YOKOHAMA MATHEMATICAL
JOURNAL VOL. 38, 1990

AVERAGING AND WEAK CONVERGENCE METHODS FOR STOCHASTIC DIFFERENTIAL EQUATIONS OF THE MCKEAN TYPE

By

K. NARITA*

(Received November 20, 1989)

Abstract. By the averaging method the weak convergence of a parameterized sequence of processes to a limit process is considered for a multi-dimensional SDE of the McKean type having the drift and diffusion coefficients with a polynomial growth condition in the phase variable. A two-dimensional SDE with mean-field containing a small parameter $\varepsilon>0$ is taken as an application, which is a random perturbation of a dynamical system having an equilibrium point $(0, 0)$ of the plane as a center. A limit process on time scales of order $1/\varepsilon$ is derived and identified for such an equation under the assumption on the existence of a suitable Lyapunov function.

0. Introduction. Hasminskii [2] shows the averaging principle for stochastic differential equations corresponding to a class of linear parabolic equations. Papanicolaou, Stroock and Varadhan [6] treat the same limit problem in a general situation to obtain results on convergence of a sequence of diffusions by martingale and perturbed test function methods. Presenting some powerful new results Kushner [3] develops known techniques so that they are of greater direct applicability in the fields of control and communication. But their results do not cover our examples in oscillations with mean-field. So we consider the averaging method for stochastic differential equation corresponding to a class of nonlinear parabolic equations. For this purpose we adopt a Lyapunov function method.

In § 1 we give a theorem on the existence and uniqueness of the solution of the stochastic differential equation of the McKean type with unbounded coefficients. Being inspired by the weak convergence method in Kushner [3, pp. 34–55], in §§ 2 and 3 we give some convergence theorems on identification

* This research was supported in part by Grant-in-Aid for General Scientific Research (No. 1540203), The Ministry of Education, Science and Culture.

Key words. SDE of the McKean type, averaging method, weak convergence, martingale problem, Lyapunov function, Liénard oscillator, quasiharmonic oscillator.

AMS(MOS) subject classifications. 60H10, 60H20, 34F05.

Abbreviated title: AVERAGING AND WEAK CONVERGENCE METHODS.

of a limit process for a parameterized sequence of processes so that they are applicable for random oscillations with mean-field. In § 4 we obtain a limit process on the time scale $1/\varepsilon$ for a solution of a two-dimensional stochastic differential equation with mean-field containing a small parameter $\varepsilon>0$. In § 5 we give the examples of the Liénard oscillator and the quasiharmonic oscillator with fluctuations strengthened by a random noise depending on the phase variable.

We shall use the following notation:

R^d is the Euclidean d-space.

$\langle x, y\rangle$ is the inner product of $x\in R^d$ and $y\in R^d$.

$|x|$ is the Euclidean norm of $x\in R^d$.

σ^* means the transposed matrix of σ.

$\mathrm{tr}\,\sigma$ is the trace of the $d\times d$-matrix σ.

$|\sigma|$ is the norm of the $d\times d$-matrix $\sigma=(\sigma_{ij})$;

$$|\sigma|^2=\sum_{i,j=1}^{d}\sigma_{ij}^2=\mathrm{tr}(\sigma\sigma^*).$$

$C^{1,2}([0, \infty)\times R^d)$ is the space of functions $f: [0, \infty)\times R^d\to R^1$ that are once continuously differentiable with respect to the time variable and twice with respect to the space variable. $C^2(R^d)$ is the space of functions $f: R^d\to R^1$ that are twice continuously differentiable.

$C_0^\infty(R^d)$ is the space of functions $f: R^d\to R^1$ that are infinitely differentiable and have compact support in R^d. $C(I; R^d)$ for $I\subseteq[0, \infty)$ is the space of R^d-valued functions on I into R^d.

$\mathcal{P}(R^d)$ is the space of probability measures on R^d.

$\langle\Psi, \mu\rangle$ is the integral $\int_{R^d}\Psi(x)\mu(dx)$ for a scalar function Ψ on R^d and $\mu\in\mathcal{P}(R^d)$.

$\mathcal{M}(\Psi)=\{\mu\in\mathcal{P}(R^d); \langle\Psi, \mu\rangle<\infty\}$ for a scalar function $\Psi\geqq 0$ on R^d.

$\mathcal{L}(Z; P)$ is the probability distribution of the random variable Z under the probability measure P in the underlying probability space.

For $\mu\in\mathcal{P}(R^d)$ and $\nu\in\mathcal{P}(R^d)$, $\|\mu-\nu\|$ is defined by

$$\|\mu-\nu\|=\Big(\inf_{Q\in\mathcal{P}_{\mu\nu}}\int_{R^d\times R^d}|x-y|^2Q(dx, dy)\Big)^{1/2},$$

where $\mathcal{P}_{\mu\nu}$ is the space of probability measures on $R^d\times R^d$ such that for any Borel set A in R^d, $Q(A\times R^d)=\mu(A)$ and $Q(R^d\times A)=\nu(A)$. Namely, $\|\mu-\nu\|=(\inf E[|X-Y|^2])^{1/2}$, where the infimum is taken over the set of random variables X and Y having the probability distributions μ and ν, respectively, and $E[\ ]$ denotes the mathematical expectation.

1. SDE of the McKean type. Let $(\Omega, \boldsymbol{F}, P)$ be a probability space with an increasing family $\{\boldsymbol{F}_t; t \geqq 0\}$ of sub-σ-algebras of $\boldsymbol{F}$ and let $W(t)=(W_i(t))_{i=1,\cdots,d}$ be a d-dimensional Brownian motion process adapted to $\boldsymbol{F}_t$. Let ϕ be a d-dimensional random vector independent of $W(t)$. Then we consider the following d-dimensional stochastic differential equation of the *McKean type*;

$$(1.1) \qquad dx(t)=b[t, X(t), u(t)]dt+\sigma[t, X((t), u(t)]dW(t),$$

$$X(0)=\phi,$$

satisfying $u(t)=\mathcal{L}(X(t); P)$ with $u(0)=\mathcal{L}(\phi; P)$. Here $b[t, x, \mu]=(b_i[t, x, \mu])_{i=1,\cdots,d}$ is a d-vector function and $\sigma[t, x, \mu]=(\sigma_{ij}[t, x, \mu])_{i,j=1,\cdots,d}$ is a $d\times d$-matrix function, that are defined on $[0, \infty)\times R^d\times\mathcal{P}(R^d)$.

