COMMUNICATIONS ON PURE AND APPLIED MATHEMATICS, VOL. XXVII, 641–668 (1974)

Asymptotic Theory of Mixing Stochastic Ordinary Differential Equations*

G. C. PAPANICOLAOU
Courant Institute

AND

W. KOHLER
Virginia Polytechnic Institute and State University

1. Introduction

The mathematical theory of stochastic differential equations is concerned almost exclusively with the study of Itô equations and the associated Markov processes. This theory has found many diverse applications and has become a powerful tool in the study of diffusion processes (cf. [1], [2], [3], [4]). Nevertheless, many of its aspects are somewhat drastic idealizations of physical processes. Our aim here is to prove a theorem which shows that a very broad class of processes defined by stochastic differential equations, not of the Itô type, converge to diffusion Markov processes, thereby providing a more acceptable framework for working with such processes. Theorems of this form were enunciated by Stratonovich [5] and later a mathematical treatment was initiated by Stratonovich [6] and Khasminskii [7]. A theory for stochastic differential equations with limited after-effect was also initiated by Gikhman and Skorokhod [2] but with somewhat different objectives. The theorem we present here is an improvement on previous attempts (cf. [6], [7], [8]) both in the technical aspects of the proof as well as regarding its applications. We wish to point out that we have found the application of this theorem remarkably effective in a variety of problems (cf. [5], [9], [10], [11], [12], [13]). The reasons for this are discussed in the remarks following the statement of the theorem in Section 2. Section 3 contains the proof.

2. Formulation and Statement of the Theorem

Let $(\Omega, \mathscr{F}, P)$ be a probability space and let $\mathscr{F}_s^t$, $0 \leqq s \leqq t \leqq \infty$, be a family of σ-algebras contained in $\mathscr{F}$ and such that

$$\mathscr{F}_{s_1}^{t_1} \subset \mathscr{F}_{s_2}^{t_2}, \qquad 0 \leqq s_2 \leqq s_1 \leqq t_1 \leqq t_2 \leqq \infty .$$

* The work of the first author was performed at the Courant Institute of Mathematical Sciences and supported by the National Science Foundation under Grant No. NSF GP 32996x3. Reproduction in whole or in part is permitted for any purpose of the United States Government.

We assume that the σ-algebras $\mathscr{F}_s^t$ are mixing relative to P in the following sense:

$$\sup_{s \geqq 0} \sup_{\substack{A \in \mathscr{F}_{s+t}^{\infty} \\ B \in \mathscr{F}_0^s}} |P(A \mid B) - P(A)| = \rho(t) \downarrow 0 \quad \text{as} \quad t \uparrow \infty . \tag{2.1}$$

The monotonically decreasing function ρ is called the mixing rate. We assume it satisfies the rate condition

$$\int_0^\infty \rho^{1/2}(s)\, ds < \infty . \tag{2.2}$$

Ergodic Markov processes on a compact state space are examples of processes mixing in the above sense with, in fact, an exponential mixing rate.

We shall denote integration over Ω relative to P by $E\{\cdot\}$. We also assume that conditional probabilities relative to $\mathscr{F}_0^s$, $0 \leqq s \leqq \infty$, have a regular version so that with probability one we have the representation (cf. [14], p. 354 or [22], p. 139)

$$E\{. \mid \mathscr{F}_0^s\} = \int_\Omega \cdot \, P_s(d\omega \mid \omega') .$$

Let R^n denote the n-dimensional Euclidean space and $|x|$ the norm of vectors in R^n. The same symbol will be used for absolute values of scalars, and (x, y) denotes inner product of vectors. Let $F(\tau, x, t, \omega)$ be a function from $[0, T] \times R^n \times [0, \infty) \times \Omega$ into R^n, where T is a fixed positive number. F is called a random vector field and we assume it satisfies the following conditions (cf. also Remark 4):

(i) F is jointly measurable with respect to its arguments. For fixed τ, x and t, $F(\tau, x, t, \omega)$ is $\mathscr{F}_t^t$ measurable as a function of $\omega \in \Omega$.

(ii) There is a constant C independent of τ, x, t, and ω such that

$$|F_i(\tau, x, t, \omega)| \leqq C(1 + |x|) , \tag{2.3}$$

$$\left| \frac{\partial F_i(\tau, x, t, \omega)}{\partial x_j} \right| \leqq C , \qquad i, j = 1, 2, \cdots, n . \tag{2.4}$$

Furthermore, there is an integer $q \geqq 0$ such that

$$\left| \frac{\partial^2 F_i(\tau, x, t, \omega)}{\partial x_j \, \partial x_k} \right| \leqq C(1 + |x|^q) , \tag{2.5}$$

$$\begin{aligned} \left| \frac{\partial^3 F_i(\tau, x, t, \omega)}{\partial x_j \, \partial x_k \, \partial x_l} \right| &\leqq C(1 + |x|^q) , \\ \left| \frac{\partial^4 F_i(\tau, x, t, \omega)}{\partial x_j \, \partial x_k \, \partial x_l \, \partial x_m} \right| &\leqq C(1 + |x|^q) . \end{aligned} \qquad i, j, k, l, m = 1, \cdots, n , \tag{2.6}$$

(iii) There is a constant C as above such that

$$E^{1/2}\{(F_i(s+h,x,t)-F_i(s,x,t))^2\} \leqq Ch(1+|x|)\,,$$

$$E^{1/2}\left\{\left(\frac{\partial F_i}{\partial x_j}(s+h,x,t)-\frac{\partial F_i}{\partial x_j}(s,x,t)\right)^2\right\} \leqq Ch\,. \tag{2.7}$$

Here s and $s+h$ are in $[0, T]$ and C is independent of s and h.

Throughout we adopt the convention that C denotes a constant, not necessarily the same each time, and usually C will be independent of variable parameters. The dependence of C on specific items is denoted explicitly as an argument.

The object of our study is the asymptotic analysis of the stochastic ordinary differential equations

$$\begin{gathered}\frac{dx(t,s,x,\omega)}{dt} = \varepsilon F(\varepsilon^2 t, x(t,s,x,\omega), t, \omega) + \varepsilon^2 G(\varepsilon^2 t, x(t,s,x,\omega), t, \omega)\,, \quad t>s\,,\\ x(s,s,x,\omega) = x \in R^n\,.\end{gathered} \tag{2.8}$$

Here $G(\tau, x, t, \omega)$ is a random vector field satisfying the same conditions as F, and ε is a real parameter in $(0, 1]$. The usual existence and uniqueness theory for ordinary differential equations gives us a solution $x(t, s, x, \omega)$ of (2.8) for each $\omega \in \Omega$ which is $\mathscr{F}_s^t$ measurable and continuous as a function of t. We have therefore a well defined stochastic process. We are interested in the behavior of this process as $\varepsilon \to 0$ and $t \to \infty$ with $\varepsilon^2 t$ remaining fixed. For this purpose we define scaled variables by

$$\tau = \varepsilon^2 t, \qquad \sigma = \varepsilon^2 s, \qquad x^{(\varepsilon)}(\tau,\sigma,x) = x\left(\frac{\tau}{\varepsilon^2}, \frac{\sigma}{\varepsilon^2}, x\right), \tag{2.9}$$

and rewrite (2.8) in the form

$$\begin{gathered}\frac{dx^{(\varepsilon)}(\tau,\sigma,x)}{d\tau} = \frac{1}{\varepsilon}F\left(\tau, x^{(\varepsilon)}(\tau,\sigma,x), \frac{\tau}{\varepsilon^2}\right) + G\left(\tau, x^{(\varepsilon)}(\tau,\sigma,x), \frac{\tau}{\varepsilon^2}\right),\\ 0 \leqq \sigma < \tau \leqq T, \quad x^{(\varepsilon)}(\sigma,\sigma,x) = x \in R^n\,.\end{gathered} \tag{2.10}$$

Note that we have not displayed dependence on ω in (2.9) and (2.10) as is customary. We shall also write $x^{(\varepsilon)}(\tau, \sigma, x) = x^{(\varepsilon)}(\tau)$ or $x^{(\varepsilon)}$ when convenient. From (2.9) it follows that $x^{(\varepsilon)}(\tau, \sigma, x)$ is $\mathscr{F}_{\sigma/\varepsilon^2}^{\tau/\varepsilon^2}$ measurable.

We shall show that the process $x^{(\varepsilon)}$ converges weakly to a diffusion Markov process as $\varepsilon \to 0$. Because of the factor $1/\varepsilon$ in front of F in (2.10), it is clear that if a limit is to exist as $\varepsilon \to 0$, the random vector field F must be centered.

We assume therefore that, for each fixed τ, x and t,

$$E\{F(\tau, x, t)\} = 0 . \tag{2.11}$$

To describe the limiting diffusion process we introduce diffusion and drift coefficients by the limits (cf. also Remark 9)

$$a^{ij}(\tau, x) = \lim_{\varepsilon\to 0} \frac{1}{\varepsilon^3} \int_\tau^{\tau+\varepsilon} \int_\tau^\sigma E\left\{F_i\left(\tau, x, \frac{s}{\varepsilon^2}\right) F_j\left(\tau, x, \frac{\sigma}{\varepsilon^2}\right)\right\} ds\, d\sigma , \tag{2.12}$$

$$\begin{aligned} b^j(\tau, x) &= \lim_{\varepsilon\to 0} \frac{1}{\varepsilon^3} \int_\tau^{\tau+\varepsilon} \int_\tau^\sigma \sum_{i=1}^n E\left\{F_i\left(\tau, x, \frac{s}{\varepsilon^2}\right) \frac{\partial F_j\left(\tau, x, \frac{\sigma}{\varepsilon^2}\right)}{\partial x_i}\right\} ds\, d\sigma \\ &+ \lim_{\varepsilon\to 0} \frac{1}{\varepsilon} \int_\tau^{\tau+\varepsilon} E\left\{G_j\left(\tau, x, \frac{s}{\varepsilon^2}\right)\right\} ds , \qquad i, j = 1, 2, \cdots, n . \end{aligned} \tag{2.13}$$

The manner in which these limits are taken, along with a rate of approach, is specified by the following inequalities:

$$\left|\frac{1}{\varepsilon^3} \int_\tau^{\tau+\varepsilon} \int_\tau^\sigma E\left\{F_i\left(\tau, x, \frac{s}{\varepsilon^2}\right) F_j\left(\tau, x, \frac{\sigma}{\varepsilon^2}\right)\right\} ds\, d\sigma - a^{ij}(\tau, x)\right| \leqq \varepsilon C(1 + |x|^2) , \tag{2.14}$$

$$\begin{aligned} &\left|\frac{1}{\varepsilon^3} \int_\tau^{\tau+\varepsilon} \int_\tau^\sigma \sum_{i=1}^n E\left\{F_i\left(\tau, x, \frac{s}{\varepsilon^2}\right) \frac{\partial F_j\left(\tau, x, \frac{\sigma}{\varepsilon^2}\right)}{\partial x_i}\right\} ds\, d\sigma \right. \\ &\left. + \frac{1}{\varepsilon} \int_\tau^{\tau+\varepsilon} E\left\{G_j\left(\tau, x, \frac{s}{\varepsilon^2}\right)\right\} ds - b^j(\tau, x)\right| \leqq \varepsilon C(1 + |x|) . \end{aligned} \tag{2.15}$$

As before the constant C is independent of $\tau \in [0, T]$, $x \in R^n$, $\varepsilon \in (0, 1]$ and $i, j = 1, 2, \cdots, n$. The diffusion and drift coefficients inherit some regularity properties from F and G. We require however other properties as well; therefore we specify them directly below. From the definition (2.12) it follows that the diffusion matrix $(a^{ij}(\tau, x))$ is non-negative definite but we do not assume that it is nondegenerate.

