* $\{z^{(n)}(t)\}_{n=1}^{\infty}$ *converges weakly in* $D[0, T]$, *to the diffusion process* $z(t)$.

Proof. [T-1] is obvious, as we assume (ii) of [II]. We show that [T-2] is satisfied.
By the definition of $z^{(n)}(t)$, we have

$$\begin{aligned} \sup_{t \leqslant s \leqslant t+\delta} |z^{(n)}(s) - z^{(n)}(t)| &= \sup_{t \leqslant s \leqslant t+\delta} |Z^{(n)}_{[s/\varepsilon_n^2 v_n]} - Z^{(n)}_{[t/\varepsilon_n^2 v_n]}| \\ &\leqslant \max(|\Delta_k Z^{(n)}_{[t/\varepsilon_n^2 v_n]}|;\ 1 \leqslant k \leqslant [\delta/\varepsilon_n^2 v_n] + 1). \end{aligned} \tag{3.21}$$

If we put $m = [\delta/\varepsilon_n^2 v_n] + 1$, then the condition $m < \varepsilon/M\varepsilon_n^2$ in Lemma A.6 is satisfied for all m and n $(n \geqslant n_1)$ for sufficiently small δ $(\delta > 0)$ and large n_1. We choose such a δ and consider n such that $n \geqslant n_1$. Using Lemma A.6, we have

$$\begin{aligned} &P\left(\sup_{t \leqslant s \leqslant t+\delta} |z^{(n)}(s) - z^{(n)}(t)| \geqslant \varepsilon \right) \\ &\quad \leqslant P(\max(|\Delta_k Z^{(n)}_{[t/\varepsilon_n^2 v_n]}|;\ 1 \leqslant k \leqslant [\delta/\varepsilon_n^2 v_n] + 1) \geqslant \varepsilon) \\ &\quad \leqslant \delta^2 K([\delta/\varepsilon_n^2 v_n]\varepsilon_n^2 v_n/\delta + \varepsilon_n^2 v_n/\delta)^2 / \{\varepsilon - ([\delta/\varepsilon_n^2 v_n]\varepsilon_n^2 + \varepsilon_n^2)M\}^4, \end{aligned} \tag{3.22}$$

where K and M are positive constants. Since ε_n and $\varepsilon_n^2 v_n$ tend to 0 as $n \to \infty$, we can choose a δ with $0 < \delta < 1$ and an integer n_0 $(n_0 \geqslant n_1)$ such that (3.20) is satisfied for all $t \in [0, T]$.

4. General Case

In this section, we consider a sequence of discrete time processes

$$\{X_k^{(n)}, k = 0, 1, 2, \ldots\}_{n=1}^{\infty},$$

defined by solutions to (1.4). We prove that a sequence of continuous time processes

$$x^{(n)}(t) = X^{(n)}_{[t/\varepsilon_n^2 v_n]} \qquad (t \in [0, T]) \tag{4.1}$$

converges weakly to a diffusion process $x(t)$. Considering (1.4), we introduce the following conditions:

[III] $f(x)$ is a function satisfying

(i) There exists an open interval $L = (x_1, x_2)$, $(-\infty \leqslant x_1 < x_2 \leqslant \infty)$ such that $f(x) > 0$ on L.

(ii) If $|x_i| < \infty$, then $f(x_i) = 0$ and there exists a positive constant B such that $f(x) < B|x - x_i|$ for any $x \in L$ in some neighbourhood of x_i. If $|x_i| = \infty$, then $\lim_{x \to x_i} (-1)^i \int_c^x (1/f(y))\, dy = \infty$ for fixed $c \in L$ $(i = 1, 2)$.

(iii) $f(x)$ is fifth continuously differentiable on L. $f(x)$, together with its derivatives up to fifth order inclusive, are continuous and uniformly bounded on L.

[IV] $g(x, s, \varepsilon)$ is a function on $L \times [-S, S] \times (0, 1)$ satisfying

(i)
$$g(x, s, \varepsilon) = h(x, s) + r(x, s, \varepsilon). \tag{4.2}$$

Here $r(x, s, \varepsilon)$ is a uniformly bounded function, and $\lim_{\varepsilon \to 0} \tilde{r}(\varepsilon) = 0$, where $\tilde{r}(\varepsilon) = \sup_{x,s} |r(x, s, \varepsilon)|$.

(ii) $h(x, s)$ is fourth continuously differentiable with respect to x, and together with its derivatives up to fourth order inclusive are continuous and uniformly bounded.

(iii) $h(x, s)$ is continuous and uniformly bounded on $L \times [-S, S]$.

[V] $X_0^{(n)} \in L$ $(n = 1, 2, \ldots)$ are random variables satisfying

(i) $X_0^{(n)}$ is independent of $\{s_k^{(n)}\}_{k=-\infty}^{\infty}$ $(n = 1, 2, \ldots)$.

(ii) $X_0^{(n)}$ converges weakly to a random variable $x_0 \in L$ as $n \to \infty$.

Applying Theorem 3.2, we prove the following theorem.

Theorem 4.1. *Assume* [S], [E], [III], [IV] *and* [V], *then* $\{x^{(n)}(t), (t \in [0, T])\}_{n=1}^{\infty}$ *converges weakly in* $D[0, T]$ *to a diffusion process* $x(t)$ *whose infinitesimal generator* $\mathcal{G}$ *is*

$$\mathcal{G} = \frac{\tilde{V}}{2} f^2(x) \frac{\partial^2}{\partial x^2} + f(x) \left\{ \frac{\tilde{V} - \mu v}{2} \frac{df}{dx}(x) + \mu h_s(x) \right\} \frac{\partial}{\partial x}, \tag{4.3}$$

and the distribution of $x(0)$ *is that of* x_0. *Here we put*

$$h_s(x) = E\{h(x, s_0^{(n)}(\omega))\} \qquad (x \in L) \tag{4.4}$$

and

$$v = E\{(s_0^{(n)})^2\}. \tag{4.5}$$

Remark. (i) The solution $X_k^{(n)}$ of (1.4) always stays in the interval L for sufficiently small ε_n, because of [III] and [IV].