We shall need the following definition and assumptions.

For $\mu\in\mathcal{P}(R^d)$, $t\geqq 0$, $x=(x_1, \cdots, x_d)\in R^d$ and a scalar function $\Psi\in C^{1,2}([0, \infty)\times R^d)$, define the differential generator $L_t(\mu)$ by

$$L_t(\mu)\Psi(t, x)=\left[\frac{\partial\Psi}{\partial t}+\langle b, \operatorname{grad}_x\Psi\rangle+\frac{1}{2}\operatorname{tr}(a\Psi_{xx})\right](t, x, \mu),$$

where $a[t, x, \mu]=\sigma[t, x, \mu]\sigma^*[t, x, \mu]$, and

$$\operatorname{grad}_x\Psi=\left(\frac{\partial\Psi}{\partial x_i}\right)_{i=1,\cdots,d}, \quad \Psi_{xx}=\left(\frac{\partial^2\Psi}{\partial x_i\partial x_j}\right)_{i,j=1,\cdots,d}.$$

Assumption 1.0. The d-vector $b[t, x, \mu]$ and the $d\times d$-matrix $\sigma[t, x, \mu]$ satisfy the following conditions:

(i) There exist a constant $K>0$ and a constant $K_M>0$ depending on M such that

$$|b[t, x, \mu]-b[t, y, \nu]|+|\sigma[t, x, \mu]-\sigma[t, y, \nu]|\leqq K_M|x-y|+K\|\mu-\nu\|$$

for all $t\geqq 0$, $|x|\leqq M$, $|y|\leqq M$, $\mu\in\mathcal{P}(R^d)$ and $\nu\in\mathcal{P}(R^d)$.

(ii) There exist a constant $c>0$ and an integer $p\geqq 0$ such that

$$|b[t, x, \mu]|+|\sigma[t, x, \mu]|\leqq c(1+|x|^p+\langle\kappa, \mu\rangle)$$

for all $t\geqq 0$, $x\in R^d$ and $\mu\in\mathcal{M}(\kappa)$, where $\kappa(x)=|x|^p$.

Assumption 1.1. There exists a nonnegative function Ψ in $C^{1,2}([0, \infty)\times R^d)$ satisfying the following conditions:

$$\text{(i)} \qquad L_t(\mu)\Psi(t, x)+\frac{1}{2}|\sigma^*[t, x, \mu]\operatorname{grad}_x\Psi(t, x)|^2$$

$$\leqq c_1+c_2\Psi(t, x)+c_3\beta(\langle\Psi(t, \cdot), \mu\rangle)$$

for all $t \geqq 0$, $x \in R^d$ and $\mu \in \mathcal{M}(\Psi(t, \cdot))$ with constants $c_1 \geqq 0$, $c_2 \geqq 0$ and $c_3 > 0$, where $\mathcal{M}(\Psi(t, \cdot)) = \{\mu \in \mathcal{P}(R^d); \langle \Psi(t, \cdot), \mu \rangle < \infty\}$ and $\beta : [0, \infty) \to [0, \infty)$ is non-decreasing and continuous function such that $\int_0^\infty \frac{dr}{1+r+\beta(r)} = \infty$.

(ii) $|x|^2 \leqq l\Psi(t, x)$ for all $t \geqq 0$, $x \in R^d$ with a constant $l > 0$.

Theorem 1.1. *Suppose that Assumption 1.0 and Assumption 1.1 hold. Let ϕ be any d-dimensional random vector indpendent of $W(t)$, such that $E[\Psi(0, \phi)^{2q}] < \infty$ with $q = \max\{2p, 2\}$, where p is as in Assumption 1.0 and Ψ is as in Assumption 1.1. Then there exists a pathwise unique solution $X(t)$ of (1.1) with the initial state $X(0) = \phi$. Define $U(t) = E[\Psi(t, X(t))]$. Then*

$$E[U(t)] \leqq J(t), \tag{1.2}$$

where $J(t) = f^{-1}(f(r_0) + \hat{c}t)$, $r_0 = E[\Psi(0, \phi)] + \Psi(0, 0)$, $\hat{c} = \max\{c_1, c_2, c_3\}$ and f^{-1} is the inverse function of $f(s) = \int_0^s \frac{dr}{1+r+\beta(r)}$.

Moreover, suppose that $E[\Psi(0, \phi)^m] < \infty$ for an integer $m \geqq 2$. Then

$$E[(1+U(t))^m] \leqq E[(1+\Psi(0, \phi)+\Psi(0, 0))^m] \exp\left\{\int_0^t I_m(s)ds\right\}, \tag{1.3}$$

where $I_m(t) = m(m-1)\{c_1 + c_2 + c_3\beta(J(t))\}$.

Proof. We construct M-sequence of truncated processes. Define $\alpha_M(x) = 1$ if $0 \leqq |x| \leqq M$; $\alpha_M(x) = 2 - |x|/M$ if $M < |x| \leqq 2M$; $\alpha_M(x) = 0$ if $|x| > 2M$. Put

$$b_M[t, x, \mu] = \alpha_M(x) b[t, x, \mu], \qquad \sigma_M[t, x, \mu] = \alpha_M(x) \sigma[t, x, \mu].$$

Further define

$$\phi_M = \phi \ \text{ if } \ |\phi| \leqq M; \quad \phi_M = 0 \ \text{ if } \ |\phi| > M.$$