Let $C^{k,p}(R^n)$ denote the collection of functions on R^n with continuous derivatives up to order $1, 2, \cdots, k$ for which there exists an integer $p \geqq 0$ such that

$$\left|\frac{\partial^\alpha f(x)}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \cdots \partial x_n^{\alpha_n}}\right| \leqq C(1 + |x|^p) , \qquad x \in R^n , \tag{2.16}$$

$$\alpha_j \text{ non-negative integers, } 0 \leqq \alpha_1 + \cdots + \alpha_n = \alpha \leqq k .$$

Let $\mathscr{L}_\sigma$ be a differential operator defined on $C^{2,p}(R^n)$, $p \geqq 0$, by

$$\mathscr{L}_\sigma f(x) = \sum_{i,j=1}^{n} a^{ij}(\sigma, x) \frac{\partial^2 f(x)}{\partial x_i \partial x_j} + \sum_{j=1}^{n} b^j(\sigma, x) \frac{\partial f(x)}{\partial x_j}, \qquad \sigma \in [0, T], \tag{2.17}$$

and consider the possibly degenerate parabolic final value problem

$$\frac{\partial u(\sigma, \tau, x)}{\partial \sigma} + \mathscr{L}_\sigma u(\sigma, \tau, x) = 0, \qquad 0 \leqq \sigma < \tau \leqq T,$$

$$u(\tau, \tau, x) = g(x), \qquad g(x) \in C^{k,p}, \qquad k \geqq 2, \quad p \geqq 0.$$

We wish to know when this partial differential equation has a classical solution which is unique and what regularity properties it has for $g \in C^{k,p}$. These facts are required in our asymptotic analysis in much the same way as similar results were required by Khinchin in proving diffusion approximations in his well known monograph [15]. It is possible to avoid using any *a priori* information about partial differential equations but we prefer the present method which we followed in [8] also. The facts we need are provided by the theory of Oleinik (cf. [16], [17], [18]; appendix) and can be modified to include the case $g \in C^{k,p}$ with $p \geqq 1$. It is also possible to use the Itô theory of stochastic differential equations as presented in [1], [2]. For this latter theory it is required that $(a^{ij}(\tau, x))$ have a sufficiently smooth symmetric square root but it is not easy to ascertain in general when this is possible (cf. [18]). We shall state our requirements here in the Itô context not losing generality in view of the above remarks (cf. also Remark 8 below).

Let $\mathscr{L}_\sigma$ be defined by (2.17) and assume the following:

(i) $(a^{ij}(\tau, x))$ has a symmetric square root $(c^{ij}(\tau, x))$,

$$a^{ij}(\tau, x) = \sum_k c^{ik}(\tau, x) c^{jk}(\tau, x). \tag{2.18}$$

(ii) $(c^{ij}(\tau, x))$ and $(b^j(\tau, x))$ satisfy the bounds

$$|c^{ij}(\tau, x)| \leqq C(1 + |x|), \qquad |b^j(\tau, x)| \leqq C(1 + |x|), \tag{2.19}$$

$$\left| \frac{\partial c^{ij}(\tau, x)}{\partial x_k} \right| \leqq C, \qquad \left| \frac{\partial b^j(\tau, x)}{\partial x_k} \right| \leqq C, \tag{2.20}$$

and there is an integer $q \geqq 0$ such that

$$c^{ij}(\tau, x) \in C^{4,q}(R^n), \tag{2.21}$$

$$b^j(\tau, x) \in C^{4,q}(R^n), \qquad i, j, k = 1, \cdots, n. \tag{2.22}$$

(iii) Moreover,

$$|a^{ij}(\tau, x) - a^{ij}(\tau + h, x)| \leqq Ch(1 + |x|^2)\,, \tag{2.23}$$

$$|b^j(\tau, x) - b^j(\tau + h, x)| \leqq Ch(1 + |x|)\,. \tag{2.24}$$

We recall here the convention about constants and note the similarity of (2.19)–(2.24) and (2.3)–(2.7) which is not accidental and explains why we state the conditions on $\mathscr{L}_\sigma$ in the Itô form.

We are ready to present the main result of this paper.

THEOREM. *Let* $x^{(\varepsilon)}(\tau, \sigma, x)$, $0 \leqq \sigma \leqq \tau \leqq T$, $x \in R^n$, $\varepsilon \in (0, 1]$, *be the process defined by* (2.10). *Let all hypotheses stated above* (*i.e.*, (2.1)–(2.7), (2.11)–(2.15) *and* (2.18)–(2.24)) *hold. Then the processes* $x^{(\varepsilon)}$ *converge weakly as* $\varepsilon \to 0$ *to a diffusion Markov process* $x^{(0)}$ *with infinitesimal generator* $\mathscr{L}_\sigma$ *defined by* (2.17). *Furthermore, let* $f(x) \in C^{4,p}(R^n)$, $p \geqq 0$, *and let* $u(\sigma, \tau, x)$ *denote the solution of*

$$\frac{\partial u(\sigma, \tau, x)}{\partial \sigma} + \mathscr{L}_\sigma u(\sigma, \tau, x) = 0\,, \qquad 0 \leqq \sigma \leqq \tau \leqq T\,, \tag{2.25}$$
$$u(\tau, \tau, x) = f(x)\,.$$

Then there exists an integer $\bar{p} \geqq p + 4$ *such that*

$$|E\{f(x^{(\varepsilon)}(\tau, \sigma, x)) \,|\, \mathscr{F}_0^{\sigma/\varepsilon^2}\} - u(\sigma, \tau, x)| \leqq \varepsilon C(f, T) \cdot (1 + |x|^{\bar{p}})\,. \tag{2.26}$$

Here $C(f, T)$ denotes a constant which depends on f and its derivatives up to order 4, on T and p (and other quantities) but is independent of x and ε. When q in (2.5), (2.6) and in (2.21), (2.22) is zero, then $\bar{p} = p + 4$.

Before continuing with the proof of this theorem in Section 3 we enumerate several remarks that may be helpful for the theory as well as the application.

Remark 1. The result (2.26) is not implied by weak convergence. In fact we use it here to prove weak convergence. As a consequence of (2.26), moments of any order of $x^{(\varepsilon)}(\tau, \sigma, x)$ converge to the corresponding moments of the diffusion process $x^{(0)}(\tau, \sigma, x)$. This is quite useful in applications (cf. [9], [10], [12], [13]). The $O(\varepsilon)$ error estimate on the right side of (2.20) is the best possible as in [8].

Remark 2. The theorem puts no restrictions of nondegeneracy at all on the operator $\mathscr{L}_\sigma$ and, in view of (2.3)–(2.7), linear equations are included. Therefore, we can apply it to linear matrix stochastic differential equations and processes on matrix Lie groups. Specifically we can apply the present theorem to recover Theorem 3 of [8] with the additional result of weak convergence and the very useful estimate (2.26) with $p \geqq 1$ (cf. [10] for an interesting application of this). Let X denote the linear space of $n \times n$ real matrices with inner product $(A, B) = \operatorname{tr}(A B^T)$ and norm $|A| = \sqrt{(A, A)}$. Consider the X-valued stochastic process $Y(t, s, Z, \omega)$ satisfying the stochastic differential equation

$$\begin{aligned} \frac{dY(t, s, Z, \omega)}{dt} &= AY(t, s, Z, \omega) + \varepsilon \tilde{M}(t, \omega) Y(t, s, Z, \omega), \qquad t > s, \\ Y(s, s, Z, \omega) &= Z \in X. \end{aligned} \tag{2.27}$$

Here A is a fixed $n \times n$ matrix and $\tilde{M}(t, \omega)$ is an $n \times n$ matrix-valued process such that

$$E\{\tilde{M}(t, \cdot)\} = 0. \tag{2.28}$$

Furthermore, we assume that there is a C independent of t and ω such that

$$|e^{-At} \tilde{M}(t, \omega) e^{At}| \leqq C, \qquad \omega \in \Omega, \ t \geqq 0. \tag{2.29}$$

We denote exponentials of matrices by e^A. The process $\tilde{M}(t, \omega)$ is assumed $\mathscr{F}_t^t$ measurable for each t and the σ-algebras $\mathscr{F}_s^t$ satisfy all hypotheses in the beginning of the section. We introduce scaled variables by (2.9) and set (omit the ω's)

$$Y^{(\varepsilon)}(\tau, \sigma, Z) = e^{-A\tau/\varepsilon^2} Y\left(\frac{\tau}{\varepsilon^2}, \frac{\sigma}{\varepsilon^2}, Z\right) e^{A\sigma/\varepsilon^2}. \tag{2.30}$$

Clearly, $Y^{(\varepsilon)}$ satisfies the linear matrix stochastic equation

$$\begin{aligned} \frac{dY^{(\varepsilon)}(\tau, \sigma, Z)}{d\tau} &= \frac{1}{\varepsilon} e^{-A\tau/\varepsilon^2} \tilde{M}\left(\frac{\tau}{\varepsilon^2}\right) e^{A\tau/\varepsilon^2} Y^{(\varepsilon)}(\tau, \sigma, Z), \\ 0 \leqq \sigma < \tau \leqq T, &\qquad Y^{(\varepsilon)}(\sigma, \sigma, Z) = Z. \end{aligned} \tag{2.31}$$

It is convenient to define the matrix-valued process $M(t)$ by

$$M(t) = e^{-At} \tilde{M}(t) e^{At}, \tag{2.32}$$

which simplifies the notation below.

We can now apply the above theorem to study the asymptotic behavior of the processes $Y^{(\varepsilon)}$ as $\varepsilon \to 0$, $0 \leqq \sigma \leqq \tau \leqq T$ and $Z \in X$. Let $C^{k,p}(X)$ denote the space of k-times continuously differentiable functions on X satisfying (2.16).

If $Z = (Z_{ij})$, then we denote by ∂_Z the matrix of differential operators $(\partial/\partial Z_{ij})$. Assume now that for $f \in C^{2,p}(X)$ the limit

$$\mathscr{L}f(Z) = \lim_{t\to\infty} \frac{1}{t}\int_{t_0}^{t_0+t}\int_{t_0}^{s} E\{(M(\sigma)Z, \partial_Z)(M(s)Z, \partial_Z)f(Z)\}\, d\sigma\, ds \tag{2.33}$$

exists uniformly in t_0 and is approached faster than $(C/t)(1 + |Z|^{2+p})$, $t \uparrow \infty$. Then, according to the theorem, the processes $Y^{(\varepsilon)}(\tau, \sigma, Z)$ converge weakly as $\varepsilon \to 0$ to the time-homogeneous diffusion process with infinitesimal generator $\mathscr{L}$ given by (2.33). Moreover, moments of any order go to the corresponding moments of the limiting Markov process (i.e., (2.26) is valid) and the error in the approximation is $O(\varepsilon)$.

The differential operator $\mathscr{L}$ is usually degenerate and in the case of matrix groups it acts tangentially on the group manifold which is thought of as a submanifold in X. Thus the probability distribution will be entirely concentrated on the group manifold. In the applications it is important to select coordinate systems in X adapted to the group manifold so that redundant variables are eliminated (cf. for example [8], [9], [10], [11]).

Moments of the $Y^{(\varepsilon)}$ process can be computed easily by using (2.26). By direct computation we find that, for $f(Z) = Z$,

$$(M(\sigma)Z, \partial_Z)\ (M(s)Z, \partial_Z)Z = M(s)M(\sigma)Z$$

and, for $f(Z) = Z \otimes Z$,

$$\begin{aligned}&(M(\sigma)Z, \partial_Z)\ (M(s)Z, \partial_Z) \cdot Z \otimes Z \\ &\qquad = (M(s) \otimes I + I \otimes M(s))\ (M(\sigma) \otimes I + I \otimes M(\sigma))Z \otimes Z\,.\end{aligned} \tag{2.34}$$

Let us denote by $\Lambda_1 \in X$ and $\Lambda_2 \in X \otimes X$ the limits

$$\begin{aligned}\Lambda_1 &= \lim_{t\uparrow\infty} \frac{1}{t}\int_{t_0}^{t_0+t}\int_{t_0}^{s} E\{M(s)M(\sigma)\}\, d\sigma\, ds\,, \\ \Lambda_2 &= \lim_{t\uparrow\infty} \frac{1}{t}\int_{t_0}^{t_0+t}\int_{t_0}^{s} E\{(M(s) \otimes I + I \otimes M(s))(M(\sigma) \otimes I + I \otimes M(\sigma))\}\, d\sigma\, ds\,.\end{aligned} \tag{2.35}$$

Then, from (2.35), (2.34) and (2.33), we obtain

$$\mathscr{L}Z = \Lambda_1 Z\,, \qquad \mathscr{L}Z \otimes Z = \Lambda_2 Z \otimes Z\,. \tag{2.36}$$

Similar formulas can easily be obtained for higher-order tensor products. The algebraic identities (2.36) and the limit theorem lead to the conclusion that

$$
\begin{aligned}
&\lim_{\varepsilon \downarrow 0} E\{Y^{(\varepsilon)}(\tau, \sigma, Z) \mid \mathscr{F}_0^{\sigma/\varepsilon^2}\} = e^{\Lambda_1(\tau-\sigma)}Z, \\
(2.37) \qquad &\lim_{\varepsilon \downarrow 0} E\{Y^{(\varepsilon)}(\tau, \sigma, Z) \otimes Y^{(\varepsilon)}(\tau, \sigma, Z) \mid \mathscr{F}_0^{\sigma/\varepsilon^2}\} = e^{\Lambda_2(\tau-\sigma)}Z \otimes Z.
\end{aligned}
$$

The result (2.37) was also obtained in [19] by a formal perturbation argument. A group-theoretic version is contained in Theorem 4 of [8] but the fact that moments converge to moments was not proved there. The role played by A in (2.31) is exploited in [9], [10], [11], [12].