(ii) Weak convergence in $D[0, T]$ implies the convergence of the finite-dimensional distribution in this case, since $x(t)$ is continuous with probability one (see Billingsley (1968), p. 124).

(iii) Recently, Kushner and Huang (1981), using a method of the martingale problem of Stroock and Varadhan (1979), obtained general results for weak convergence of a sequence of stochastic difference equations defined on $\mathbb{R}^d$ to a diffusion process, where the mixing rate $\phi_n(k)$ is independent of n. In this case, for $L = \mathbb{R}$, Theorem 4.1 reduces to a special case of their results.

Proof. Define a sequence of discrete time processes $\{Z_k^{(n)}, k = 0, 1, 2, \ldots\}_{n=1}^{\infty}$ by a transformation of the sequence of processes $\{X_k^{(n)}, k = 0, 1, 2, \ldots\}_{n=1}^{\infty}$ as follows.

$$Z_k^{(n)} = F(X_k^{(n)}) \qquad (k \geqslant 0, n \geqslant 1), \tag{4.6}$$

where

$$F(x) = \int_c^x \frac{1}{f(y)} dy \qquad (x \in L), \tag{4.7}$$

for fixed $c \in L$. By [III], $\{Z_k^{(n)}\}$ is a sequence of real-valued processes, $F(x)$ is sixth differentiable on L, and $F(x)$ is invertible. Let $\tilde{F} = \tilde{F}(z)$ be the inverse function of $F(x)$.

First, we prove weak convergence of $z^{(n)}(t) = Z_{[t/\varepsilon_n^2 v_n]}^{(n)}$ to some diffusion process applying Theorem 3.2. The increment of $Z_k^{(n)}$ is

$$\begin{aligned} \Delta Z_k^{(n)} &= Z_{k+1}^{(n)} - Z_k^{(n)} = F(X_{k+1}^{(n)}) - F(X_k^{(n)}) \\ &= \frac{dF}{dx}(X_k^{(n)})\,\Delta X_k^{(n)} + \frac{1}{2}\frac{d^2F}{dx^2}(X_k^{(n)})(\Delta X_k^{(n)})^2 + Q(X_k^{(n)}, s_k^{(n)}, \varepsilon_n), \end{aligned} \tag{4.8}$$

where

$$\Delta X_k^{(n)} = X_{k+1}^{(n)} - X_k^{(n)}, \tag{4.9}$$

and

$$Q(X_k^{(n)}, s_k^{(n)}, \varepsilon_n) = F(X_{k+1}^{(n)}) - F(X_k^{(n)}) - \frac{dF}{dx}(X_k^{(n)})\,\Delta X_k^{(n)} - \frac{1}{2}\frac{d^2F}{dx^2}(X_k^{(n)})(\Delta X_k^{(n)})^2. \tag{4.10}$$

Q is a function of $(X_k^{(n)}, s_k^{(n)}, \varepsilon_n)$, since $X_{k+1}^{(n)}$ and $\Delta X_k^{(n)}$ are functions of $(X_k^{(n)}, s_k^{(n)}, \varepsilon_n)$. We show in Appendix that

$$\sup_{x,s,\varepsilon} |Q(x, s, \varepsilon)|/\varepsilon^3 < \infty \tag{4.11}$$

for sufficiently small ε (see Lemma A.7). Substituting (1.4), (4.2) and (4.7) into (4.8), we have

$$\Delta Z_k^{(n)} = \varepsilon_n s_k^{(n)} + \varepsilon_n^2 \left\{ h(X_k^{(n)}, s_k^{(n)}) - \tfrac{1}{2}(s_k^{(n)})^2 \frac{df}{dx}(X_k^{(n)}) \right\} + \varepsilon_n^2 q(X_k^{(n)}, s_k^{(n)}, \varepsilon_n), \tag{4.12}$$

where

$$q(x, s, \varepsilon) = r(x, s, \varepsilon) + Q(x, s, \varepsilon)/\varepsilon^2. \tag{4.13}$$

By condition (i) of [IV] and (4.11), $q(x, s, \varepsilon)$ is uniformly bounded and

$$\sup_{x,s} |q(x, s, \varepsilon)| \to 0 \qquad (\varepsilon \to 0). \tag{4.14}$$

Therefore, we have

$$\Delta Z_k^{(n)} = \varepsilon_n s_k^{(n)} + \varepsilon_n^2 H(Z_k^{(n)}, s_k^{(n)}) + \varepsilon_n^2 R(Z_k^{(n)}, s_k^{(n)}, \varepsilon_n), \tag{4.15}$$

where

$$H(z, s) = h(\tilde{F}(z), s) - \frac{s^2}{2} \frac{df}{dx}(\tilde{F}(z)), \tag{4.16}$$

and

$$R(z, s, \varepsilon) = q(\tilde{F}(z), s, \varepsilon). \tag{4.17}$$

By [III], [IV] and (4.14), we have

$$\sup_{z,s} |H(z, s)| < \infty, \tag{4.18}$$

$$\sup_{z,s,\varepsilon} |R(z, s, \varepsilon)| < \infty, \tag{4.19}$$

$$\sup_{z,s} |R(z, s, \varepsilon)| \to 0 \qquad (\varepsilon \to 0). \tag{4.20}$$

Using [III], [IV], [V], (4.18), (4.19) and (4.20), we can see that all the conditions of Theorem 3.2 are satisfied for the sequence of processes $\{Z_k^{(n)}\}$ defined by the transformation of $\{X_k^{(n)}\}$ by F. Therefore, $\{z^{(n)}(t)\}$ converges weakly in $D[0, T]$ to a diffusion process $z(t)$. By Corollary 1 to Theorem 5.1 of Billingsley (1968), we have weak convergence of $\{x^{(n)}(t)\}$ in $D[0, T]$, since $\tilde{F}(z)$ is continuous.