Then there exists a pathwise unique solution $X_M(t)$ of (1.1) with $b[t, x, \mu]$ and $\sigma[t, x, \mu]$ replaced by $b_M[t, x, \mu]$ and $\sigma_M[t, x, \mu]$, respectively, having the initial state $X_M(0) = \phi_M$ such that

$$u_M(t) = \mathcal{L}(X_M(t); P) \quad \text{with} \quad u_M(0) = \mathcal{L}(\phi_M; P).$$

Set $e_M = \inf\{t;\ |X_M(t)| \geqq M\}$. For each $t \geqq 0$, set $t_M = t \wedge e_M$, where $a \wedge b$ is the smaller of the numbers $a, b \in R^1$. For the function Ψ given by Assumption 1.1, define $U_M(t) = \Psi(t_M, X_M(t_M))$. Then Ito's formula concerning stochastic differentials yields

$$U_M(t) = \Psi(0, \phi_M) + \int_0^{t_M} A_M(s)ds + W_M(t_M), \tag{1.4}$$

where $A_M(t)$ is a random function satisfying

$$A_M(s)=L_s(u_M(s))\Psi(s,\ X_M(s)) \qquad \text{for} \quad 0\leqq s\leqq t_M$$

and $W_M(t)$ is a local martingale.

It follows from (i) of Assumption 1.1 that

$$A_M(s)=A_M(s_M)\leqq c_1+c_2U_M(s)+c_3\beta(E[U_M(s)]) \tag{1.5}$$
$$\text{for} \quad 0\leqq s\leqq t_M,\ \text{where}\ s_M=s\wedge e_M.$$

Take the expectations on (1.4) and set $m(t)=E[U_M(t)]$. Then by (1.5) we get

$$m(t)\leqq r_0+\hat{c}\int_0^t g(m(s))ds\,, \tag{1.6}$$

where $r_0=E[\Psi(0,\ \phi)]+\Psi(0,\ 0)$, $\hat{c}=\max\{c_1,\ c_2,\ c_3\}$ and $g(r)=1+r+\beta(r)$.

By $r(t)$ denote the right-hand side of (1.6), so that

$$m(t)\leqq r(t) \quad \text{and hence} \quad r'(t)\leqq \hat{c}g(r(t)).$$

This inequality implies that (1.2) holds for $U(t)$ replaced by $U_M(t)$. Ito's formula applying to $(1+\Psi(t,\ x))^m$ and the condition (i) of Assumption 1.1 yield that (1.3) holds for $U(t)$ replaced by $U_M(t)$. So, by the same argument as in the proof of Theorem 2.1 of Narita [5] we find a solution $X(t)$ of (1.1) with the initial state $X(0)=\phi$ satisfying (1.2) and (1.3), for which $X_{M_j}(t\wedge e_{M_j})\to X(t)$ with probability 1 uniformly for each finite time interval as $j\to\infty$ for a certain subsequence $\{M_j\}_{j=1,2,\dots}$ of $\{M\}$. Let $Y(t)$ be another solution of (1.1) with the initial state $Y(0)=\phi$, having the probability distribution $v(t)=\mathcal{L}(Y(t)\,;\,P)$ with $v(0)=\mathcal{L}(\phi\,;\,P)=u(0)$. Set $\Delta(t)=X(t)-Y(t)$, and also put

$$\tau_M^X=\inf\{t\,;\ |X(t)|\geqq M\} \quad \text{and} \quad \tau_M^Y=\inf\{t\,;\ |Y(t)|\geqq M\}\,.$$

Then, since $X(t)$ and $Y(t)$ are the global solutions of (1.1), $\tau_M^X\to\infty$ and $\tau_M^Y\to\infty$ with probability 1 as $M\to\infty$. Moreover, by the same argument as in Narita [4, pp. 69-71] we find a constant $c_M>0$ depending on M such that

$$E[|\Delta(t_M)|^2]\leqq(c_M+1)\int_0^t E[|\Delta(s_M)|^2]ds\,,$$

where $t_M=t\wedge\tau_M^X\wedge\tau_M^Y$ and $s_M=s\wedge\tau_M^X\wedge\tau_M^Y$.

From this we get the pathwise uniqueness of the solution. Hence the proof is complete.

2. Weak convergence method. Let Λ denote a numerical set and $\lambda_0\in\Lambda$ be a limit point of Λ. Let $\{b^\lambda[t,\ x,\ \mu]\}_{\lambda\in\Lambda}$ be a family of d-vector functions and $\{\sigma^\lambda[t,\ x,\ \mu]\}_{\lambda\in\Lambda}$ be a family of $d\times d$-matrix functions, that are defined for $t\geqq 0$, $x\in R^d$ and $\mu\in\mathcal{P}(R^d)$. Introduce a family $\{\phi^\lambda\}_{\lambda\in\Lambda}$ of d-dimensional random vectors such that each ϕ^λ is independent of the d-dimensional Brownian motion process. Then we consider the following system of stochastic differential equations;

$$(2.1)^{\lambda} \qquad dX^{\lambda}(t)=b^{\lambda}[t, X^{\lambda}(t), u^{\lambda}(t)]dt+\sigma^{\lambda}[t, X^{\lambda}(t), u^{\lambda}(t)]dW(t),$$
$$X^{\lambda}(0)=\phi^{\lambda}, \quad \text{where } \lambda\in\Lambda\setminus\{\lambda_0\},$$

satisfying $u^{\lambda}(t)=\mathcal{L}(X^{\lambda}(t); P)$ with $u^{\lambda}(0)=\mathcal{L}(\phi^{\lambda}; P)$. Introducing some d-dimensional Brownian motion process $B(t)$ we consider the following stochastic differential equation;

$$(2.1)^{\lambda_0} \qquad dX^{\lambda_0}(t)=b^{\lambda_0}[t, X^{\lambda_0}(t), u^{\lambda_0}(t)]dt+\sigma^{\lambda_0}[t, X^{\lambda_0}(t), u^{\lambda_0}(t)]dB(t),$$
$$X^{\lambda_0}(0)=\phi^{\lambda_0},$$

satisfying $u^{\lambda_0}(t)=\mathcal{L}(X^{\lambda_0}(t); P)$ with $u^{\lambda_0}(0)=\mathcal{L}(\phi^{\lambda_0}; P)$. Here we find under what conditions the processes $X^{\lambda}(t)$ converge weakly for $\lambda\to\lambda_0$ to $X^{\lambda_0}(t)$.