Remark 3. The theorem of Khasminskii [7] differs from the present one in the following respects. Linear problems are ruled out in this theorem since, instead of (2.3), F and G are uniformly bounded in x. The vector fields F and G are also assumed independent of τ. The estimate (2.26) holds only insofar as it is implied by the fourth moment estimate used to show weak convergence (i.e., for low $p \geqq 0$) and the right side of (2.26) is $o(1)$ instead of $O(\varepsilon)$.

Remark 4. Hypotheses (2.3) and (2.4) seem excessively stringent, that is, the constant C should have been allowed to depend on ω and t and $C(\omega, t)$ should perhaps have uniformly bounded (in t) moments of sufficiently high order. We have not succeeded in weakening (2.3) and (2.4). It is interesting to note however that in [2], Part II, Chapter 1, a theory of stochastic equations with limited after-effect is developed in which conditions analogous to (2.3) and (2.4) are imposed, i.e., stringent uniform bounds relative to ω are required. The theory of [2] can nevertheless handle stochastic equations for diffusions as well as jump Markov processes. It seems to us that (2.8) or (2.10) as formulated here fit many physical problems rather well and in particular conditions such as (2.3) do not affect the applicability of the results seriously.

Remark 5. An important aspect of the limit theorem is that the random vector fields $F(\tau, x, t)$ and $G(\tau, x, t)$ are not assumed to be stationary processes for fixed τ and x. Indeed, (2.32) shows that, for example, even if $\tilde{M}(t)$ were stationary, $M(t)$ would not be. Instead of stationarity, (2.12) and (2.13) are assumed to hold uniformly in $\tau \in [0, T]$, or the limit in (2.33) is assumed to be independent of t_0 uniformly for $t_0 \geqq 0$. We are thus performing an averaging, as in Bogolyubov's method of averaging [20] for deterministic ordinary differential equations, *simultaneously* with the diffusion Markov process approximation. This is the reason why the theorem, when applicable, is so effective (cf. [5], [9]–[13]). If we set $F \equiv 0$ in (2.10) and assume that G is deterministic, then the above theorem reduces to the method of averaging for deterministic equations (cf. [20]). Other aspects of the averaging principle for stochastic equations along with references to earlier work of Bogolyubov, Krylov and Gikhman can be found in Section 14, Part II of [2].

Remark 6. The present theorem and its proof can be formulated in terms of spaces and operators as in [8]. We do not do this here because it would require a somewhat elaborate asbstract framework which seems of little use beyond problem (2.10). Our proof however, especially Lemmas 5–8, is almost identical with the proof of Theorem 1 in [8]. An interesting survey of operator methods for stochastic equations and other applications is given by Hersh [21] along with many references to related work.

Remark 7. Discrete time problems, i.e., random difference equations or, in the case of the example in Remark 2, products of random matrices are included in the present theorem. One simply has to take the random fields F and G piecewise constant in some appropriate way and, usually, statistically independent in different segments. In this kind of setup, proofs become considerably simpler and many hypotheses introduced above are superfluous.

Remark 8. Let F and G in (2.10) have the form

$$\text{(2.38)}\qquad \begin{aligned} F_i\left(\tau, x, \frac{\tau}{\varepsilon^2}\right) &= \sum_{k=1}^{N} \tilde{F}_i^k(\tau, x)\eta_k\left(\frac{\tau}{\varepsilon^2}\right), \\ G_i\left(\tau, x, \frac{\tau}{\varepsilon^2}\right) &= \tilde{G}_i(\tau, x). \end{aligned} \qquad i = 1, 2, \cdots, n,$$

Here $\tilde{F}^k(\tau, x)$ and $\tilde{G}(\tau, x)$, $k = 1, \cdots, N$, are deterministic vector fields satisfying (2.3)–(2.6) and $\eta_k(t) = \eta_k(t, \omega)$ are stationary zero-mean processes, $\mathscr{F}_t^t$ measurable and bounded. The σ-algebras $\mathscr{F}_s^t$ satisfy the mixing hypothesis stated at the beginning of this section. We define the noise intensity matrix (γ_{kl}) by

$$\text{(2.39)}\qquad \gamma_{kl} = \int_0^\infty R_{kl}(s)\, ds, \qquad R_{kl}(s) = E\{\eta_k(s+t)\eta_l(t)\}, \qquad k, l = 1, 2, \cdots, N.$$

In view of (2.2) the integral in (2.39) converges.

We compute now the diffusion and drift coefficients defined by (2.12) and (2.13) when F and G are given by (2.38):

$$\text{(2.40)}\qquad a^{ij}(\tau, x) = \sum_{k,l=1}^{N} \gamma_{kl} \tilde{F}_i^k(\tau, x) \tilde{F}_j^l(\tau, x),$$

$$\text{(2.41)}\qquad b^j(\tau, x) = \sum_{k,l=1}^{N} \sum_{i=1}^{n} \gamma_{kl} \tilde{F}_i^k(\tau, x) \frac{\partial \tilde{F}_j^l(\tau, x)}{\partial x_i} + \tilde{G}_j(\tau, x), \qquad i, j = 1, 2, \cdots, n.$$

According to our theorem, the solution process $x^{(\varepsilon)}(\tau)$ of the system of stochastic differential equations

$$\frac{dx^{(\varepsilon)}(\tau)}{d\tau} = \tilde{G}(\tau, x^{(\varepsilon)}(\tau)) + \sum_{k=1}^{N} \tilde{F}^k(\tau, x^{(\varepsilon)}(\tau)) \frac{1}{\varepsilon} \eta_k\left(\frac{\tau}{\varepsilon^2}\right) \tag{2.42}$$

converges weakly to a diffusion Markov process with diffusion and drift coefficients given by (2.40) and (2.41). Furthermore, moments of any order converge to the corresponding moments of the diffusion process i.e., (2.26) holds. This result gives a physically acceptable interpretation of solutions to stochastic equations with white noise coefficients. Indeed, the processes $\eta_k^{(\varepsilon)}(\tau) = (1/\varepsilon)\eta_k(\tau/\varepsilon^2)$, $k = 1, \cdots, N$, converge to Gaussian white noise by the central limit theorem. It is actually more appropriate to consider integrals of $\eta_k^{(\varepsilon)}(\tau)$ which tend to Brownian motion. Note however that the solutions $x^{(\varepsilon)}(\tau)$ tend to a limit that differs from that of the Itô theory because of the differentiated terms that appear in (2.41). Note further the formal similarity of the white noise approximations $\eta_k^{(\varepsilon)}(\tau)$ to approximations of delta functions by integrable functions.

The problem of analyzing the limit as $\varepsilon \to 0$ in (2.42) has received considerable attention (cf. [6], [24], and references therein). In many applications it is the averaging, as we pointed out in Remark 5, that is the principal reason for the effectiveness of the theorem. More specifically, suppose that instead of (2.42) we have the system

$$\frac{dx^{(\varepsilon)}(\tau)}{d\tau} = \tilde{G}\left(\tau, x^{(\varepsilon)}(\tau), \frac{\tau}{\varepsilon^2}\right) + \sum_{k=1}^{N} \tilde{F}^k\left(\tau, x^{(\varepsilon)}(\tau), \frac{\tau}{\varepsilon^2}\right) \frac{1}{\varepsilon} \eta_k\left(\frac{\tau}{\varepsilon^2}\right), \tag{2.43}$$

where $\tilde{F}^k(\tau, x, t)$ and $\tilde{G}(\tau, x, t)$ are deterministic vector fields satisfying (2.3)–(2.7). For the asymptotic analysis of (2.43) one needs the full strength of the present theorem and (2.12), (2.13) cannot be significantly simplified. It may be possible to assume, however, that there exist two parameters ε_1 and ε_2 such that $\varepsilon_1 \ll \varepsilon_2$ and in (2.43) we have

$$\tilde{G} \to \tilde{G}\left(\tau, x, \frac{\tau}{\varepsilon_2^2}\right), \qquad \tilde{F} \to \tilde{F}\left(\tau, x, \frac{\tau}{\varepsilon_2^2}\right),$$

$$\frac{1}{\varepsilon}\eta_k\left(\frac{\tau}{\varepsilon^2}\right) \to \frac{1}{\varepsilon_1}\eta_k\left(\frac{\tau}{\varepsilon_1^2}\right).$$

Then we may apply the theorem first with respect to ε_1, in a manner similar to (2.42), and subsequently use the method of averaging to the diffusion Markov process relative to ε_2 (cf. [2]). This yields a result that is different from the case $\varepsilon_1 \sim \varepsilon_2$ to which our theorem applies (cf. the example in Remark 2). Since the separation of scales ε_1 and ε_2 described above does not lead to anything simpler and since our theorem covers the full range $\varepsilon_1 \sim \varepsilon_2$, the separation of scales approach is clearly less useful. From the physical point of view, in examples such as the one in Remark 2, one of the most interesting features emerging in the limit is the way the stochastic fluctuations and the deterministic rapid variations interact. This is suppressed in the separation of scales approach.

The symmetric part of the noise intensity matrix (γ_{kl}) in (2.39), which enters in the diffusion coefficient matrix in (2.40), is non-negative definite since it is a power spectrum matrix. Therefore it has a non-negative symmetric square root. From (2.40) and this observation it follows that (2.18) is automatically satisfied for the system (2.42). This is also the case in many other applications. Furthermore, for (2.42) the conditions (2.19)–(2.21) are also satisfied automatically and (2.7) is unnecessary.

3. Proof of the Theorem

The proof is divided into several steps, some of which are stated as lemmas. To simplify the notation we assume without loss of generality that G in (2.10) is identically zero. The difficulty of the theorem is due to F. The first lemma is well known but we include it here for completeness. It is used frequently in the sequel. Lemma 3 is an important estimate in the proof. Lemma 4 summarizes some results about (2.25) which are needed. Lemmas 5–8 are almost a step by step adaptation to the present context of the proof of Theorem 1 in [8]. That proof should be consulted for motivation if necessary. The fourth moment condition for weak compactness which we use here is well known (cf. [22], p. 450).

LEMMA 1. *Let $F(\omega, \omega')$ be a function on $\Omega \times \Omega$ such that, for fixed ω', $F(\cdot, \omega')$ is $\mathscr{F}_{t+s}^{\infty}$ measurable and, for fixed ω, $F(\omega, \cdot)$ is $\mathscr{F}_0^s$ measurable and $|F(\omega, \omega')| \leqq \phi(\omega')$. Let $\mathscr{F}_s^t$ and P satisfy the hypotheses of Section 2 and set*

$$\bar{F}(\omega') = E\{F(\cdot, \omega')\} = \int_\Omega F(\omega, \omega') P(d\omega) .$$

Then

$$|E\{F(\cdot, \omega') \mid \mathscr{F}_0^s\} - \bar{F}(\omega')| \leqq 2\rho(t)\phi(\omega') . \tag{3.1}$$

Proof: We have

$$\begin{aligned} |E\{F(\cdot, \omega') \mid \mathscr{F}_0^s\} - \bar{F}(\omega')| &= \left|\int_\Omega F(\omega, \omega')[P_s(d\omega \mid \omega') - P(d\omega)]\right| \\ &= \left|\int_\Omega F(\omega, \omega')\mu_s(d\omega \mid \omega')\right| . \end{aligned} \tag{3.2}$$

Here $\mu_s(A \mid \omega') = P_s(A \mid \omega') - P(A)$ is a signed measure and, from (2.1) and the Hahn decomposition theorem, its variation $|\mu_s|\,(A \mid \omega')$ satisfies

$$\sup_{A\in\mathscr{F}_{t+s}^{\infty}} |\mu_s|\,(A \mid \omega') \leqq 2\rho(t)\,. \tag{3.3}$$

This estimate applied to (3.2) together with the hypothesis $|F(\omega, \omega')| \leqq \phi(\omega')$ completes the proof of (3.1).

We state without proof (cf. [23], p. 222) a related estimate which we use in Lemma 6.