Finally, we decide the infinitesimal generator of the limit process $x(t)$. We can represent $z(t)$ by the following stochastic differential equation of Ito type,

$$dz(t) = \tilde{V}^{1/2}\, dB_t + \mu H_s(z(t))\, dt, \tag{4.21}$$

where $H_s(z) = E\{H(z, s_0^{(n)})\}$, $(z \in \mathbb{R})$, and B_t is a one-dimensional Brownian motion. Using Ito's formula [Ito (1961), p. 187], we have a stochastic differential equation for $x(t) = \tilde{F}(z(t))$ as follows.

$$\begin{aligned} dx(t) &= \frac{d\tilde{F}}{dz} dz(t) + \frac{1}{2} \frac{d^2\tilde{F}}{dz^2} \tilde{V}\, dt \\ &= f(x(t))\{\tilde{V}^{1/2}\, dB_t + \mu H_s(z(t))\, dt\} + \frac{\tilde{V}}{2} f(x(t)) \frac{df}{dx}(x(t))\, dt \\ &= \tilde{V}^{1/2} f(x(t))\, dB_t + f(x(t)) \left\{ \frac{\tilde{V} - \mu v}{2} \frac{df}{dx}(x(t)) + \mu h_s(x(t)) \right\} dt. \end{aligned} \tag{4.22}$$

Therefore,

$$\mathscr{G} = \frac{\tilde{V}}{2} f^2(x) \frac{\partial^2}{\partial x^2} + f(x) \left\{ \frac{\tilde{V} - \mu v}{2} \frac{df}{dx}(x) + \mu h_s(x) \right\} \frac{\partial}{\partial x} \tag{4.23}$$

is the infinitesimal generator of the limiting diffusion process $x(t)$.

5. Stochastic Selection Models in Population Genetics

Let us interpret and apply the result of the preceding section to stochastic selection models in population genetics.

First, consider the haploid model represented by (1.2). In order to include the case $E\{s_k\} \neq 0$, we write the random variable s_k of the stationary stochastic process in (1.2) as

$$s_k = \varepsilon\sigma_k + \varepsilon^2\tilde{\sigma}, \tag{5.1}$$

where

$$E\{\sigma_k\} = 0, \tag{5.2}$$

ε is a positive constant much smaller than 1, and $\tilde{\sigma}$ is a constant of the order of 1. Random variables $\{\sigma_k\}$ are uniformly bounded, and $\{\sigma_k\}_{k=-\infty}^{\infty}$ is a strictly stationary ϕ-mixing process with $v = \sum_{k=0}^{\infty} \phi^{1/2}(k) < \infty$. The parameter v is a kind of relaxation time.

Then, (1.2) can be rewritten as

$$\begin{aligned} \Delta X_k = X_{k+1} - X_k &= s_k X_k(1 - X_k)\left\{1 - s_k X_k + \frac{(s_k X_k)^2}{1 + s_k X_k}\right\} \\ &= X_k(1 - X_k)\left\{\varepsilon\sigma_k - (\varepsilon\sigma_k)^2 X_k + \varepsilon^2\tilde{\sigma}\right. \\ &\quad \left. - \varepsilon^3\left[2\tilde{\sigma}\sigma_k + \tilde{\sigma}^2 - \frac{(\sigma_k + \varepsilon\tilde{\sigma})^3 X_k}{1 + \varepsilon(\sigma_k + \varepsilon\tilde{\sigma})X_k}\right]\right\}. \end{aligned} \tag{5.3}$$

Taking σ_k for $s_k^{(n)}$ in (1.4), we get comparing (5.3) with (1.4):

$$f(x) = x(1 - x), \tag{5.4}$$

$$g(x, \sigma, \varepsilon) = \tilde{\sigma} - \sigma^2 x + \varepsilon\left\{-(2\tilde{\sigma}\sigma + \tilde{\sigma}^2) + \frac{(\sigma + \varepsilon\tilde{\sigma})^3 x}{1 + \varepsilon(\sigma + \varepsilon\tilde{\sigma})x}\right\}. \tag{5.5}$$

These functions are easily verified to satisfy conditions [III] and [IV] of Theorem 4.1 by setting $x_1 = 0$, and $x_2 = 1$. Therefore, if εv is sufficiently smaller than 1, the diffusion approximation is justified by Theorem 4.1 noting the condition [E].

In (4.1) the continuous time t is associated with the $[t/\varepsilon_n^2 v_n]$th generation of the original discrete model. However, in the diffusion approximation of population genetics it is natural and convenient to take a generation as a unit of time. In order to obtain an appropriate infinitesimal generator of the approximative diffusion process in this time unit, we only have to multiply $\mathscr{G}$ in (4.3) by $\varepsilon_n^2 v_n$. Thus, from (4.3), (5.4) and (5.5) we obtain the diffusion and drift term of the generator corresponding to the case $\mu \neq 0$ as

$$\mathscr{G} = \frac{a(x)}{2}\frac{\partial^2}{\partial x^2} + b(x)\frac{\partial}{\partial x}, \tag{5.6}$$

$$a(x) = V\{x(1 - x)\}^2, \tag{5.7}$$

and

$$b(x) = x(1 - x)\left\{\bar{s} - vx + \frac{V - v}{2}(1 - 2x)\right\}, \tag{5.8}$$

where

$$\bar{s} = \varepsilon^2 \tilde{\sigma} = E\{s_0\}, \tag{5.9}$$

$$v = \varepsilon^2 E\{\sigma_0^2\} = \mathrm{Var}\{s_0^2\}, \tag{5.10}$$

$$V = \varepsilon^2 \sum_{k=-\infty}^{\infty} E\{\sigma_0 \sigma_k\} = \sum_{k=-\infty}^{\infty} \mathrm{Cov}\{s_0, s_k\}. \tag{5.11}$$

The case $\mu = 0$ can be regarded as an extreme case of $\mu \neq 0$ such that $\max(|\bar{s}|, v) \ll V$ as a result of a significant autocorrelation. Therefore, in the following we apply our theorem for $\mu \neq 0$.