We begin with the martingale problem approach. Given $\mu\in\mathcal{P}(R^d)$, let $L_t^{\lambda}(\mu)$ be the differential generator associated with $(2.1)^{\lambda}$. Namely, for $t\geqq 0$, $x\in R^d$, $\mu\in\mathcal{P}(R_p)$ and $f\in C^2(R^d)$, $L_t^{\lambda}(\mu)$ has the form

$$L_t^{\lambda}(\mu)f(x)=\left[\langle b^{\lambda}, \operatorname{grad}_x f\rangle+\frac{1}{2}\operatorname{tr}(a^{\lambda}f_{xx})\right](t, x, \mu),$$

where $a^{\lambda}[t, x, \mu]=\sigma^{\lambda}[t, x, \mu](\sigma^{\lambda}[t, x, \mu])^*$.

By $L_t^{\lambda_0}(\mu)$ denote the differential generator associated with $(2.1)^{\lambda_0}$, that is defined by $L_t^{\lambda}(\mu)$ with λ replaced by λ_0.

Let $\Omega=C([0, \infty); R^d)$. For each $\omega\in\Omega$ and $t\geqq 0$ we denote $X(t, \omega)=\omega(t)$. Let $\boldsymbol{F}_t$ and $\boldsymbol{F}$ be the σ-algebras generated by $\{X(s); 0\leqq s\leqq t\}$ and $\{X(s); 0\leqq s<\infty\}$, respectively. For $f\in C_0^{\infty}(R^d)$, put

$$M_f^{\lambda_0}(t)=f(X(t))-\int_0^t L_s^{\lambda_0}(u^{\lambda_0}(s))f(X(s))ds,$$

where $u^{\lambda_0}(s)=u^{\lambda_0}(s, dx)\in\mathcal{P}(R^d)$.

When an initial distribution $\nu^{\lambda_0}\in\mathcal{P}(R^d)$ is specified, we say that a probability measure P^{λ_0} on Ω is a solution of the martingale problem for $(L_t^{\lambda_0}(u^{\lambda_0}), \nu^{\lambda_0})$ if $\{M_f^{\lambda_0}(t), \boldsymbol{F}_t, P^{\lambda_0}; 0\leqq t<\infty\}$ is a martingale for $f\in C_0^{\infty}(R^d)$, satisfying

$$\mathcal{L}(X(t); P^{\lambda_0})=u^{\lambda_0}(t) \text{ with } u^{\lambda_0}(0)=\nu^{\lambda_0}.$$

Remark 2.1. Let $\{P^{\lambda}\}_{\lambda\in\Lambda\setminus\{\lambda_0\}}$ of probability measures in $C([0, \infty); R^d)$ induced by $X^{\lambda}(t)$ be relatively weakly compact. Then, by the representing theorem of Skorokhod [8], without loss of generality, we can assume that there exist a subsequnce $\{\lambda_j\}_{j=1,2,\dots}$ of $\{\lambda\}_{\lambda\in\Lambda\setminus\{\lambda_0\}}$ and a random process $\tilde{X}^{\lambda_0}(t)$, such that $X^{\lambda_j}(t)\to\tilde{X}^{\lambda_0}(t)$ with probability 1 uniformly for each finite time interval as $j\to\infty$.

Theorem 2.1. *Suppose that the following conditions hold*: (0) $Eq(2.1)^{\lambda}$ *has a pathwise unique solution* $X^{\lambda}(t)$ *with the initial state* $X^{\lambda}(0)=\phi^{\lambda}$, *such that* $u^{\lambda}(t)$ *denotes the probability distribution of* $X^{\lambda}(t)$ *with the initial distribution* $u^{\lambda}(0)=\mathcal{L}(\phi^{\lambda}; P)$. *The family* $\{P^{\lambda}\}_{\lambda\in\Lambda\setminus\{\lambda_0\}}$ *of probability measures in* $C([0, \infty); R^d)$ *induced by* $X^{\lambda}(\cdot)$ *is relatively weakly compact, and* $u^{\lambda}(0)$ *converges weakly to* $u^{\lambda_0}(0)=\mathcal{L}(\phi^{\lambda_0}; P)$ *as* $\lambda\to\lambda_0$.

(I) *The martingale problem for* $(L_t^{\lambda_0}(u^{\lambda_0}), \nu^{\lambda_0})$ *has a unique solution in* $C([0, \infty); R^d)$ *for the initial distribution* ν^{λ_0}, *where* $\nu^{\lambda_0}=\mathcal{L}(\phi^{\lambda_0}; P)$.

(II) *There exist a subsequence* $\{\lambda_j\}_{j=1,2,\dots}$ *of* $\{\lambda\}_{\lambda\in\Lambda\setminus\{\lambda_0\}}$ *and a random process* $\tilde{X}^{\lambda_0}(t)$ *for which Remark* 2.1 *holds, satisfying for each* $T<\infty$ *and* $f\in C^2(R^d)$ *with compact support in* R^d

$$E\left[\left|\int_t^T L_s^{\lambda_j}(u^{\lambda_j}(s))f(X^{\lambda_j}(s))-L_s^{\lambda_0}(\tilde{u}^{\lambda_0}(s))f(X^{\lambda_j}(s))\}ds\right|\right]\longrightarrow 0$$

for each $t\leqq T$ *as* $j\to\infty$, *where* $\tilde{u}^{\lambda_0}(s)=\mathcal{L}(\tilde{X}^{\lambda_0}(s); P)$. *Then,* $X^{\lambda}(t)$, $t\geqq 0$, *with the initial state* $X^{\lambda}(0)=\phi^{\lambda}$, *converges weakly in* $C([0, T]; R^d)$, $T<\infty$, *but arbitrary, as* $\lambda\to\lambda_0$ *to the solution* $X^{\lambda_0}(t)$ *of* $(2.1)^{\lambda_0}$, $t\geqq 0$, *with the initial state* $X^{\lambda_0}(0)=\phi^{\lambda_0}$.