If $|F(\omega, \omega')| \leqq \psi(\omega)\phi(\omega')$ and $E\{\phi^p\} < \infty$, $E\{\psi^q\} < \infty$ with $1/p + 1/q = 1$, $p,q > 1$, then

$$|E\{F\} - E\{\bar{F}\}| \leqq 2\rho^{1/p}(t)E^{1/p}\{\phi^p\}E^{1/q}\{\psi^q\}. \tag{3.4}$$

We rewrite (2.10) as an integral equation:

$$x^{(\varepsilon)}(\tau, \sigma, x) = x + \frac{1}{\varepsilon}\int_\sigma^\tau F\left(s, x^{(\varepsilon)}(s, \sigma, x), \frac{s}{\varepsilon^2}\right) ds\,. \tag{3.5}$$

We have the following simple estimates.

LEMMA 2. *For* $0 \leqq \tau - \sigma \leqq \varepsilon$,

$$\begin{gathered} |x^{(\varepsilon)}(\tau, \sigma, x)| \leqq C(1 + |x|), \qquad \left|\frac{\partial x^{(\varepsilon)}(\tau, \sigma, x)}{\partial x}\right| \leqq C\,, \\ \left|\frac{\partial^2 x^{(\varepsilon)}(\tau, \sigma, x)}{\partial x^2}\right| \leqq C(1 + |x|^q)\,. \end{gathered} \tag{3.6}$$

Proof: First we recall the convention about constants stated in Section 2 and write $\partial/\partial x$ for any first partial and $\partial^2/\partial x^2$ for any second partial derivative. The estimates (3.6) follow from Gronwall's inequality applied to (3.5) after using (2.3), (2.4) and (2.5).

We come now to a basic lemma in our proof.

LEMMA 3. *For any integer* $p \geqq 1$ *and* $0 \leqq \sigma \leqq \tau \leqq T$,

$$E\{|x^{(\varepsilon)}(\tau, \sigma, x)|^p \mid \mathscr{F}_0^{\sigma/\varepsilon^2}\} \leqq C(1 + |x|^p)\,, \tag{3.7}$$

where the constant C *is independent of* ε, τ *and* σ *but depends on* T *and* p.

Proof: We consider first the case when p is an even integer, say $2p$. Divide the interval $[\sigma, \tau]$ into segments of length ε and assume with no loss of generality that ε takes on values for which $m = (\tau - \sigma)/\varepsilon$ is an integer going to $+\infty$ as ε goes to 0.

From (3.5) we deduce by differentiation and integration that the following identity holds:

$$\begin{aligned}(3.8)\qquad &|x^{(\varepsilon)}(\tau, \sigma, x)|^{2p} \\ &= |x|^{2p} + \frac{2p}{\varepsilon}\int_\sigma^\tau \left(F\left(s, x^{(\varepsilon)}(s, \sigma, x), \frac{s}{\varepsilon^2}\right), x^{(\varepsilon)}(s, \sigma, x)\right) |x^{(\varepsilon)}(s, \sigma, x)|^{2(p-1)}\, ds\,.\end{aligned}$$

Using the decomposition of the interval $[\sigma, \tau]$ we rewrite (3.8), with slightly simplified notation, in the form

(3.9)

$$|x^{(\varepsilon)}(\tau)|^{2p} = |x|^{2p} + \frac{2p}{\varepsilon}\sum_{k=0}^{m-1}\int_{\sigma+k\varepsilon}^{\sigma+(k+1)\varepsilon} \left(F\left(s, x^{(\varepsilon)}(s), \frac{s}{\varepsilon^2}\right), x^{(\varepsilon)}(s)\right) |x^{(\varepsilon)}(s)|^{2(p-1)}\, ds\,.$$

To simplify the notation further we write

$$\begin{aligned}&\sigma_k = \sigma + k\varepsilon, \qquad x_k^{(\varepsilon)} = x^{(\varepsilon)}(\sigma_k, \sigma, x), \qquad \mathscr{F}_k = \mathscr{F}_0^{\sigma_k/\varepsilon^2}, \\ (3.10)\qquad &I_k^{(\varepsilon)}(x) = \int_{\sigma_k}^{\sigma_{k+1}} \left(F\left(s, x^{(\varepsilon)}(s), \frac{s}{\varepsilon^2}\right), x^{(\varepsilon)}(s)\right) |x^{(\varepsilon)}(s)|^{2(p-1)}\, ds\,, \\ &\qquad\qquad k = 0, 1, \cdots, m-1\,.\end{aligned}$$

For $I_0^{(\varepsilon)}(x)$ we have the simple estimate

$$(3.11)\qquad |I_0^{(\varepsilon)}(x)| \leqq \varepsilon C(1 + |x|^{2p})\,.$$

Here we use (2.3) and the local estimate (3.6) of Lemma 2. In the notation (3.10) and in view of (3.11), (3.9) yields

$$(3.12)\qquad E\{|x_m^{(\varepsilon)}|^{2p} \mid \mathscr{F}_0\} \leqq C(1 + |x|^{2p}) + \frac{2p}{\varepsilon}\sum_{k=1}^{m-1} |E\{I_k^{(\varepsilon)}(x) \mid \mathscr{F}_0\}|\,.$$

We wish to show that

$$\begin{aligned}(3.13)\qquad &|E\{I_k^{(\varepsilon)}(x) \mid \mathscr{F}_0\}| \leqq \varepsilon^2 C(1 + E\{|x_{k-1}^{(\varepsilon)}|^{2p} \mid \mathscr{F}_0\})\,, \\ &\qquad\qquad k = 1, 2, \cdots, m-1\,.\end{aligned}$$

From this estimate and (3.12) the desired result (3.7) follows immediately by the discrete time version of Gronwall's inequality (cf. [2], p. 221). We proceed therefore to prove (3.13).

The following identities are verified by direct computation:

$$
\begin{aligned}
x^{(\varepsilon)}(s) &= x_{k-1}^{(\varepsilon)} + \frac{1}{\varepsilon}\int_{\sigma_{k-1}}^{s} F\left(\sigma, x^{(\varepsilon)}(\sigma), \frac{\sigma}{\varepsilon^2}\right) d\sigma \\
&\equiv x_{k-1}^{(\varepsilon)} + \frac{1}{\varepsilon}\,\xi_{k-1}(s)\,,
\end{aligned}
\tag{3.14}
$$

$$
\begin{aligned}
&|x^{(\varepsilon)}(s)|^{2(p-1)} \\
&= |x_{k-1}^{(\varepsilon)}|^{2(p-1)} + \frac{2(p-1)}{\varepsilon}\int_{\sigma_{k-1}}^{s}\left(F\left(\sigma, x^{(\varepsilon)}(\sigma), \frac{\sigma}{\varepsilon^2}\right), x^{(\varepsilon)}(\sigma)\right)|x^{(\varepsilon)}(\sigma)|^{2(p-2)}\, d\sigma\,,
\end{aligned}
\tag{3.15}
$$

$$
\begin{aligned}
&F\left(s, x^{(\varepsilon)}(s), \frac{s}{\varepsilon^2}\right) \\
&= F\left(s, x_{k-1}^{(\varepsilon)}, \frac{s}{\varepsilon^2}\right) + \frac{1}{\varepsilon}\int_{\sigma_{k-1}}^{s} \frac{\partial F}{\partial x}\left(s, x^{(\varepsilon)}(\sigma), \frac{s}{\varepsilon^2}\right) F\left(\sigma, x^{(\varepsilon)}(\sigma), \frac{\sigma}{\varepsilon^2}\right) d\sigma \\
&\equiv F\left(s, x_{k-1}^{(\varepsilon)}, \frac{s}{\varepsilon^2}\right) + \frac{1}{\varepsilon}\eta_{k-1}(s)\,, \quad \sigma_{k-1} \leqq s \leqq \sigma_{k+1}\,, \quad k = 1, 2, \cdots, m-1
\end{aligned}
\tag{3.16}
$$

In (3.16), $\partial F/\partial x$ denotes the matrix of derivatives of F with respect to x. We insert (3.14)–(3.16) into (3.10) and rearrange. This yields

$$
\begin{aligned}
I_k^{(\varepsilon)}(x) = \int_{\sigma_k}^{\sigma_{k+1}} &\Bigg[\left(F\left(s, x_{k-1}^{(\varepsilon)}, \frac{s}{\varepsilon^2}\right), x_{k-1}^{(\varepsilon)}\right) + \frac{1}{\varepsilon}\,(\eta_{k-1}(s), x_{k-1}^{(\varepsilon)}) \\
&+ \frac{1}{\varepsilon}\left(F\left(s, x_{k-1}^{(\varepsilon)}, \frac{s}{\varepsilon^2}\right), \xi_{k-1}(s)\right) + \frac{1}{\varepsilon^2}\,(\eta_{k-1}(s), \xi_{k-1}(s))\Bigg] \\
&\times \Bigg[|x_{k-1}^{(\varepsilon)}|^{2(p-1)} + \frac{2(p-1)}{\varepsilon} \\
&\times \int_{\sigma_{k-1}}^{s}\left(F\left(\sigma, x^{(\varepsilon)}(\sigma), \frac{\sigma}{\varepsilon^2}\right), x^{(\varepsilon)}(\sigma)\right)|x^{(\varepsilon)}(\sigma)|^{2(p-2)}\, d\sigma\Bigg]\, ds\,.
\end{aligned}
\tag{3.17}
$$

There are 8 terms in (3.17) that must be estimated. We shall estimate 2 typical terms in detail. The others are estimated in the same manner. In the following we use Lemma 1, Lemma 2, (2.2), (2.3) and (2.11):

$$
\begin{aligned}
&\left|E\left\{\int_{\sigma_k}^{\sigma_{k+1}}\left(F\left(s, x_{k-1}^{(\varepsilon)}, \frac{s}{\varepsilon^2}\right), x_{k-1}^{(\varepsilon)}\right) |x_{k-1}^{(\varepsilon)}|^{2(p-1)}\, ds \,\middle|\, \mathscr{F}_0\right\}\right| \\
&\qquad \leqq E\left\{\int_{\sigma_k}^{\sigma_{k+1}} \left|E\left\{\left(F\left(s, x_{k-1}^{(\varepsilon)}, \frac{s}{\varepsilon^2}\right), x_{k-1}^{(\varepsilon)}\right) \,\middle|\, \mathscr{F}_{k-1}\right\}\right| ds\, |x_{k-1}^{(\varepsilon)}|^{2(p-1)} \,\middle|\, \mathscr{F}_0\right\} \qquad (3.18)\\
&\qquad \leqq \varepsilon\rho\left(\frac{1}{\varepsilon}\right) C(1 + E\{|x_{k-1}^{(\varepsilon)}|^{2p} \mid \mathscr{F}_0\}) \\
&\qquad \leqq \varepsilon^2 C(1 + E\{|x_{k-1}^{(\varepsilon)}|^{2p} \mid \mathscr{F}_0\}), \qquad k = 1, 2, \cdots, m-1 .
\end{aligned}
$$

In the last inequality we use (2.2) and the monotonicity of ρ to deduce that $(1/\varepsilon)\rho(1/\varepsilon)$ is bounded in $[0, 1]$.

Let $\mathscr{F}_s = \mathscr{F}_0^{s/\varepsilon^2}$, $\sigma \leqq s \leqq \tau$. We continue by considering the following relations:

$$
\begin{aligned}
&\left|E\left\{\frac{1}{\varepsilon}\int_{\sigma_k}^{\sigma_{k-1}}\int_{\sigma_{k-1}}^{s}\left(\frac{\partial F}{\partial x}\left(s, x^{(\varepsilon)}(\sigma), \frac{s}{\varepsilon^2}\right) F\left(\sigma, x^{(\varepsilon)}(\sigma), \frac{\sigma}{\varepsilon^2}\right), x_{k-1}^{(\varepsilon)}\right) |x_{k-1}^{(\varepsilon)}|^{2(p-1)}\, d\sigma\, ds \,\middle|\, \mathscr{F}_0\right\}\right| \\
&\qquad \leqq E\left\{\frac{1}{\varepsilon}\int_{\sigma_k}^{\sigma_{k+1}}\int_{\sigma_{k-1}}^{s}\left|E\left\{\left(\frac{\partial F}{\partial x}\left(s, x^{(\varepsilon)}(\sigma), \frac{s}{\varepsilon^2}\right)\right.\right.\right.\right. \\
&\qquad\qquad \left.\left.\left.\left. \times F\left(\sigma, x^{(\varepsilon)}(\sigma), \frac{\sigma}{\varepsilon^2}\right), x_{k-1}^{(\varepsilon)}\right) \,\middle|\, \mathscr{F}_0\right\}\right| \cdot |x_{k-1}^{(\varepsilon)}|^{2(p-1)}\, d\sigma\, ds \,\middle|\, \mathscr{F}_0\right\} \qquad (3.19)\\
&\qquad \leqq \frac{1}{\varepsilon}\int_{\sigma_k}^{\sigma_{k+1}}\int_{\sigma_{k-1}}^{s}\rho\left(\frac{s-\sigma}{\varepsilon^2}\right) d\sigma\, ds \cdot C(1 + E\{|x_{k-1}^{(\varepsilon)}|^{2p} \mid \mathscr{F}_0\}) \\
&\qquad \leqq \varepsilon^2 C(1 + E\{|x_{k-1}^{(\varepsilon)}|^{2p} \mid \mathscr{F}_0\} .
\end{aligned}
$$

Here we have used again Lemmas 1 and 2, hypotheses (2.2), (2.3) and (2.11). The analysis of the other terms in (3.17) proceeds in much the same way as (3.18) or (3.19). Thus (3.13) follows and the proof of Lemma 3 is complete for p even. For general p we use Schwarz' inequality and the even p result:

$$
\begin{aligned}
E\{|x^{(\varepsilon)}(\tau, \sigma, x)|^p \mid \mathscr{F}_0^{\sigma/\varepsilon^2}\} &\leqq E^{1/2}\{|x^{(\varepsilon)}(\tau, \sigma, x)|^{2p} \mid \mathscr{F}_0^{\sigma/\varepsilon^2}\} \\
&\leqq C^{1/2}(1 + |x|^{2p})^{1/2} \leqq C(1 + |x|^p) .
\end{aligned}
$$

We note that the decomposition (3.9) can be avoided which yields a more direct proof for the lemma.