Remark. The cases $\mu \neq 0$ and $\mu = 0$ correspond to the "mildly autocorrelated" and the "moderately autocorrelated" cases of Gillespie and Guess (1978).

Next, let us consider the diploid model. Let $W_{AA}(k)$, $W_{Aa}(k)$ and $W_{aa}(k)$ be the fitness of genotypes AA, Aa, and aa in the kth generation. We give these fitnesses in the form

$$\begin{aligned} W_{AA}(k) &= 1 + \alpha s_k, \\ W_{Aa}(k) &= 1 + \beta s_k, \\ W_{aa}(k) &= 1 + \gamma s_k, \end{aligned} \tag{5.12}$$

where α, β, and γ are constant parameters. Assuming random mating, we obtain the gene frequency change of A-allele as

$$\begin{aligned} \Delta X_k &= \frac{W_{AA} X_k^2 + W_{Aa} X_k (1 - X_k)}{W_{AA} X_k^2 + 2 W_{Aa} X_k (1 - X_k) + W_{aa} (1 - X_k)^2} - X_k \\ &= \frac{\alpha - \gamma}{2} \frac{s_k X_k (1 - X_k)}{1 + s_k J(X_k)} \{1 + \theta(2X_k - 1)\}, \end{aligned} \tag{5.13}$$

where

$$J(x) = \alpha x^2 + 2\beta x(1 - x) + \gamma(1 - x)^2, \tag{5.14}$$

and

$$\theta = \frac{\alpha + \gamma - 2\beta}{\alpha - \gamma}. \tag{5.15}$$

Here, we restrict ourselves to the case

$$\alpha \neq \gamma \qquad \text{and} \qquad |\theta| \leqslant 1, \tag{5.16}$$

in order that Theorem 4.1 can be used.

When the selection is genic, we have $\theta = 0$ by definition, so that we may call the constant parameter θ "non-genicity index". It is related to the degree of dominance h of A-allele as

$$\theta = 1 - 2h, \tag{5.17}$$

where h is defined by

$$h = \frac{W_{Aa}(k) - W_{aa}(k)}{W_{AA}(k) - W_{aa}(k)}. \tag{5.18}$$

In the same way as we have obtained (5.4) and (5.5) from (5.3), we get from (5.13):

$$f(x) = \frac{\alpha - \gamma}{2} x(1-x)\{1 - \theta(1-2x)\}, \tag{5.19}$$

$$h_s(x) = \frac{\alpha - \gamma}{2} x(1-x)\{1 - \theta(1-2x)\}\{\tilde{\sigma} - E[\sigma_0^2]J(x)\}. \tag{5.20}$$

Therefore, by Theorem 4.1 the appropriate infinitesimal generator for the above diploid model has the diffusion and drift term given by

$$a(x) = V\left(\frac{\alpha - \gamma}{2}\right)^2 \{x(1-x)[1 - \theta(1-2x)]\}^2, \tag{5.21}$$

$$\begin{aligned} b(x) = \frac{\alpha - \gamma}{2} x(1-x)\{1 - \theta(1-2x)\}\Big\{\bar{s} - vJ(x) \\ + \frac{\alpha - \gamma}{2}\frac{V - v}{2}[(1-2x) - \theta(1 - 6x + 6x^2)]\Big\}, \end{aligned} \tag{5.22}$$

where $\bar{s}$, v and V are given by (5.9), (5.10) and (5.11).

Noting in (5.12) that without loss of generality we can put

$$\frac{\alpha - \gamma}{2} = 1, \tag{5.23}$$

we can write instead of (5.18) and (5.19) as

$$a(x) = V\{x(1-x)[1 - \theta(1-2x)]\}^2, \tag{5.24}$$

$$\begin{aligned} b(x) = x(1-x)\{1 - \theta(1-2x)\}\Big\{\bar{s} - vJ(x) \\ + \frac{V - v}{2}[(1-2x) - \theta(1 - 6x + 6x^2)]\Big\}. \end{aligned} \tag{5.25}$$

When the stochastic variables have no autocorrelation, we have by (5.10) and (5.11) $V = v$. For $V = v$, the drift term may give a delicate effect dependent on the detail of the selection scheme. Indeed, even for the case of genic selection, $\theta = 0$, where we have

$$b(x) = x(1-x)\{\bar{s} + v(1 - 2x - \beta)\}, \tag{5.26}$$

this drift term acts either centrifugally or centripetally to gene frequency x according to the value of β.

On the other hand, when the autocorrelation is significant to the extent $\max(|\bar{s}|, v) \ll V$, and the non-genicity index θ is close to zero, then the diffusion and drift term $a(x)$ and $b(x)$ are invariably approximated by

$$a_0(x) = V\{x(1-x)\}^2, \tag{5.27}$$

$$b_0(x) = \frac{V}{2}x(1-x)(1-2x). \tag{5.28}$$

Using the diffusion model for stochastic selection essentially based on the diffusion and drift term given by (5.27) and (5.28), Matsuda and Gojobori (1979) analyzed the data of protein polymorphism.