Proof. By the condition (0) there are a subsequence $\{\lambda_j\}_{j=1,2,\dots}$ of $\Lambda\setminus\{\lambda_0\}$ and a probability measure P^{λ_0} on $C([0, \infty); R^d)$ for which P^{λ_j} converges weakly as $j\to\infty$ to P^{λ_0}. Since the condition (I) holds, we have only to identify the limit measure by showing that it solves the martingale problem for $(L_t^{\lambda_0}(u^{\lambda_0}), \nu^{\lambda_0})$, satisfying $u^{\lambda_0}(t)=\mathcal{L}(X(t); P^{\lambda_0})$ with $u^{\lambda_0}(0)=\nu^{\lambda_0}$. For the sake of convenience, we will use $\{\lambda\}$ to denote the subsequence $\{\lambda_j\}$. From the martingale characterization for the solution $X^{\lambda}(t)$ of $(2.1)^{\lambda}$ it follows that for $f\in C_0^{\infty}(R^d)$

$$f(X^{\lambda}(t))-f(X^{\lambda}(s))-\int_s^t L_r^{\lambda_0}(\tilde{u}^{\lambda_0}(r))f(X^{\lambda}(r))dr \tag{2.2}$$

$$=\int_s^t\{L_r^{\lambda}(u^{\lambda}(r))f(X^{\lambda}(r))-L_r^{\lambda_0}(\tilde{u}^{\lambda_0}(r))f(X^{\lambda}(r))\}dr$$

$$+M_f^{\lambda}(t)-M_f^{\lambda}(s)$$

for $0\leqq s\leqq t\leqq T$, where $M_f^{\lambda}(t)$ is a bounded martingale. When superscript λ appears on the paths, we do not refer explicity to the measure P^{λ}. Multiply (2.2) by any continuous functional $H(s)$ which is $\boldsymbol{F}_s$-measurable, and then take the expectation. Then, by the condition (II) we get

$$E[H(s)\{f(X^{\lambda}(t))-f(X^{\lambda}(s))-\int_s^t L_r^{\lambda_0}(\tilde{u}^{\lambda_0}(r))f(X^{\lambda}(r))dr\}]\longrightarrow 0$$

for $0\leqq s\leqq t\leqq T$.

Hence the proof is complete.

3. Averaging method. Here we give some useful criteria for satisfaction of the conditions (0), (I) and (II) of Theorem 2.1. For simplicity we consider $(2.1)^{\lambda}$ with the initial state that does not depend on λ. In the following let ϕ be a d-dimensional random vector independent of the d-dimensional Brownian motion process.

Theorem 3.1. *Suppose that the following conditions hold:*

(0) *For all* $\lambda\in\Lambda$, $b^{\lambda}[t, x, \mu]$ *and* $\sigma^{\lambda}[t, x, \mu]$ *satisfy Assumption* 1.0 *with the family* $\{K_M, K, c, p\}$ *of constants independent of* λ. *For* $\lambda\neq\lambda_0$, $(2.1)^{\lambda}$ *has a pathwise unique solution* $X^{\lambda}(t)$ *with the initial state* $X^{\lambda}(0)=\phi$, *such that for any* $T<\infty$

$$\sup_{0\leq t\leq T} E[|X^{\lambda}(t)|^{4q}]<\infty \qquad \text{uniformly for} \quad \lambda\in\Lambda\setminus\{\lambda_0\},$$

where $q=\max\{2p, 2\}$ *and* p *is the integer given in Assumption* 1.0.

(I) *Eq*$(2.1)^{\lambda_0}$ *has a pathwise unique solution* $X^{\lambda_0}(t)$ *with the initial state* $X^{\lambda_0}(0)=\phi$.

(II)
$$\lim_{\lambda\to\lambda_0}\int_{t_1}^{t_2} c^{\lambda}[t, x, \mu]dt=\int_{t_1}^{t_2} c^{\lambda_0}[t, x, \mu]dt$$

for $0\leq t_1\leq t_2\leq T$, *where* $c^{\lambda}[t, x, \mu]$ *denotes the vector function* $b^{\lambda}[t, x, \mu]$ *and the matrix function* $a^{\lambda}[t, x, \mu]=\sigma^{\lambda}[t, x, \mu](\sigma_{\lambda}[t, x, \mu])^*$. *Then,* $X^{\lambda}(t)$, $t\geq 0$, *with the initial state* $X^{\lambda}(0)=\phi$, *converges weakly in* $C([0, T]; R^d)$, $T<\infty$, *but arbitrary, as* $\lambda\to\lambda_0$ *to the solution* $X^{\lambda_0}(t)$ *of* $(2.1)^{\lambda_0}$, $t\geq 0$, *with the initial state* $X^{\lambda_0}(0)=\phi^{\lambda_0}$.

For the proof of the theorem we prepare the following lemmas.

Lemma 3.1. *Under the same assumptions as in Theorem* 3.1, *let* $\{k^{\lambda}[t, x, \mu]\}_{\lambda\in\Lambda}$ *be a family of scalar functions satisfying the following conditions:*

(i)
$$\sup_{\lambda\in\Lambda}|k^{\lambda}[t, x, \mu]-k^{\lambda}[t, y, \nu]|\leq K_M|x-y|+K\|\mu-\nu\|$$

for all $t\geq 0$, $|x|\leq M$, $|y|\leq M$, $\mu\in\mathcal{P}(R^d)$ *and* $\nu\in\mathcal{P}(R^d)$ *with some constants* $K_M>0$ *depending on* M *and* $K>0$.

(ii)
$$\sup_{\lambda\in\Lambda}|k^{\lambda}[t, x, \mu]|\leq\tilde{c}(1+\tilde{\kappa}+\langle\tilde{\kappa}, \mu\rangle)$$

for all $t\geq 0$, $x\in R^d$ *and* $\mu\in\mathcal{M}(\tilde{\kappa})=\{\mu\in\mathcal{P}(R^d); \langle\tilde{\kappa}, \mu\rangle<\infty\}$, *where* $\tilde{\kappa}(x)=|x|^{2q}$, $q=\max\{2p, 2\}$ *and* p *is as in Assumption* 1.0.