Before continuing further with the proof of the theorem we summarize in the following lemma some facts about the solution $u(\sigma, \tau, x)$ of (2.25) which we shall use in the sequel. As we pointed out before, the statements in Lemma 4 are seen to be valid without (2.18), by using Oleinik's theory (cf. [16], [17],

[18]). With (2.18) the results can be derived by using the Itô calculus (cf. [1], [2]).

LEMMA 4. *Let* (2.18)–(2.22) *hold and let* $x^{(0)}(\tau, \sigma, x)$ *be the diffusion Markov process with infinitesimal generator* $\mathscr{L}_\sigma$ *given by* (2.17). *Then the partial differential equation* (2.25) *has a unique classical solution for* $f(x) \in C^{k,\bar{p}}$, $2 \leqq k \leqq 4$, $p \geqq 0$, *and there is a* $\bar{p} \geqq p + 4$ *such that* $u(\sigma, \tau, x) \in C^{k,p}$.

Following the convention for the constants C we shall not be precise with values of $\bar{p}$ also.

Lemmas 5–8 which follow deal with (2.26). Evidently, (2.26) implies that the finite-dimensional distributions of $x^{(\varepsilon)}$ converge to the finite-dimensional distributions of the diffusion Markov process $x^{(0)}$ with infinitesimal generator $\mathscr{L}_\sigma$ given by (2.17). The processes $x^{(\varepsilon)}(\tau, \sigma, x)$ are continuous. Thus in order to show weak convergence of $x^{(\varepsilon)}$ to $x^{(0)}$ it is sufficient to show that the family $x^{(\varepsilon)}$ is weakly compact. After Lemma 8 we shall employ a slightly improved version of (2.26), along with some additional calculations, to show weak compactness.

As in Lemma 3, we fix σ and τ in $[0, T]$ and decompose the interval $[\sigma, \tau]$ into $m = m(\varepsilon)$ sections of length ε. We also employ the abbreviations (3.10) for σ_k, $x_k^{(\varepsilon)}$ and $\mathscr{F}_k$, $k = 0, 1, 2, \cdots, m$. We fix a function $f(x)$ in $C^{4,p}(R^n)$, $p \geqq 0$, and define $I^{(\varepsilon)}(\sigma, \tau, x)$ as follows:

$$I^{(\varepsilon)}(\sigma, \tau, x) = |E\{f(x^{(\varepsilon)}(\tau, \sigma, x)) \mid \mathscr{F}_0^{\sigma/\varepsilon^2}\} - u(\sigma, \tau, x)| . \tag{3.20}$$

Note that $I^{(\varepsilon)}$ is an $\mathscr{F}_0^{\sigma/\varepsilon^2}$ measurable random variable. As in Lemma 3, we shall abbreviate in the sequel $\mathscr{F}_0^{s/\varepsilon^2}$ by $\mathscr{F}_s$ also for variable s, $\sigma \leqq s \leqq \tau$, but we continue to denote $\mathscr{F}_0^{\sigma/\varepsilon^2}$ by $\mathscr{F}_0$ adhering to (3.10):

$$\bar{u}(\sigma_k, \tau, x) = E\{u(\sigma_k, \tau, x^{(\varepsilon)}(\sigma_k, \sigma_{k-1}, x))\} . \tag{3.21}$$

The following relations can be verified easily:

$$\begin{aligned}
&I^{(\varepsilon)}(\sigma, \tau, x) \\
&= \left| \sum_{k=1}^{m} E\{u(\sigma_k, \tau, x_k^{(\varepsilon)}) - u(\sigma_{k-1}, \tau, x_{k-1}^{(\varepsilon)}) \mid \mathscr{F}_0\} \right| \\
&\leqq \sum_{k=1}^{m} \left| E\{u(\sigma_k, \tau, x^{(\varepsilon)}(\sigma_k, \sigma_{k-1}, x_{k-1}^{(\varepsilon)})) \mid \mathscr{F}_0\} - E\{\bar{u}(\sigma_k, \tau, x_{k-1}^{(\varepsilon)}) \mid \mathscr{F}_0\} \right| \\
&\quad + \sum_{k=1}^{m} \left| E\{\bar{u}(\sigma_k, \tau, x_{k-1}^{(\varepsilon)}) \mid \mathscr{F}_0\} - E\{u(\sigma_{k-1}, \tau, x_{k-1}^{(\varepsilon)}) \mid \mathscr{F}_0\} \right| \\
&\leqq \sum_{k=1}^{m} \left| E\{u(\sigma_k, \tau, x^{(\varepsilon)}(\sigma_k, \sigma_{k-1}, x_{k-1}^{(\varepsilon)})) \mid \mathscr{F}_0\} - E\{\bar{u}(\sigma_k, \tau, x_{k-1}^{(\varepsilon)} \mid \mathscr{F}_0\} \right| \\
&\quad + \sum_{k=1}^{m} E\{|E\{\bar{u}(\sigma_k, \tau, x_{k-1}^{(\varepsilon)}) \mid \mathscr{F}_{k-1}\} - E\{u(\sigma_{k-1}, \tau, x_{k-1}^{(\varepsilon)}) \mid \mathscr{F}_{k-1}\}| \mid \mathscr{F}_0\} .
\end{aligned} \tag{3.22}$$

Let

$$
\begin{aligned}
(3.23)\qquad I_{1,k}^{(\varepsilon)}(x) &= |E\{u(\sigma_k, \tau, x^{(\varepsilon)}(\sigma_k, \sigma_{k-1}, x_{k-1}^{(\varepsilon)})) \mid \mathscr{F}_0\} - E\{\bar{u}(\sigma_k, \tau, x_{k-1}^{(\varepsilon)}) \mid \mathscr{F}_0\}| , \\
I_{2,k}^{(\varepsilon)}(x) &= |\bar{u}(\sigma_k, \tau, x) - u(\sigma_{k-1}, \tau, x)| .
\end{aligned}
$$

Note that $I_{1,k}^{(\varepsilon)}(x)$ is an $\mathscr{F}_\sigma$ measurable random variable.

LEMMA 5. *For* $k = 2, \cdots, m$ *we have*

$$
(3.24)\qquad I_{1,k}^{(\varepsilon)}(x) \leqq \varepsilon^2 C(1 + |x|^{\bar{p}}) .
$$

Here C *is independent of* k, x, ε *and* $\omega \in \Omega$ *and* $\bar{p} \geqq p$ *is some integer. For* $k = 1$ *we have* $I_{1,1}^{(\varepsilon)}(x) \leqq C\varepsilon(1 + |x|^{\bar{p}})$.

Proof: First we rewrite $I_{1,k}^{(\varepsilon)}(x)$ using (3.21):

$$
\begin{aligned}
(3.25)\qquad I_{1,k}^{(\varepsilon)}(x) = \Big| &\int u(\sigma_k, \tau, x^{(\varepsilon)}(\sigma_k, \sigma_{k-1}, x^{(\varepsilon)}(\sigma_{k-1}, \sigma, x, \omega), \omega)) P_{\sigma/\varepsilon^2}(d\omega \mid \omega'') \\
&- \iint u(\sigma_k, \tau, x^{(\varepsilon)}(\sigma_k, \sigma_{k-1}, x^{(\varepsilon)}(\sigma_{k-1}, \sigma, x, \omega'), \omega)) P(d\omega) P_{\sigma/\varepsilon^2}(d\omega' \mid \omega'') \Big|
\end{aligned}
$$

Dependence of $I_{1,k}^{(\varepsilon)}(x)$ on ω'' will not be indicated explicitly. Moreover, since $x^{(\varepsilon)}(\sigma_{k-1}, \sigma, x)$ is $\mathscr{F}_{\sigma/\varepsilon^2}^{\sigma_{k-1}/\varepsilon^2}$ measurable and $x^{(\varepsilon)}(\sigma_k, \sigma_{k-1}, x)$ is $\mathscr{F}_{\sigma_{k-1}/\varepsilon^2}^{\sigma_k/\varepsilon^2}$ measurable, we may rewrite (3.25) further as follows:

$$
\begin{aligned}
(3.26)\qquad I_{1,k}^{(\varepsilon)}(x) = \Big| \iint & u(\sigma_k, \tau, x^{(\varepsilon)}(\sigma_k, \sigma_{k-1}, x^{(\varepsilon)}(\sigma_{k-1}, \sigma, x, \omega'), \omega)) \\
& \times [P_{\sigma_{k-1}/\varepsilon^2}(d\omega \mid \omega') - P(d\omega)] P_{\sigma/\varepsilon^2}(d\omega' \mid \omega'') \Big| .
\end{aligned}
$$

We wish to apply Lemma 1 but we note that this would be unproductive in (3.26) since there is no gap between the σ-algebras of $x^{(\varepsilon)}(\sigma_k, \sigma_{k-1}, x)$ and $x^{(\varepsilon)}(\sigma_{k-1}, \sigma, x)$. We must therefore rewrite (3.26) in a more convenient form. As pointed out at the beginning of Section 3, motivation for the identity that follows is provided by equation (2.24) of [8]. Here we shall omit derivations and simply write the desired result which can be verified directly.

To simplify the notation we denote partial derivatives by a comma followed by a subscript, employ the summation convention and set $g^{(k)}(x) = u(\sigma_k, \tau, x)$.

From (3.14) and (3.16) we have

$$
\begin{aligned}
I_{1,k}^{(\varepsilon)}(x) \leqq & \left| \frac{1}{\varepsilon} \int_{\sigma_{k-1}}^{\sigma_k} \int F_i\left(t, x, \frac{t}{\varepsilon^2}, \omega\right) g_{,j}^{(k)}(x^{(\varepsilon)}(\sigma_k, t, x, \omega)) \right. \\
& \left. \times x_{j,i}^{(\varepsilon)}(\sigma_k, t, x, \omega) \cdot [P_{\sigma/\varepsilon^2}(d\omega \mid \omega'') - P(d\omega)]\, dt \right| \\
& + \left| \frac{1}{\varepsilon^2} \int_{\sigma}^{\sigma_{k-1}} \int_{\sigma_{k-1}}^{\sigma_k} \iint F_i\left(s, x^{(\varepsilon)}(s, \sigma, x, \omega'), \frac{s}{\varepsilon^2}, \omega'\right) \right. \\
& \times \left[F_{j,i}\left(t, x^{(\varepsilon)}(s, \sigma, x, \omega'), \frac{t}{\varepsilon^2}, \omega\right) \right. \\
& \times g_{,\gamma}^{(k)}(x^{(\varepsilon)}(\sigma_k, t, x^{(\varepsilon)}(s, \sigma, x, \omega'), \omega)) x_{\gamma,j}^{(\varepsilon)}(\sigma_k, t, x^{(\varepsilon)}(s, \sigma, x, \omega'), \omega) \\
& + F_j\left(t, x^{(\varepsilon)}(s, \sigma, x, \omega'), \frac{t}{\varepsilon^2}, \omega\right) g_{,l\gamma}^{(k)}(x^{(\varepsilon)}(\sigma_k, t, x^{(\varepsilon)}(s, \sigma, x, \omega'), \omega)) \\
& \times x_{\gamma,i}^{(\varepsilon)}(\sigma_k, t, x^{(\varepsilon)}(s, \sigma, x, \omega'), \omega) \cdot x_{l,j}^{(\varepsilon)}(\sigma_k, t, x^{(\varepsilon)}(s, \sigma, x, \omega'), \omega) \\
& + F_j\left(t, x^{(\varepsilon)}(s, \sigma, x, \omega'), \frac{t}{\varepsilon^2}, \omega\right) g_{,\gamma}^{(k)}(x^{(\varepsilon)}(\sigma_k, t, x^{(\varepsilon)}(s, \sigma, x, \omega'), \omega)) \\
& \left. \times x_{\gamma,ji}^{(\varepsilon)}(\sigma_k, t, x^{(\varepsilon)}(s, \sigma, x, \omega'), \omega) \right] \\
& \left. \times [P_{s/\varepsilon^2}(d\omega \mid \omega') - P(d\omega)] P_{\sigma/\varepsilon^2}(d\omega' \mid \omega'')\, dt\, ds \right| .
\end{aligned} \tag{3.27}
$$

Note that the second term on the right side of (3.27) is equal to zero when $k = 1$.