When the effect of mutation is included, $g(x, s, \varepsilon)$ in (1.4) cannot satisfy condition [IV], since $f(x) \to 0$ as $x \to 0$ or $x \to 1$, while $f(x)g(x, s, \varepsilon)$ does not vanish there in the presence of mutation. In this sense, there still remains a logical gap in the mathematical foundation of the diffusion approximation of stochastic selection models. Apart such drawbacks our result shows that the diffusion approximation is indeed valid even in the presence of autocorrelation so long as we have $\varepsilon v \ll 1$, that is so long as the gene frequency change during the time span of the relaxation time of the mixing process is very small compared to 1. It also gives the correct infinitesimal generator of the approximating diffusion process to be used.

Acknowledgements. We are grateful to Professor H. Watanabe, Professor H. Kunita, and Dr. T. Yamada for their advice from mathematical points of view. We are indebted to Dr. K. Ishii for his interest and comments.

Appendix

We prove several lemmas that are necessary for the estimate in §3 and §4. Without loss of generality, we can assume that all the bounded functions appearing in these sections have common bound M ($M < \infty$). All the quantities that are not defined in this section are defined in these sections. First, we estimate the cummulative sum $\Delta_m Z_k^{(n)} = Z_{k+m}^{(n)} - Z_k^{(n)}$.

Lemma A.1. *For each positive* δ,

$$P(|\Delta_m Z_k^{(n)}| > \delta)/m\varepsilon_n^2 v_n \leqslant \{cS/(\delta - Mm\varepsilon_n^2)\}^4 m\varepsilon_n^2 v_n \qquad (1 \leqslant m < \delta/M\varepsilon_n^2,\ k \geqslant 0) \tag{A.1}$$

where c is a positive constant.

Proof. For each positive integer m ($m < \delta/M\varepsilon_n^2$), we have

$$P(|\Delta_m Z_k^{(n)}| > \delta) \leqslant P\left(\varepsilon_n \left|\sum_{j=0}^{m-1} s_{k+j}^{(n)}\right| + Mm\varepsilon_n^2 > \delta\right) = P\left(\left|\sum_{j=0}^{m-1} s_{k+j}^{(n)}\right| > \delta/\varepsilon_n - Mm\varepsilon_n\right)$$
$$\leqslant \{cS/(\delta - Mm\varepsilon_n^2)\}^4 m^2 \varepsilon_n^4 v_n. \tag{A.2}$$

Here, we used Lemma 2.8 for the estimate of the last inequality.

Lemma A.2. *For each integer* k ($k \geqslant 0$),

$$E\{|E[(\Delta_m Z_k^{(n)})^2 | \mathscr{F}_{k-1}^{(n)}] - m\varepsilon_n^2 V_n|\}$$
$$\leqslant 20S^2\varepsilon_n^2(v_n - 1)^2 + 2cSMm^{3/2}\varepsilon_n^3 v_n^{1/2} + M^2 m^2 \varepsilon_n^4 \qquad (m \geqslant 0). \tag{A.3}$$

Proof. For each integer m ($m \geqslant 1$) and k ($k \geqslant 0$), we have

$$|E\{(\Delta_m Z_k^{(n)})^2 | \mathscr{F}_{k-1}^{(n)}\} - m\varepsilon_n^2 V_n|$$
$$\leqslant \varepsilon_n^2 \left| E\left\{\left(\sum_{j=0}^{m-1} s_{k+j}^{(n)}\right)^2 \middle| \mathscr{F}_{k-1}^{(n)}\right\} - mV_n \right| + 2Mm\varepsilon_n^3 E\left\{\left|\sum_{j=0}^{m-1} s_{k+j}^{(n)}\right| \middle| \mathscr{F}_{k-1}^{(n)}\right\} + M^2 m^2 \varepsilon_n^4$$

$$\leqslant 20S^2\varepsilon_n^2(v_n-1)^2 + 2Mm\varepsilon_n^3\left(E\left\{\left(\sum_{j=0}^{m-1} s_{k+j}^{(n)}\right)^4\Big|\mathscr{F}_{k-1}^{(n)}\right\}\right)^{1/4} + M^2m^2\varepsilon_n^4. \tag{A.4}$$

We applied Lemma 2.7 for the last inequality. Here, by Lemma 20.4 of Billingsley (1968),

$$E\left\{\left(E\left[\left(\sum_{j=0}^{m-1} s_{k+j}^{(n)}\right)^4\Big|\mathscr{F}_{k-1}^{(n)}\right]\right)^{1/4}\right\} \leqslant \left(E\left\{E\left[\left(\sum_{j=0}^{m-1} s_{k+j}^{(n)}\right)^4\Big|\mathscr{F}_{k-1}^{(n)}\right]\right\}\right)^{1/4}$$

$$= \left(E\left\{\left(\sum_{j=0}^{m-1} s_{k+j}^{(n)}\right)^4\right\}\right)^{1/4} \leqslant cS(mv_n)^{1/2}. \tag{A.5}$$

We have the conclusion from (A.4) and (A.5).