(iii)
$$\lim_{\lambda\to\lambda_0}\int_{t_1}^{t_2} k^{\lambda}[t, x, \mu]dt=\int_{t_1}^{t_2} k^{\lambda_0}[t, x, \mu]dt$$

uniformly with respect to $0<t_2-t_1<T$ *for each* x *and* μ. *Let* $\{\lambda_j\}$ *and* $\tilde{X}^{\lambda_0}(t)$

be a subsequence and a random process, respectively, for which Remark 2.1 *holds. For the sake of simplicity, let us consider that* $\{\lambda_j\}=\{\lambda\}$. *Then*

$$\lim_{\lambda\to\lambda_0} E\left[\left|\int_0^T \{k^\lambda[s, X^\lambda(s), u^\lambda(s)]-k^{\lambda_0}[s, X^\lambda(s), \tilde{u}^{\lambda_0}(s)]\}ds\right|\right]=0,$$

where $u^\lambda(s)$ *and* $\tilde{u}^{\lambda_0}(s)$ *denote the probability distributions of* $X^\lambda(s)$ *and* $\tilde{X}^{\lambda_0}(s)$ *with the same initial states* $X^\lambda(0)=\tilde{X}^{\lambda_0}(0)=\phi$, *respectively.*

Proof. First note the following inequality:

$$\left|\int_0^T \{k^\lambda[s, X^\lambda(s), u^\lambda(s)]-k^{\lambda_0}[s, X^\lambda(s), \tilde{u}^{\lambda_0}(s)]\}ds\right|$$

$$\leqq \int_0^T |I_1|ds+\left|\int_0^T I_2 ds\right|+\int_0^T |I_3|ds,$$

where

$$I_1=k^\lambda[s, X^\lambda(s), u^\lambda(s)]-k^\lambda[s, \tilde{X}^{\lambda_0}(s), \tilde{u}^{\lambda_0}(s)],$$

$$I_2=k^\lambda[s, \tilde{X}^{\lambda_0}(s), \tilde{u}^{\lambda_0}(s)]-k^{\lambda_0}[s, \tilde{X}^{\lambda_0}(s), \tilde{u}^{\lambda_0}(s)],$$

$$I_3=k^{\lambda_0}[s, \tilde{X}^{\lambda_0}(s), \tilde{u}^{\lambda_0}(s)]-k^{\lambda_0}[s, X^\lambda(s), \tilde{u}^{\lambda_0}(s)].$$

Secondly note that there is a constant $c'>0$ satisfying

$$\sup_{\lambda\in\Lambda}|k^\lambda[t, x, \mu]|^2\leqq c'(1+\psi+\langle\psi, \mu\rangle)$$

for all $t\geqq 0$, $x\in R^d$ and $\mu\in\mathcal{M}(\psi)=\{\mu\in\mathcal{P}(R^d); \langle\psi, \mu\rangle<\infty\}$, where $\psi(x)=|x|^{4q}$, $q=\max\{2p, 2\}$ and p is as in Assumption 1.0. Then, since $\sup_{0\leq t\leq T} E[|X^\lambda(t)|^{4q}]<\infty$ uniformly for $\lambda\in\Lambda\setminus\{\lambda_0\}$ and since $X^\lambda(t)\to\tilde{X}^{\lambda_0}(t)$ with probability 1 as $\lambda\to\lambda_0$ by Remark 2.1, we obtain

(3.1) $E[|I_i|^2]<\infty$ uniformly for $\lambda\in\Lambda$, where $i=1, 2$ and 3.

Introduce the indicator function

$$\chi_r(x)=\begin{cases} 1 & \text{if } |x|\leqq r, \\ 0 & \text{otherwise}. \end{cases}$$

Then by the condition (i) and (3.1) we get the following estimates;

$$E[\chi_r(X^\lambda(s))\chi_r(\tilde{X}^{\lambda_0}(s))I_1] \longrightarrow 0 \quad \text{as } \lambda\to\lambda_0,$$

$$E[\{1-\chi_r(X^\lambda(s))\chi_r(\tilde{X}^{\lambda_0}(s))\}I_1] \longrightarrow 0 \quad \text{as } r\to\infty,$$

$$E[\chi_r(\tilde{X}^{\lambda_0}(s))\chi_r(X^\lambda(s))I_3] \longrightarrow 0 \quad \text{as } \lambda\to\lambda_0,$$

$$E[\{1-\chi_r(\tilde{X}^{\lambda_0}(s))\chi_r(X^\lambda(s))\}I_3] \longrightarrow 0 \quad \text{as } r\to\infty.$$

On the other hand, since $\tilde{X}^{\lambda_0}(s)$ is continuous in s with probability 1 and since the family $\{|\tilde{X}^{\lambda_0}(s)|^2; 0 \leqq s \leqq T\}$ is uniformly integrable, the probability distribution $\tilde{u}^{\lambda_0}(s)$ of the process $\tilde{X}^{\lambda_0}(s)$ is continuous in s with respect to the norm $\| \ \|$. For a moment assume that the following Lemma 3.2 holds. Then Lemma 3.2 applies with $\tilde{X}(s)=\tilde{X}^{\lambda_0}(s)$ and $\tilde{u}(s)=\tilde{u}^{\lambda_0}(s)$. So, by the dominated convergence theorem we get

$$E\left[\left|\int_0^T I_2 ds\right|\right] \longrightarrow 0 \quad \text{as } \lambda \to \lambda_0 .$$

Hence the proof is complete.