In (3.27) we have effected the desired gap because functions of ω are $\mathscr{F}_{t/\varepsilon^2}^{\sigma_k/\varepsilon^2}$ measurable and functions of ω' are $\mathscr{F}_{\sigma/\varepsilon^2}^{s/\varepsilon^2}$ measurable, where $\sigma \leqq s \leqq \sigma_{k-1} \leqq t \leqq \sigma_k$. We can now apply Lemma 1. For the first term on the right side of (3.27) we use Lemmas 1, 2, 4, (2.2) and (2.4) and obtain

$$
\begin{aligned}
& \left| \frac{1}{\varepsilon} \int_{\sigma_{k-1}}^{\sigma_k} \int F_i\left(t, x, \frac{t}{\varepsilon^2}, \omega\right) g_{,j}^{(k)}(x^{(\varepsilon)}(\sigma_k, t, x, \omega)) x_{j,i}^{(\varepsilon)}(\sigma_k, t, x, \omega) \right. \\
& \qquad \left. \times [P_{\sigma/\varepsilon^2}(d\omega \mid \omega') - P(d\omega)]\, dt \right| \\
& \qquad \leqq \varepsilon^2 C \cdot \frac{1}{\varepsilon} \rho^{1/2}\left(\frac{1}{\varepsilon}\right)(1 + |x|^{\bar{p}}) \leqq \varepsilon^2 C(1 + |x|^{\bar{p}}), \qquad k = 2, \cdots, m .
\end{aligned} \tag{3.28}
$$

The term corresponding to $k = 1$ can be shown to be at most $C\varepsilon(1 + |x|^{\bar{p}})$. Here we have also used the monotonicity of ρ and the fact that $(1/\varepsilon)\rho^{1/2}(1/\varepsilon)$ is uniformly bounded for $\varepsilon \in [0, 1]$. Note that Lemma 3 is not used in (3.28) and $\bar{p} \geqq p$ is some integer.

Let $\tilde{I}_{1,k}^{(\varepsilon)}(x)$ denote the second integral in (3.27). By using Lemmas 1, 2, 4, (2.2), (2.3) and (2.4), we obtain, for some $\bar{p} \geqq p$,

$$\tilde{I}_{1,k}^{(\varepsilon)}(x) \leqq \frac{1}{\varepsilon^2} \int_{\sigma}^{\sigma_{k-1}} \int_{\sigma_{k-1}}^{\sigma_k} \rho\left(\frac{t-s}{\varepsilon^2}\right) dt\, ds\, C \int (1 + |x^{(\varepsilon)}(s, \sigma, x, \omega')|^{\bar{p}}) P_{\sigma/\varepsilon^2}(d\omega' \mid \omega'') , \qquad k \geqq 2 .$$

The t, s integral, with the factor $1/\varepsilon^2$ included, is estimated in (2.28)–(2.30) of [8] and is $O(\varepsilon^2)$. On using that estimate and Lemma 3 we obtain

$$\tilde{I}_{1,k}^{(\varepsilon)}(x) \leqq \varepsilon^2 C(1 + |x|^{\bar{p}}) , \qquad k \geqq 2 . \tag{3.29}$$

This last inequality and (3.28) yield (3.24). The proof of Lemma 5 is complete.

We proceed now with the estimation of $I_{2,k}^{(\varepsilon)}(x)$ in (3.23).

LEMMA 6. *There is a constant C independent of k, x, ε and $\omega \in \Omega$ and a $\bar{p} \geqq p$ such that*

$$I_{2,k}^{(\varepsilon)}(x) \leqq \varepsilon^2 C(1 + |x|^{\bar{p}}) . \tag{3.30}$$

Proof: The proof of (3.30) is a minor modification of the argument beginning with (2.33) in [8]. Here we must account for the time inhomogeneity of the limiting diffusion process. First we rewrite $I_{2,k}^{(\varepsilon)}(x)$ using (3.21):

$$\begin{aligned} I_{2,k}^{(\varepsilon)}(x) &= \left| \int u(\sigma_k, \tau, x^{(\varepsilon)}(\sigma_k, \sigma_{k-1}, x, \omega)) P(d\omega) - u(\sigma_{k-1}, \tau, x) \right| \\ &= |E\{g^{(k)}(x^{(\varepsilon)}(\sigma_k, \sigma_{k-1}, x))\} - u(\sigma_{k-1}, \tau, x)| . \end{aligned} \tag{3.31}$$

We use the abbreviation $g^{(k)}(x) = u(\sigma_k, \tau, x)$. Next we iterate the integrated form of (2.25) once, a legitimate procedure in view of Lemma 4, and obtain

$$u(\sigma_{k-1},\tau,x)=g^{(k)}(x)+\int_{\sigma_{k-1}}^{\sigma_k}\mathscr{L}_\sigma g^{(k)}(x)\,d\sigma+\int_{\sigma_{k-1}}^{\sigma_k}\int_{\sigma}^{\sigma_k}\mathscr{L}_\sigma\mathscr{L}_s u(s,\tau,x)\,ds\,d\sigma\,. \tag{3.32}$$

Thus, by Lemma 4 there is a $\bar{p}\geqq p$ such that

$$\left|u(\sigma_{k-1},\tau,x)-g^{(k)}(x)-\int_{\sigma_{k-1}}^{\sigma_k}\mathscr{L}_\sigma g^{(k)}(x)\,d\sigma\right|\leqq\varepsilon^2C(1+|x|^{\bar{p}})\,. \tag{3.33}$$

We shall show in Lemmas 7 and 8 that there is a $\bar{p}\geqq p$ such that we have, along with (3.33), the estimate

$$\begin{aligned}&\Bigg|E\{g^{(k)}(x^{(\varepsilon)}(\sigma_k,\sigma_{k-1},x))\}-g^{(k)}(x)\\ &\quad-\frac{1}{\varepsilon^2}\int_{\sigma_{k-1}}^{\sigma_k}\int_{\sigma_{k-1}}^{s}\left[E\left\{F_i\left(\sigma,x,\frac{\sigma}{\varepsilon^2}\right)F_j\left(s,x,\frac{s}{\varepsilon^2}\right)\right\}g^{(k)}_{,ij}(x)\right.\\ &\quad\left.+E\left\{F_i\left(\sigma,x,\frac{\sigma}{\varepsilon^2}\right)F_{j,i}\left(s,x,\frac{s}{\varepsilon^2}\right)\right\}g^{(k)}_{,j}(x)\right]\Bigg|\,d\sigma\,ds\leqq\varepsilon^2C(1+|x|^{\bar{p}})\,.\end{aligned} \tag{3.34}$$

Now we use (3.33) and (3.34) in (3.31) to obtain

$$\begin{aligned}I^{(\varepsilon)}_{2,k}(x)\leqq&\left|\left[\frac{1}{\varepsilon^2}\int_{\sigma_{k-1}}^{\sigma_k}\int_{\sigma_{k-1}}^{s}E\left\{F_i\left(\sigma,x,\frac{\sigma}{\varepsilon^2}\right)F_j\left(s,x,\frac{s}{\varepsilon^2}\right)\right\}d\sigma\,ds\right.\right.\\ &\left.\left.-\int_{\sigma_{k-1}}^{\sigma_k}a^{ij}(\sigma,x)\,d\sigma\right]g^{(k)}_{,ij}(x)\right|\\ &+\left|\left[\frac{1}{\varepsilon^2}\int_{\sigma_{k-1}}^{\sigma_k}\int_{\sigma_{k-1}}^{s}E\left\{F_i\left(\sigma,x,\frac{\sigma}{\varepsilon^2}\right)F_{j,i}\left(s,x,\frac{s}{\varepsilon^2}\right)\right\}d\sigma\,ds\right.\right.\\ &\left.\left.-\int_{\sigma_{k-1}}^{\sigma_k}b^{j}(\sigma,x)\,d\sigma\right]g^{(k)}_{,j}(x)\right|+\varepsilon^2C(1+|x|^{\bar{p}})\,.\end{aligned} \tag{3.35}$$

To complete the proof of Lemma 6, assuming (3.34) for the moment, it remains to estimate the first two terms on the right side of (3.35). We shall only estimate the first term since the second is treated in the same way:

$$
\begin{aligned}
&\left|\left[\frac{1}{\varepsilon^2}\int_{\sigma_{k-1}}^{\sigma_k}\int_{\sigma_{k-1}}^{s} E\left\{F_i\left(\sigma, x, \frac{\sigma}{\varepsilon^2}\right)F_j\left(s, x, \frac{s}{\varepsilon^2}\right)\right\} d\sigma\, ds - \int_{\sigma_{k-1}}^{\sigma_k} a^{ij}(\sigma, x)\, d\sigma\right] g^{(k)}_{,ij}(x)\right| \\
&\leqq C(1+|x|^{\bar{p}})\left\{\sup_{i,j}\left|\frac{1}{\varepsilon^2}\int_{\sigma_{k-1}}^{\sigma_k}\int_{\sigma_{k-1}}^{s}\left[E\left\{F_i\left(\sigma, x, \frac{\sigma}{\varepsilon^2}\right)F_j\left(s, x, \frac{s}{\varepsilon^2}\right)\right\}\right.\right.\right. \\
&\qquad \left.\left. - E\left\{F_i\left(\sigma_{k-1}, x, \frac{\sigma}{\varepsilon^2}\right)F_j\left(\sigma_{k-1}, x, \frac{s}{\varepsilon^2}\right)\right\}\right] d\sigma\, ds \\
&\qquad + \sup_{i,j}\left|\int_{\sigma_{k-1}}^{\sigma_k}[a^{ij}(\sigma, x) - a^{ij}(\sigma_{k-1}, x)]\, d\sigma\right| \\
&\qquad \left. + \sup_{i,j}\left|\frac{1}{\varepsilon^2}\int_{\sigma_{k-1}}^{\sigma_k}\int_{\sigma_{k-1}}^{s} E\left\{F_i\left(\sigma_{k-1}, x, \frac{\sigma}{\varepsilon^2}\right)F_j\left(\sigma_{k-1}, x, \frac{s}{\varepsilon^2}\right)\right\} d\sigma\, ds - \varepsilon a^{ij}(\sigma_{k-1}, x)\right|\right\} \qquad (3.36) \\
&\leqq C(1+|x|^{\bar{p}})\left\{\sup_{i,j}\left|\frac{1}{\varepsilon^2}\int_{\sigma_{k-1}}^{\sigma_k}\int_{\sigma_{k-1}}^{s}\left[E\left\{\left(F_i\left(\sigma, x, \frac{\sigma}{\varepsilon^2}\right)\right.\right.\right.\right.\right. \\
&\qquad \left. \left. - F_i\left(\sigma_{k-1}, x, \frac{\sigma}{\varepsilon^2}\right)\right)F_j\left(s, x, \frac{s}{\varepsilon^2}\right)\right\} \\
&\qquad \left.\left. + E\left\{F_i\left(\sigma_{k-1}, x, \frac{\sigma}{\varepsilon^2}\right)\left(F_j\left(s, x, \frac{s}{\varepsilon^2}\right) - F_j\left(\sigma_{k-1}, x, \frac{s}{\varepsilon^2}\right)\right)\right\}\right] d\sigma\, ds\right| + \varepsilon^2 C(1+|x|^2)\right\} \\
&\leqq \varepsilon^2 C(1+|x|^{\bar{p}+2}) + C(1+|x|^{\bar{p}}) \\
&\qquad \times\left[\frac{4}{\varepsilon^2}\int_{\sigma_{k-1}}^{\sigma_k}\int_{\sigma_{k-1}}^{s}\rho\left(\frac{s-\sigma}{\varepsilon^2}\right) d\sigma\, ds \cdot C(1+|x|)\varepsilon C(1+|x|)\right] \\
&\leqq \varepsilon^2 C(1+|x|^{\bar{p}})\,. \qquad\qquad \text{(Recall the convention about } \bar{p}.\text{)}
\end{aligned}
$$

In the next to the last inequality in (3.36), we have used (3.4) with $p = q = \frac{1}{2}$, (2.7) (and earlier (2.23)) and, in the last inequality, (2.2). From (3.36) and (3.35) it follows that the proof of Lemma 6 is complete, assuming the validity of (3.34).