Lemma A.3. *For each positive δ,*

$$E\{|E[\chi_\delta(\Delta_m Z_k^{(n)})(\Delta_m Z_k^{(n)})^2|\mathscr{F}_{k-1}^{(n)}] - m\varepsilon_n^2 V_n|\}/m\varepsilon_n^2 v_n$$
$$\leqslant 20S^2(v_n-1)^2/mv_n + 2cSM(m\varepsilon_n^2/v_n)^{1/2} + M^2m\varepsilon_n^2/v_n$$
$$+ (S+M\varepsilon_n)^2\{cS/(\delta - Mm\varepsilon_n^2)\}^4 m^3\varepsilon_n^4 v_n \qquad (1 \leqslant m < \delta/M\varepsilon_n^2,\ k \geqslant 0). \tag{A.6}$$

Proof. For each integer m ($m \geqslant 1$), we have

$$|E\{\chi_\delta(\Delta_m Z_k^{(n)})(\Delta_m Z_k^{(n)})^2|\mathscr{F}_{k-1}^{(n)}\} - m\varepsilon_n^2 V_n|$$
$$\leqslant |E\{(\Delta_m Z_k^{(n)})^2|\mathscr{F}_{k-1}^{(n)}\} - m\varepsilon_n^2 V_n| + (Sm\varepsilon_n + Mm\varepsilon_n^2)^2 E\{1 - \chi_\delta(\Delta_m Z_k^{(n)})|\mathscr{F}_{k-1}^{(n)}\}. \tag{A.7}$$

By Lemma A.1,

$$E\{E[1 - \chi_\delta(\Delta_m Z_k^{(n)})|\mathscr{F}_{k-1}^{(n)}]\} = P(|\Delta_m Z_k^{(n)}| > \delta) \leqslant \{cS/(\delta - Mm\varepsilon_n^2)\} m^2\varepsilon_n^4 v_n^2. \tag{A.8}$$

We have the conclusion from Lemma A.2 and (A.8).

Lemma A.4. *For each integer k ($k \geqslant 0$),*

$$E\{|E[\Delta_m Z_k^{(n)}|\mathscr{F}_{k-1}^{(n)}] - m\varepsilon_n^2 H_s(Z_k^{(n)})|\}$$
$$\leqslant 2S\varepsilon_n(v_n-1) + 2LM^2\varepsilon_n^2(v_n-1) + cSMm^{3/2}\varepsilon_n^3 v_n^{1/2} + M^2m^2\varepsilon_n^4 + \tilde{R}(\varepsilon_n)m\varepsilon_n^2 \qquad (m \geqslant 0), \tag{A.9}$$

where L is the number of terms of linear combination in $H(z,s)$.

Proof. For simplicity, we consider the case of $L = 1$. We can write

$$H(z,s) = f(z)g(s), \tag{A.10}$$

where $f(z)$ and $g(s)$ are uniformly bounded functions on $\mathbb{R}$ and $[-S, S]$ respectively.

For each integer m ($m \geqslant 1$) and k ($k \geqslant 0$), let

$$J_{m,k}^{(n)} = |E\{\Delta_m Z_k^{(n)}|\mathscr{F}_{k-1}^{(n)}\} - m\varepsilon_n^2 H_s(Z_k^{(n)})|. \tag{A.11}$$

We expand $H(Z_{k+j}^{(n)}, s_{k+j}^{(n)})$ as follows.

$$H(Z_{k+j}^{(n)}, s_{k+j}^{(n)}) = H(Z_k^{(n)}, s_{k+j}^{(n)}) + \frac{\partial H}{\partial z}(Z_k^{(n)} + \theta(\omega)\Delta_j Z_k^{(n)}, s_{k+j}^{(n)})\,\Delta_j Z_k^{(n)}, \tag{A.12}$$

where $0 \leqslant \theta(\omega) \leqslant 1$. Substituting (A.12) into (A.11), we have

$$J_{m,k}^{(n)} \leqslant \sum_{j=0}^{m-1}\{\varepsilon_n|E[s_{k+j}^{(n)}|\mathscr{F}_{k-1}^{(n)}]| + \varepsilon_n^2|E[H(Z_{k+j}^{(n)}, s_{k+j}^{(n)})|\mathscr{F}_{k-1}^{(n)}] - H_s(Z_k^{(n)})|\} + \tilde{R}(\varepsilon_n)m\varepsilon_n^2$$
$$\leqslant 2S\varepsilon_n\sum_{j=1}^{m}\phi_n(j) + \varepsilon_n^2\sum_{j=1}^{m-1}\left\{|E[H(Z_k^{(n)}, s_{k+j}^{(n)})|\mathscr{F}_{k-1}^{(n)}] - H_s(Z_k^{(n)})|\right.$$
$$\left. + \left|E\left[\frac{\partial H}{\partial z}(Z_k^{(n)} + \theta(\omega)\,\Delta_j Z_k^{(n)}, s_{k+j}^{(n)})\,\Delta_j Z_k^{(n)}\,\Big|\,\mathscr{F}_{k-1}^{(n)}\right]\right|\right\} + \tilde{R}(\varepsilon_n)m\varepsilon_n^2$$

$$\leqslant 2S\varepsilon_n(v_n - 1) + \varepsilon_n^2 \sum_{j=0}^{m-1} \{|E[H(Z_k^{(n)}, s_{k+j}^{(n)})|\mathscr{F}_{k-1}^{(n)}] - H_s(Z_k^{(n)})| + ME[|\Delta_j Z_k^{(n)}|\,|\mathscr{F}_{k-1}^{(n)}]\} + \tilde{R}(\varepsilon_n)m\varepsilon_n^2. \tag{A.13}$$

Since $Z_k^{(n)}$ is $\mathscr{F}_{k-1}^{(n)}$-measurable, we have

$$|E\{H(Z_k^{(n)}, s_{k+j}^{(n)})|\mathscr{F}_{k-1}^{(n)}\} - H_s(Z_k^{(n)})| \leqslant f(Z_k^{(n)})E\{g(s_{k+j}^{(n)})|\mathscr{F}_{k-1}^{(n)}\} - f(Z_k^{(n)})E\{g(s_0^{(n)})\}| \leqslant M|E\{g(s_{k+j}^{(n)})|\mathscr{F}_{k-1}^{(n)}\} - E\{g(s_{k+j}^{(n)})\}| \leqslant 2M^2\phi_n(j+1). \tag{A.14}$$