Lemma 3.2. *Let* $\{k^\lambda[t, x, \mu]\}_{\lambda \in \Lambda}$ *be a family of scalar functions satisfying the conditions* (i) *and* (iii) *of Lemma* 3.1. *Let* $\tilde{X}(s)$ *be a random process such that* $\tilde{X}(t)$ *is continuous with probability* 1 *and suppose that the probability distribution* $\tilde{u}(t)$ *of* $\tilde{X}(t)$ *is continuous with respect to the norm* $\| \ \|$. *Then*

$$\lim_{\lambda \to \lambda_0} \int_0^t k^\lambda[s, \tilde{X}(s), \tilde{u}(s)]ds = \int_0^t k^{\lambda_0}[s, \tilde{X}(s), \tilde{u}(s)]ds$$

with probability 1 *for each* $t \geqq 0$.

Proof. By the same argument as in the proof of Gikhman and Skorokhod's [1, p. 344] lemma we can get the conclusion, and so we omit the details.

Proof of Theorem 3.1. Step 0 (Relative compactness). Since $b^\lambda[t, x, \mu]$ and $\sigma^\lambda[t, x, \mu]$ satisfy Assumption 1.0 with the family $\{K_M, K, c, p]$ of constants independent of λ, there is a constant $\tilde{c}>0$ being independent of $\lambda \in \Lambda$ such that

$$|b^\lambda[t, x, \mu]|^4 + |\sigma^\lambda[t, x, \mu]|^4 \leqq \tilde{c}(1+\tilde{\kappa}+\langle \tilde{\kappa}, \mu \rangle) \tag{3.2}$$

for all $t \geqq 0$, $x \in R^d$ and $\mu \in \mathcal{M}(\tilde{\kappa})$, where $\tilde{\kappa}(x)=|x|^{2q}$, $q=\max\{2p, 2\}$ and p is as in Assumption 1.0. Let $T<\infty$ be arbitrary and fixed. Observe $(2.1)^\lambda$ and use the Schwarz inequality, so that

$$|X^\lambda(t)-X^\lambda(s)|^4 \leqq 8\left[(t-s)^3 \int_s^t |b^\lambda[r, X^\lambda(r), u^\lambda(r)]|^4 dr + \left|\int_s^t \sigma^\lambda[r, X^\lambda(r), u^\lambda(r)]dW(r)\right|^4\right].$$

The assumption that $\sup_{0 \leq r \leq T} E[|X^\lambda(r)|^{4q}]<\infty$ uniformly for $\lambda \in \Lambda \setminus \{\lambda_0\}$ and the estimate (3.2) imply

$$E[|X^\lambda(t)-X^\lambda(s)|^4] \leqq D(t-s)^2 \qquad \text{for all } 0 \leqq s \leqq t \leqq T$$

with a constant $D>0$ being independent of $\lambda\in\Lambda\setminus\{\lambda_0\}$. So, by Prokhorov [7] the family $\{P^\lambda\}_{\lambda\in\Lambda\setminus\{\lambda_0\}}$ of probability measures in $C([0, \infty); R^d)$ induced by $X^\lambda(t)$ is relatively weakly compact. Namely the condition (0) of Theorem 2.1 is satisfied.

Step 1 (Limit measure). The pathwise uniqueness for solutions of $(2.1)^{\lambda_0}$ implies the uniqueness in the sense of distributions which is equivalent to the uniqueness for solutions of the martingale problem, and so the condition (I) of Theorem 2.1 is satisfied.

Step 2 (Perturbed test). Let $\{\lambda_j\}_{j=1,2,\dots}$ and $\tilde{X}^{\lambda_0}(t)$ be a sequence and a random process, respectively, for which Remark 2.1 hold. Let f be in $C^2(R^d)$, having compact support in R^d. Further observe the components of the form

$$L_s^\lambda(u^\lambda(s))f(X^\lambda(s))-L_s^{\lambda_0}(\tilde{u}^{\lambda_0}(s))f(X^\lambda(s)),\quad \text{where}\quad \tilde{u}^{\lambda_0}(s)=\mathcal{L}(\tilde{X}^{\lambda_0}(t)\,;\,P).$$

Then, by the assumption, since all coefficients b^λ and σ^λ, where $\lambda\in\Lambda$, satisfy Assumption 1.0 with the constants independent of λ, Lemma 3.1 implies that the condition (II) of Theorem 2.1 is satisfied. Hence the proof is complete.

The following convergence theorem for SDE of the McKean type is an analogue of that for SDE of the Ito type given in Gikhman and Skorokhod [1, p. 338].

Theorem 3.2. *Suppose that* $(2.1)^\lambda$ *satisfies Assumption* 1.0 *with the family* $\{K_M, K, c, p\}$ *of constants independent of* $\lambda\in\Lambda\setminus\{\lambda_0\}$, *and Assumption* 1.1 *with the family* $\{c_1, c_2, c_3\}$ *of constants replaced by* $\{c_1(\lambda), c_2(\lambda), c_3(\lambda)\}$ *depending on* $\lambda\in\Lambda\setminus\{\lambda_0\}$, *and the function* $\beta(\cdot)$ *and the constant* l *independent of* $\lambda\in\Lambda\setminus\{\lambda_0\}$. *By* $\Psi_\lambda(t, x)$ *denote the function* $\Psi(t, x)$ *given in Assumption* 1.1 *for* $(2.1)^\lambda$, *considering the dependence on* $\lambda\in\Lambda\setminus\{\lambda_0\}$. *Suppose that there exist a family* $\{\hat{c}_1, \hat{c}_2, \hat{c}_3\}$ *of constants and a scalar function* $\hat{\Psi}(t, x)$ *such that*

$$\sup_{\lambda\in\Lambda\setminus\{\lambda_0\}} c_i(\lambda)\leqq\hat{c}_i,\quad \textit{where } i=1, 2 \textit{ and } 3,$$

and $\sup_{\lambda\in\Lambda\setminus\{\lambda_0\}} \Psi_\lambda(t, x)\leqq\hat{\Psi}(t, x).$

Further suppose that $(2.1)^{\lambda_0}$ *satisfies Assumption* 1.0 *and Assumption* 1.1. *By* $\bar{\Psi}(t, x)$ *denote the function* $\Psi(t, x)$ *given in Assumption* 1.1 *for* $(2.1)^{\lambda_0}$.