We proceed therefore with the proof of (3.34). As for (3.27), the motivation for what follows comes from Theorem 1 in [8]. It can also be verified by direct computation that the following identity holds:

$$
\begin{aligned}
&E\{g^{(k)}(x^{(\varepsilon)}(\sigma_k, \sigma_{k-1}, x))\} \\
&= g^{(k)}(x) + \frac{1}{\varepsilon}\int_{\sigma_{k-1}}^{\sigma_k} E\left\{F_i\left(s_1, x, \frac{s_1}{\varepsilon^2}\right)\right\} g^{(k)}_{,i}(x)\, ds_1 \\
&\quad + \frac{1}{\varepsilon^2}\int_{\sigma_{k-1}}^{\sigma_k}\int_{s_1}^{\sigma_k} E\left\{F_i\left(s_1, x, \frac{s_1}{\varepsilon^2}\right)\left[F_j\left(s_2, x, \frac{s_2}{\varepsilon^2}\right) g^{(k)}_{,j}(x)\right]_{,i}\right\} ds_1\, ds_2 \\
&\quad + \frac{1}{\varepsilon^3}\int_{\sigma_{k-1}}^{\sigma_k}\int_{s_1}^{\sigma_k}\int_{s_2}^{\sigma_k} E\left\{F_i\left(s_1, x, \frac{s_1}{\varepsilon^2}\right)\left[F_j\left(s_2, x, \frac{s_2}{\varepsilon^2}\right)\right.\right. \\
&\quad \times \left.\left.\left(F_l\left(s_3, x, \frac{s_3}{\varepsilon^2}\right) g^{(k)}_{,l}(x)\right)_{,j}\right]_{,i}\right\} ds_1\, ds_2\, ds_3 \qquad (3.37) \\
&\quad + \frac{1}{\varepsilon^4}\int_{\sigma_{k-1}}^{\sigma_k}\int_{s_1}^{\sigma_k}\int_{s_2}^{\sigma_k} E\left\{F_i\left(s_1, x, \frac{s_1}{\varepsilon^2}\right)\right. \\
&\quad \times \left(F_j\left(s_2, x, \frac{s_2}{\varepsilon^2}\right)\right)\int_{s_2}^{s_3}\left(F_l\left(s_4, x^{(\varepsilon)}(s_4, s_2, x), \frac{s_4}{\varepsilon^2}\right) ds_4\right. \\
&\quad \times \left.\left.\left(F_\gamma\left(s_3, x^{(\varepsilon)}(s_4, s_2, x), \frac{s_3}{\varepsilon^2}\right) g^{(k)}_{,\gamma}(x)\right)_{,l}\right)_{,j}\right]_{,i}\right\} ds_3\, ds_2\, ds_1 \,.
\end{aligned}
$$

The second term on the right side of (3.37) is zero on account of (2.11). After bringing the first and third terms of the right side to the left side and taking absolute values it becomes clear, with an interchange of integration, that (3.34) holds with the help of the following two lemmas. Recall that, by Lemma 4, $g^{(k)}(x)$ is differentiable.

LEMMA 7. *There is an integer $\bar{p} \geqq p$ and a C independent of x, ε and k such that the absolute value of the fourth term on the right side of* (3.37) *is at most* $\varepsilon^2 C(1 + |x|^{\bar{p}})$.

LEMMA 8. *There is an integer $\bar{p} \geqq p$ and a C independent of x, ε and k such that the absolute value of the last term on the right side of* (3.37) *is at most* $\varepsilon^2 C(1 + |x|^{\bar{p}})$.

The proofs of these lemmas are almost identical with the proofs of Lemmas 2 and 3 in [8]. Therefore we shall not repeat them here. We note only that Lemma 3 is not required for either Lemma 7 or Lemma 8 since the estimates are local, i.e., within an interval $[\sigma_{k-1}, \sigma_k]$ of length ε.

Let us return finally to (3.22). From Lemma 5 and Lemma 6 it follows that there is a constant C independent of ε, x and ω and an integer $\bar{p} \geqq p$ such that

$$I^{(\varepsilon)}(\sigma, \tau, x) \leqq m\varepsilon^2 C(1 + |x|^{\bar{p}}) + \varepsilon^2 C \sum_{k=2}^{m} [1 + E\{|x_{k-1}^{(\varepsilon)}|^{\bar{p}} \mid \mathscr{F}_0\}] + \varepsilon C(1 + |x|^{\bar{p}}). \tag{3.38}$$

On using Lemma 3 in (3.38) and the fact that $m\varepsilon = \tau - \sigma$, we obtain

$$I^{(\varepsilon)}(\sigma, \tau, x) \leqq \varepsilon TC(1 + |x|^{\bar{p}}) . \tag{3.39}$$

This completes the proof of (2.26). If we have $q = 0$ in (2.5), (2.6) and (2.21), (2.22), then it is easy to check that $\bar{p} = p + 4$.

It remains to prove that the continuous processes $\{x^{(\varepsilon)}(\tau, \sigma, x),\ \varepsilon \in (0, 1],\ 0 \leqq \sigma \leqq \tau \leqq T,\ x \in R^n\}$, whose finite-dimensional distributions have been shown to converge to those of the diffusion Markov process $x^{(0)}$, are weakly compact in the space of continuous trajectories on R^n. We shall show that for each compact $U \subset R^n$ the corresponding $x^{(\varepsilon)}(\tau, \sigma, x)$, $x \in U$, $\varepsilon \in (0, 1]$, $0 \leqq \sigma \leqq \tau \leqq T$, are weakly compact on $C([\sigma, T], R^n)$, σ fixed in $[0, T]$. We take $\sigma = 0$ without loss in generality.

We wish to show that for any $0 \leqq \sigma \leqq \tau \leqq T$ there is a constant C indedependent of $x \in U$, $\varepsilon \in [0, 1]$, σ and $\tau \in [0, T]$ such that, for some $\alpha > 0$,

$$E\{|x^{(\varepsilon)}(\tau, 0, x) - x^{(\varepsilon)}(\sigma, 0, x)|^4\} \leqq C(\tau - \sigma)^{1+\alpha} . \tag{3.40}$$

By a well known theorem (see [22], p. 450) the estimate (3.40) yields weak compactness and hence, by the above remarks, weak convergence. We have

$$\begin{aligned} &E\{|x^{(\varepsilon)}(\tau, 0, x) - x^{(\varepsilon)}(\sigma, 0, x)|^4\} \\ &\qquad = E\{|x^{(\varepsilon)}(\tau, \sigma, x^{(\varepsilon)}(\sigma, 0, x)) - x^{(\varepsilon)}(\sigma, 0, x)|^4\} \\ &\qquad = E\{E\{|x^{(\varepsilon)}(\tau, \sigma, x^{(\varepsilon)}(\sigma, 0, x)) - x^{(\varepsilon)}(\sigma, 0, x)|^4 \mid \mathscr{F}_0^{\sigma/\varepsilon^2}\}\} . \end{aligned} \tag{3.41}$$

Define $I(y)$ by

$$I(y) = E\{|x^{(\varepsilon)}(\tau, \sigma, y) - y|^4 \mid \mathscr{F}_0^{\sigma/\varepsilon^2}) . \tag{3.42}$$

For given ε, τ and σ, we consider the cases $(\tau - \sigma)^{5/7} \geqq \varepsilon$ and $(\tau - \sigma)^{5/7} < \varepsilon$ separately. Suppose first that $(\tau - \sigma)^{5/7} \geqq \varepsilon$. We shall obtain the desired estimate (3.40) with $\alpha = \frac{1}{7}$ by using a refined version of (3.38).

Let $f(x) = |x - y|^4$ and denote the solution of (2.25), with y regarded as a parameter, by $u(\sigma, \tau, x;\ y)$. We have the following inequalities:

$$\begin{aligned} I(y) &\leqq |E\{|x^{(\varepsilon)}(\tau, \sigma, y) - y|^4 \mid \mathscr{F}_0^{\sigma/\varepsilon^2}\} - E\{|x^{(0)}(\tau, \sigma, y) - y|^4\}| \\ &\quad + E\{|x^{(0)}(\tau, \sigma, y) - y|^4\} \\ &\leqq |E\{|x^{(\varepsilon)}(\tau, \sigma, y) - y|^4 \mid \mathscr{F}_0^{\sigma/\varepsilon^2}\} - u(\sigma, \tau, y; y)| \\ &\quad + C(\tau - \sigma)^2(1 + |y|^{\bar{p}}) . \end{aligned} \tag{3.43}$$

Here we have used the well known estimate

$$E\{|x^{(0)}(\tau, \sigma, y) - y|^4\} \leqq C(\tau - \sigma)^2 (1 + |y|^{\bar{p}})$$

which can be obtained by Itô's calculus (cf. [2]) or otherwise.

We wish to use (3.38) to estimate the first term on the right side of (3.43), i.e., to use (3.38) with $f(x) = |x - y|^4$, y regarded as a parameter, and then set $x = y$. This is possible except that the term $\varepsilon C(1 + |y|^{\bar{p}})$ on the right side of (3.38) is too crude and must be refined. The remaining terms in (3.38) yield the estimate $\leqq C(\tau - \sigma)^{1+5/7}(1 + |y|^{\bar{p}})$ (recall that $\varepsilon \leqq (\tau - \sigma)^{5/7}$) which is adequate. By re-examining $I_{1,1}^{(\varepsilon)}$ of (3.28), which yielded the crude term in (3.38), we shall show that, if $I^{(\varepsilon)}(\sigma, \tau, x;\ y)$ denotes the first term on the right side of (3.43) (which is the same as (3.20) with $f(x) = |x - y|^4$), then

$$\begin{aligned} I^{(\varepsilon)}(\sigma, \tau, y; y) &\leqq C(\tau - \sigma)^{1+5/7}(1 + |y|^{\bar{p}}) + C(\tau - \sigma)^{1+1/7}(1 + |y|^{\bar{p}}) \\ &\leqq C(1 + |y|^{\bar{p}})(\tau - \sigma)^{1+1/7} . \end{aligned} \tag{3.44}$$

Inserting (3.44) into (3.41), using Lemma 3 and remembering that $x \in U$, a compact subset of R^n, the estimate (3.40) follows with $\alpha = \frac{1}{7}$ and under the hypothesis $\varepsilon \leqq (\tau - \sigma)^{5/7}$.

We proceed now with the proof of (3.44). The case $(\tau - \sigma)^{5/7} < \varepsilon$ will be examined later. Let $I_{1,1}^{(\varepsilon)}(x; y)$ denote (3.26) with $k = 1$ and with $f(x) = |x - y|^4$. It follows that

$$I_{1,1}^{(\varepsilon)}(y; y) = \left|\int u(\sigma + \varepsilon, \tau, x^{(\varepsilon)}(\sigma + \varepsilon, \sigma, y, \omega); y)[P_{\sigma/\varepsilon^2}(d\omega \mid \omega') - P(d\omega)]\right|. \tag{3.45}$$

We have assumed here that $\tau - \sigma \geqq \varepsilon$. The intermediate case $\varepsilon^{7/5} \leqq \tau - \sigma \leqq \varepsilon$ will not be treated explicitly since similar considerations apply. Let us denote by $E^{(0)}$ expectation relative to the measure of the limiting Markov process $x^{(0)}$.