We used Lemma 2.2 for the estimate of the last inequality. On the other hand, by Lemma 20.4 of Billingsley (1968), we have

$$E\{|\Delta_j Z_k^{(n)}|\} \leqslant \varepsilon_n E\left\{\left|\sum_{i=0}^{j-1} s_{k+i}^{(n)}\right|\right\} + Mj\varepsilon_n^2 \leqslant \varepsilon_n\left(E\left\{\left(\sum_{i=0}^{j-1} s_{k+i}^{(n)}\right)^4\right\}\right)^{1/4} + Mj\varepsilon_n^2 \leqslant \varepsilon_n cS(jv_n)^{1/2} + Mj\varepsilon_n^2. \tag{A.15}$$

By (A.13), (A.14) and (A.15), we have

$$E\{J_{m,k}^{(n)}\} \leqslant 2S\varepsilon_n(v_n - 1) + 2M^2\varepsilon_n^2(v_n - 1) + cSMm^{3/2}\varepsilon_n^3 v_n^{1/2} + M^2m^2\varepsilon_n^4 + \tilde{R}(\varepsilon_n)m\varepsilon_n^2. \tag{A.16}$$

For the case of $L \geqslant 2$, the proof is analogous to the case of $L = 1$.

The following lemma is easily proved as the same way to the proof of Lemma A.3.

Lemma A.5. *For each positive* δ,

$$E\{|E[\chi_\delta(\Delta_m Z_k^{(n)})\,\Delta_m Z_k^{(n)}|\mathscr{F}_{k-1}^{(n)}] - m\varepsilon_n^2 H_s(Z_k^{(n)})\}/m\varepsilon_n^2 v_n \leqslant 2S(v_n - 1)/m\varepsilon_n v_n + 2LM^2(v_n - 1)/mv_n + cSM(m\varepsilon_n^2/v_n)^{1/2} + M^2m\varepsilon_n^2/v_n + \tilde{R}(\varepsilon_n)/v_n + (S + M\varepsilon_n)\{cS/(\delta - Mm\varepsilon_n^2)\}^4 m^2\varepsilon_n^4 v_n \quad (1 \leqslant m < \delta/M\varepsilon_n^2,\ k \geqslant 0), \tag{A.17}$$

where L is given in Lemma A.4.

The next lemma is necessary for proving the tightness condition.

Lemma A.6. *For each positive* ε,

$$P\left(\max_{1\leqslant j\leqslant m}(|\Delta_j Z_k^{(n)}|) \geqslant \varepsilon\right) \leqslant Km^2\varepsilon_n^4 v_n^2/(\varepsilon - Mm\varepsilon_n^2)^4 \qquad (1 \leqslant m < \varepsilon/M\varepsilon_n^2,\ k \geqslant 0), \tag{A.18}$$

where K is a constant.

Proof. We apply Lemma 2.10. For each positive integer m ($m < \varepsilon/M\varepsilon_n^2$), we have

$$P\left(\max_{1\leqslant j\leqslant m}(|\Delta_j Z_k^{(n)}|) \geqslant \varepsilon\right) \leqslant P\left(\varepsilon_n \max_{1\leqslant j\leqslant m}\left(\left|\sum_{i=1}^{j} s_{k+i}^{(n)}\right|\right) + Mm\varepsilon_n^2 \geqslant \varepsilon\right) = P\left(\max_{1\leqslant j\leqslant m}\left(\left|\sum_{i=1}^{j} s_{k+i}^{(n)}\right|\right) \geqslant (\varepsilon - Mm\varepsilon_n^2)/\varepsilon_n\right) \leqslant Km^2\varepsilon_n^4 v_n^2/(\varepsilon - Mm\varepsilon_n^2)^4. \tag{A.19}$$

Finally, we prove the following lemma that is necessary for the estimate in §4.

Lemma A.7. *For* $Q(x, s, \varepsilon)$ *given by* (4.10),

$$\sup_{x,s,\varepsilon}|Q(x, s, \varepsilon)|/\varepsilon^3 < \infty. \tag{A.20}$$

Proof. We can write $Q(x, s, \varepsilon)$ as follows.

$$Q(x, s, \varepsilon) = \frac{d^3F}{dx^3}(x + \theta\,\Delta x)(\Delta x)^3/6, \tag{A.21}$$

where θ $(0 \leqslant \theta \leqslant 1)$ is a constant. Substituting

$$(\Delta x)^3 = f^3(x)\{\varepsilon s + \varepsilon^2 g(x,s,\varepsilon)\}^3, \tag{A.22}$$

$$\frac{d^3F}{dx^3}(x) = \left\{2\left(\frac{df}{dx}(x)\right)^2 - f(x)\frac{d^2f}{dx^2}(x)\right\}\Big/ f^3(x), \tag{A.23}$$

into (A.21), we have

$$Q(x,s,\varepsilon) = \varepsilon^3\{s + \varepsilon g(x,s,\varepsilon)\}^3 \left\{2\left(\frac{df}{dx}(y)\right)^2 - f(y)\frac{d^2f}{dx^2}(y)\right\}\{f(x)/f(y)\}^3/6, \tag{A.24}$$

where we put $y = x + \theta\Delta x$. On the other hand, we can write

$$f(x+\theta\Delta x) = f(x) + \frac{df}{dx}(z)\theta\Delta x = f(x)\left\{1 + \theta[\varepsilon s + \varepsilon^2 g(x,s,\varepsilon)]\frac{df}{dx}(z)\right\}, \tag{A.25}$$

where we put $z = x + \theta_1\theta\Delta x$, θ_1 $(0 \leqslant \theta_1 \leqslant 1)$ is a constant. (A.24) and (A.25) give the following estimate for sufficiently small ε.

$$|Q(x,s,\varepsilon)| \leqslant \varepsilon^3 M^2 (s + M\varepsilon)^3, \tag{A.26}$$

since the right-hand side of (A.25) is larger than $f(x)/2$ for sufficiently small ε. (A.26) gives the conclusion.