Then we have the following relations:

$$
\begin{aligned}
& u(\sigma+\varepsilon, \tau, x^{(\varepsilon)}(\sigma+\varepsilon, \sigma, y, \omega); y) \\
&\quad = E^{(0)}\{|x^{(0)}(\tau, \sigma+\varepsilon, x^{(\varepsilon)}(\sigma+\varepsilon, \sigma, y, \omega)) - y|^4\} \\
&\quad \leqq 8E^{(0)}\{|x^{(0)}(\tau, \sigma+\varepsilon, x^{(\varepsilon)}(\sigma+\varepsilon, \sigma, y, \omega)) - x^{(\varepsilon)}(\sigma+\varepsilon, \sigma, y, \omega)|^4\} \\
&\qquad + 8\,|x^{(\varepsilon)}(\sigma+\varepsilon, \sigma, y, \omega) - y|^4 \\
&\quad \leqq 8C(\tau-\sigma-\varepsilon)^2(1+|x^{(\varepsilon)}(\sigma+\varepsilon, \sigma, y, \omega)|^{\bar{p}}) \\
&\qquad + 8\,|x^{(\varepsilon)}(\sigma+\varepsilon, \sigma, y, \omega) - y|^4 \\
&\quad \leqq C(\tau-\sigma)^2(1+|y|^{\bar{p}}) + 8\,|x^{(\varepsilon)}(\sigma+\varepsilon, \sigma, y, \omega) - y|^4 .
\end{aligned} \tag{3.46}
$$

Here we have used the same property of $x^{(0)}$ as employed in (3.43). From (3.46) and (3.45) we obtain

$$
\begin{aligned}
I_{1,1}^{(\varepsilon)}(y; y) \leqq{} & C(\tau-\sigma)^2(1+|y|^{\bar{p}}) \\
& + C\left|\int |x^{(\varepsilon)}(\sigma+\varepsilon, \sigma, y, \omega) - y|^4\,[P_{\sigma/\varepsilon^2}(d\omega \mid \omega') - P(d\omega)]\right|,
\end{aligned} \tag{3.47}
$$

and proceed now with the estimation of the last term in this inequality.

Note first that

$$
\begin{aligned}
& |x^{(\varepsilon)}(\sigma+\varepsilon, \sigma, y) - y|^4 \\
&\quad \leqq 8\,|x^{(\varepsilon)}(\sigma+\varepsilon, \sigma+\varepsilon^{7/5}, x^{(\varepsilon)}(\sigma+\varepsilon^{7/5}, \sigma, y)) - x^{(\varepsilon)}(\sigma+\varepsilon^{7/5}, \sigma, y)|^4 \\
&\qquad + 8\,|x^{(\varepsilon)}(\sigma+\varepsilon^{7/5}, \sigma, y) - y|^4 .
\end{aligned} \tag{3.48}
$$

In order to estimate

$$
\int |x^{(\varepsilon)}(\sigma+\varepsilon^{7/5}, \sigma, y, \omega) - y|^4\,[P_{\sigma/\varepsilon^2}(d\omega \mid \omega') - P(d\omega)] ,
$$

we use (3.37) with $g^{(k)}(x) = |x-y|^4$, y regarded as a parameter, and with $\sigma_{k-1} = \sigma$, σ_k replaced by $\sigma + \varepsilon^{7/5}$. Inserting this $g^{(k)}(x)$ into (3.37) and setting $x = y$, all terms on the right side become zero except the last one. Because of the conditioning relative to $\mathscr{F}_0^{\sigma/\varepsilon^2}$, we cannot use Lemma 8 to estimate the surviving last term but we have the elementary estimate

$$
\begin{aligned}
& \left|\int |x^{(\varepsilon)}(\sigma+\varepsilon^{7/5}, \sigma, y, \omega) - y|^4\,[P_{\sigma/\varepsilon^2}(d\omega \mid \omega') - P(d\omega)]\right| \\
&\quad \leqq C\left(\frac{\varepsilon^{7/5}}{\varepsilon}\right)^4 (1+|y|^{\bar{p}}) = C\varepsilon^{8/5}(1+|y|^{\bar{p}}) \\
&\quad \leqq C(\tau-\sigma)^{1+1/7}(1+|y|^{\bar{p}}) ,
\end{aligned} \tag{3.49}
$$

which is what we sought. Here we have used the hypothesis that $\varepsilon \leqq (\tau-\sigma)^{5/7}$.

We estimate next the integral of the first term on the right side of (3.48). We rewrite this term in more convenient form by the usual iteration procedures which yielded (3.27) and (3.37) in the following way (we use the notation of (3.27) and (3.37)):

$$\begin{aligned}
&|x^{(\varepsilon)}(\sigma+\varepsilon, \sigma+\varepsilon^{7/5}, x^{(\varepsilon)}(\sigma+\varepsilon^{7/5}, \sigma, y)) - x^{(\varepsilon)}(\sigma+\varepsilon^{7/5}, \sigma, y)|^4 \\
&\quad = \frac{1}{\varepsilon}\int_{\sigma+\varepsilon^{7/5}}^{\sigma+\varepsilon} F_i\left(s, x^{(\varepsilon)}(\sigma+\varepsilon^{7/5}, \sigma, y), \frac{s}{\varepsilon^2}\right) |x^{(\varepsilon)}(\sigma+\varepsilon, s, x^{(\varepsilon)}(\sigma+\varepsilon^{7/5}, \sigma, y)) \\
&\quad\quad - x^{(\varepsilon)}(\sigma+\varepsilon^{7/5}, \sigma, y)|^4_{,j}\, x^{(\varepsilon)}_{j,1}(\sigma+\varepsilon, s, x^{(\varepsilon)}(\sigma+\varepsilon^{7/5}, \sigma, y))\, ds \\
&\quad = \frac{1}{\varepsilon}\int_{\sigma+\varepsilon^{7/5}}^{\sigma+\varepsilon} F_i\left(s, y, \frac{s}{\varepsilon^2}\right) |x^{(\varepsilon)}(\sigma+\varepsilon, s, x) - y|^4_{,j}\, x^{(\varepsilon)}_{j,i}(\sigma+\varepsilon, s, y)\, ds \\
&\quad\quad + \frac{1}{\varepsilon^2}\int_{\sigma}^{\sigma+\varepsilon^{7/5}} \int_{\sigma+\varepsilon^{7/5}}^{\sigma+\varepsilon} F_l\left(t, x^{(\varepsilon)}(t, \sigma, y), \frac{t}{\varepsilon^2}\right) \\
&\quad\quad \times \Bigg[F_i\left(s, x^{(\varepsilon)}(t, \sigma, y), \frac{s}{\varepsilon^2}\right) |x^{(\varepsilon)}(\sigma+\varepsilon, s, x^{(\varepsilon)}(t, \sigma, y)) - x^{(\varepsilon)}(t, \sigma, y)|^4_{,j} \\
&\quad\quad\quad \times x^{(\varepsilon)}_{j,i}(\sigma+\varepsilon, s, x^{(\varepsilon)}(t, \sigma, y))\Bigg]_{,l}\, ds\, dt\,.
\end{aligned} \tag{3.50}$$

We integrate now both sides of (3.50) with respect to $P_{\sigma/\varepsilon^2}(d\omega \mid \omega') - P(d\omega)$. In a manner entirely analogous to the estimate (3.28) $(k \geqq 2)$ we find that the single integral in (3.50) is less than or equal to

$$\rho\varepsilon^{1/2}(\varepsilon^{-3/5})\, C(1+|y|^{\bar{p}}) \leqq \varepsilon^{8/5}\, C(1+|y|^{\bar{p}}) \leqq (\tau-\sigma)^{1+1/7}\, C(1+|y|^{\bar{p}}).$$

The double integral in (3.50) is estimated like $\tilde{I}_{1,k}$, $k \geqq 2$, in (3.29), and we find that the integral of this term is less than or equal to

$$\varepsilon^2 C(1+|y|^{\bar{p}}) \leqq C(\tau-\sigma)^{1+1/7}(1+|y|^{\bar{p}})\,.$$

Thus (3.44) has been shown to hold when $\varepsilon \leqq (\tau-\sigma)^{5/7}$. As we indicated above this implies that (3.40) holds with $\alpha = \frac{1}{7}$ when $\varepsilon \leqq (\tau-\sigma)^{5/7}$.

It remains to show that (3.40) holds, also with $\alpha = \frac{1}{7}$, when $(\tau-\sigma)^{5/7} < \varepsilon$. This follows easily by estimating $I(y)$ of (3.42) directly using (3.37). Thus, we set $g^{(k)}(x) = |x-y|^4$, $\sigma_{k-1} = \sigma$, $\sigma_k = \tau$, insert the conditioning $\mathscr{F}_0^{\sigma/\varepsilon^2}$ and note that with $x = y$ all terms except the last become zero on the right side of (3.37). This last term is less than or equal to

$$\left(\frac{\tau-\sigma}{\varepsilon}\right)^4 C(1+|y|^{\bar{p}}) \leqq (\tau-\sigma)^{1+1/7} C(1+|y|^{\bar{p}})\,,$$

which follows easily (we use the hypothesis $(\tau-\sigma)^{5/7} < \varepsilon$). From this estimate for $I(y)$, we deduce (3.40) as before. The proof of the theorem is complete.

Bibliography

[1] McKean, H. P. Jr., *Stochastic Integrals*, Academic Press, New York, 1969.
[2] Gikhman, I. I., and Skorokhod, A. V., *Stochastic Differential Equations*, Springer-Verlag, New York, 1972.
[3] Fleming, W. H., *Stochastically perturbed dynamical systems*, Rocky Mountain Math. Journal, to appear.
[4] Khasminskii, R. Z., *Stability of Systems of Differential Equations Under Random Perturbations of Their Parameters*, Nauka, Moscow, 1969. (In Russian.)
[5] Stratonovich, R. L., *Topics in the Theory of Random Noise*, Vols. 1, 2, Gordon and Breach, New York, 1963.
[6] Stratonovich, R. L., *Conditional Markov Processes and Their Application to the Theory of Optimal Control*, American Elsevier, New York, 1967.
[7] Khasminskii, R. Z., *A limit theorem for solutions of differential equations with a random right-hand side*, Theory Prob. Applications, Vol. 11, 1966, pp. 390–406.
[8] Papanicolaou, G. C., and Varadhan, S. R. S., *A limit theorem with strong mixing in Banach space and two applications to stochastic differential equations*, Comm. Pure Appl. Math., Vol. 26, 1973, pp. 497–524.
[9] Kohler, W., and Papanicolaou, G. C., *Power statistics for wave propagation in one dimension and comparison with radiative transport theory*, J. Math. Phys. Vol. 14, 1973, pp. 1733–1745.
[10] Kohler, W., and Papanicolaou, G. C., *Power statistics for wave propagation in one dimension and comparison with radiative transport theory, II*, J. Math. Phys. Vol. 15, 1974, to appear.
[11] Burridge, R., and Papanicolaou, G. C., *The geometry of coupled mode propagation in one-dimensional random media*, Comm. Pure Appl. Math., Vol. 25, 1972, pp. 715–757.
[12] Papanicolaou, G. C., *A kinetic theory for power transfer in stochastic systems*, J. Math. Phys., Vol. 13, 1972, pp. 1912–1918.
[13] Morrison, J. A., and McKenna, J., *Analysis of some stochastic ordinary differential equations*, SIAM-AMS Proc., Vol. 6, Amer. Math. Soc., Providence, R. I., 1973, pp. 97–162.
[14] Loéve, M., *Probability Theory*, Van Nostrand, Princeton, 1963.
[15] Khinchin, A., *Asymptotische Gesetze der Wahrscheinlichkeitsrechnung*, Springer, Berlin, 1933.
[16] Oleinik, O. A., *Alcuni risultati sulle equazioni lineari e quasi lineari ellitticoparaboliche a derivate parziali del segondo ordine*, Rendiconti Accad. Naz., Vol. 40, 1966, pp. 775–784.
[17] Oleinik, O. A., and Radkevich, E. V., *Second Order Equations with Non-negative Characteristic Form*, Amer. Math. Society, Providence, R. I., 1973.
[18] Stroock, R. S., and Varadhan, S. R. S., *On degenerate elliptic-parabolic operators of second order and their associated diffusions*, Comm. Pure Appl. Math., Vol. 25, 1972, pp. 651–714.
[19] Papanciolaou, G., and Keller, J. B., *Stochastic differential equations with applications to random harmonic oscillators and wave propagation in random media*, SIAM J. Appl. Math., Vol. 20, 1971, pp. 282–305.
[20] Bogolyubov, N. N., and Mitropolikii, Yu. A., *Asymptotic Methods in the Theory of Nonlinear Oscillations*, Gordon and Breach, New York, 1961.
[21] Hersh, R., *Random evolutions: a survey of results and problems*, Rocky Mountain Math. Journal, to appear.
[22] Gikhman, I. I., and Skorokhod, A.V., *Introduction to the Theory of Random Processes*, Saunders, Philadelphia, 1969.
[23] Doob, J. L., *Stochastic Processes*, J. Wiley, New York, 1953.
[24] McShane, E. J., *Stochastic differential equations and models of random processes*, Sixth Berkeley Symposium, Vol. III, pp. 263–294, University of California Press, Berkeley, California, 1972.

Received March, 1974.