References

1. Billingsley, P.: Convergence of probability measures. New York: Wiley 1968
2. Cook, R. D., Hartl, D. L.: Uncorrelated random environments and their effects on gene frequency. Evolution **28**, 265–274 (1974)
3. Crow, J. F., Kimura, M.: An introduction to population genetics theory. New York: Harper and Row 1970
4. Ewens, W. J.: Mathematical population genetics. New York: Springer 1979
5. Feller, W.: Diffusion processes in genetics. Second Berkeley Symp. Math. Stat. Prob., pp. 227–246. Berkeley and Los Angeles: University of California Press 1951
6. Gikhman, I. I.: Convergence to Markov processes. Ukrainian Math. Jour. **21**, 263–270 (1969)
7. Gillespie, J. H.: Natural selection with varying selection coefficients – A haploid model. Genet. Res., Camb. **21**, 115–120 (1973a)
8. Gillespie, J. H.: Polymorphism in random environments. Theoret. Population Biology **4**, 193–195 (1973b)
9. Gillespie, J. H.: A general model to account for enzyme variation in natural populations, V. The SAS-CFF model. Theoret. Population Biology **14**, 1–45 (1978)
10. Gillespie, J. H., Guess, H. A.: The effects of environmental autocorrelations on the progress of selection in a random environment. Amer. Natur. **112**, 897–909 (1978)
11. Guess, H. A.: On the weak convergence of Wright-Fisher models. Stoch. Processes Appl. **1**, 287–306 (1973)
12. Guess, H. A., Gillespie, J. H.: Diffusion approximations to linear stochastic difference equations with stationary coefficients. J. Appl. Prob. **14**, 58–74 (1977)
13. Hartl, D. L.: Mutation – selection balance with stochastic selection. Genetics **86**, 687–696 (1977)
14. Hartl, D. L., Cook, R. D.: Balanced polymorphisms of quasineutral alleles. Theoret. Population Biology **4**, 163–172 (1973)
15. Hartl, D. L., Cook, R. D.: Autocorrelated random environments and their effects on gene frequency. Evolution **28**, 275–280 (1974)
16. Ito, K.: Lectures on stochastic processes. Bombay: Tata Institute of Fundamental Research 1961
17. Jensen, L., Pollak, E.: Random selective advantages of a gene in a finite population. J. Appl. Prob. **6**, 19–37 (1969)
18. Karlin, S., Levikson, B.: Temporal fluctuations in selection intensities: Case of small population size. Theoret. Population Biology **6**, 383–412 (1974)

19. Karlin, S., Lieberman, U.: Random temporal variation in selection intensities: Case of large population size. Theoret. Population Biology **6**, 355–382 (1974)
20. Kimura, M.: Stochastic processes and distribution of gene frequencies under natural selection. Cold Spring Harbor Symp. **20**, 33–53 (1955)
21. Kimura, M.: Diffusion models in population genetics. J. Appl. Prob. **1**, 177–232 (1964)
22. Kimura, M., Ohta, T.: Theoretical aspects of population genetics. Princeton: Princeton University Press 1971
23. Kurtz, T. G.: Semigroup of conditioned shifts and approximation of Markov processes. Ann. Probab. **3**, 618–642 (1975)
24. 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)
25. Levikson, B., Karlin, S.: Random temporal variation in selection intensities acting on infinite diploid populations: Diffusion method analysis. Theoret. Population Biology **8**, 292–300 (1975)
26. Li, W.-H. (ed.): Stochastic models in population genetics. Pennsylvania: Dowden, Hutchinson and Ross, Inc., 1977
27. Maruyama, T.: Stochastic problems in population genetics. New York: Springer 1977
28. Matsuda, H., Gojobori, T.: Protein polymorphism and fluctuation of environments. Adv. Biophys. **12**, 53–99 (1979)
29. Norman, M. F.: Diffusion approximation of non-Markovian processes. Ann. Probab. **3**, 358–364 (1975)
30. Okada, N.: On convergence to diffusion processes of Markov chains related to population genetics. Adv. Appl. Prob. **11**, 673–700 (1979)
31. Papanicolaou, G. C., Kohler, W.: Asymptotic theory of mixing stochastic ordinary differential equations. Comm. Pure Appl. Math. **27**, 641–668 (1974)
32. Sato, K.: Diffusion processes and a class of Markov chains related to population genetics. Osaka J. Math. **13**, 631–659 (1976a)
33. Sato, K.: A class of Markov chains related to selection in population genetics. J. Math. Soc. Japan **28**, 621–637 (1976b)
34. Sato, K.: Convergence of a class of Markov chains to multi-dimensional degenerate diffusion processes. Proc. Intern. Symp. Stoch. Diff. Eq. Kyoto 367–383 (1976c)
35. Sato, K.: Convergence to a diffusion of a multiallelic model in population genetics. Adv. Appl. Prob. **10**, 538–562 (1978)
36. Serfling, R. J.: Moment inequalities for the maximum cumulative sum. Ann. Math. Statist. **41**, 1227–1234 (1970)
37. Stroock, D. W., Varadhan, S. R. S.: Multidimensional diffusion processes. Berlin: Springer 1979
38. Takahata, N., Ishii, K., Matsuda, H.: Effect of temporal fluctuation of selection coefficient on gene frequency in a population. Proc. Nat. Acad. Sci. USA **72**, 4541–4545 (1975)
39. Trotter, H. F.: Approximation of semi-groups of operators. Pacific J. Math. **8**, 887–919 (1958)
40. Watterson, G. A.: Some theoretical aspects of diffusion theory in population genetics. Ann. Math. Statist. **33**, 939–957 (1962)

Received June 17, 1981/Revised February 1, 1982