Let ϕ *be any d-dimensional random vector independent of the d-dimensional Brownian motion process, satisfying*

$$E[\hat{\Psi}(0, \phi)^{2q_1}]<\infty,\quad E[\bar{\Psi}(0, \phi)^{2q_2}]<\infty,\quad q_i=\max\{2p_i, 2\},$$

where $i=1, 2$, *and* p_1 *and* p_2 *are the integers given in Assumption* 1.0 *for* $(2.1)^\lambda$ *and* $(2.1)^{\lambda_0}$, *respectively.*

Let the condition (II) *of Theorem* 3.1 *hold.*

Then, $X^\lambda(t)$, $t\geqq 0$, *with the initial state* $X^\lambda(0)=\phi$, *converges weakly in*

$C([0, T]; R^d)$, $T<\infty$, but arbitrary, as $\lambda\to\lambda_0$ to the solution $X^{\lambda_0}(t)$ of $(2.1)^{\lambda_0}$, *$t\geqq 0$, with the initial state $X^{\lambda_0}(0)=\phi$.*

Proof. Since $(2.1)^{\lambda}$ and $(2.1)^{\lambda_0}$ satisfy Assumption 1.0 and Assumption 1.1, by Theorem 1.1 the existence and uniqueness of the solution holds for $(2.1)^{\lambda}$ and $(2.1)^{\lambda_0}$. Note that

$$|x|^{4q_1}\leqq \text{const}(1+\Psi_{\lambda}(t, x))^{2q_1}\leqq \text{const}(1+\hat{\Psi}(t, x))^{2q_1}$$

for all $t\geqq 0$ and $x\in R^d$.

Then, by the assumption, since $E[\hat{\Psi}(0, \phi)^{2q_1}]<\infty$, the estimate (1.3) of Theorem 1.1 applies to the solution $X^{\lambda}(t)$, which yields

$$\sup_{0\leq t\leq T} E[|X^{\lambda}(t)|^{4q_1}]<\infty \quad \text{uniformly for } \lambda\in\Lambda\setminus\{\lambda_0\}.$$

Consider that the condition (II) of Theorem 3.1 is assumed. Then by Theorem 3.1 we get the conclusion. Hence the proof is complete.

4. Mean-field with a small parameter. Here we give an application of Theorem 3.1. For a small parameter ε such that $0<\varepsilon\leqq 1$, we consider the following two-dimensional stochastic differential equation;

$$(4.1) \qquad dz(t)=[Az(t)+\varepsilon v(z(t))-\varepsilon\Gamma\{z(t)-E(z(t))\}]dt+\sqrt{\varepsilon}\,D(z(t))dW(t),$$

where $W(t)$ is a two-dimensional Brownian motion process, $E(\ )$ denotes the mathematical expectation, A and Γ are 2×2-matrices, v is a vector function and D is a matrix function. Hereafter we assume the following conditions:

$A=\begin{pmatrix} 0 & 1 \\ -\omega^2 & 0 \end{pmatrix}$ with a constant $\omega>0$.

$v(z)=\begin{pmatrix} v_1(z) \\ v_2(z) \end{pmatrix}$ satisfies the local Lipschitz condition in $z\in R^2$.

$\Gamma=\begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix}$ with constant components γ_{ij} satisfying either $\gamma_{11}+\gamma_{22}>0$ or $\omega\gamma_{12}-\dfrac{\gamma_{21}}{\omega}>0$.

$D(z)=\begin{pmatrix} \delta_{11}(z) & \delta_{12}(z) \\ \delta_{21}(z) & \delta_{22}(z) \end{pmatrix}$ satisfies the local Lipschitz condition in $z\in R^2$.

Our purpose is to obtain a limit diffusion for (4.1) on the time scale $1/\varepsilon$. First we observe that the deterministic equation (4.1) with $\varepsilon=0$ has the equilibrium point (0, 0) in the plane as the center, and hence we introduce the following matrix and process:

For $\omega>0$ and $t\geqq 0$, define the matrix $\Theta^{\omega}(t)$ by

$$\Theta^{\omega}(t)=\begin{pmatrix}\cos\omega t & -\frac{1}{\omega}\sin\omega t\\ \sin\omega t & \frac{1}{\omega}\cos\omega t\end{pmatrix}.$$

Let $z(t)=(x(t), y(t))$ be a solution of (4.1), and define the process $\zeta^{\varepsilon}(t)=(\xi^{\varepsilon}(t), \eta^{\varepsilon}(t))$ by

$$\begin{pmatrix}\xi^{\varepsilon}(t)\\ \eta^{\varepsilon}(t)\end{pmatrix}=\Theta^{\omega}(t/\varepsilon)\begin{pmatrix}x(t/\varepsilon)\\ y(t/\varepsilon)\end{pmatrix}.$$

We note that

$$\text{if }\begin{pmatrix}\xi\\ \eta\end{pmatrix}=\Theta^{\omega}(t)\begin{pmatrix}x\\ y\end{pmatrix},\text{ then }\omega^2x^2+y^2=\omega^2(\xi^2+\eta^2).$$

Then $\zeta^{\varepsilon}(t)$ satisfies the following stochastic differential equations;

$$(4.1)^{\varepsilon}\qquad d\zeta^{\varepsilon}(t)=\Big[b_0\Big(\frac{t}{\varepsilon}, \zeta^{\varepsilon}(t)\Big)-\Gamma_0\Big(\frac{t}{\varepsilon}\Big)\{\zeta^{\varepsilon}(t)-E(\zeta^{\varepsilon}(t))\}\Big]dt$$
$$+\sigma_0\Big(\frac{t}{\varepsilon}, \zeta^{\varepsilon}(t)\Big)dW_0(t),$$

where $W_0(t)$ is a new Brownian motion process defined by $W_0(t)=\sqrt{\varepsilon}\,W(t/\varepsilon)$. Here and below b_0 is the vector function, Γ_0 and σ_0 are the matrix functions; these are given by the following definitions:

For $t\geqq 0$ and $\zeta\in R^d$, define the vector function $b_0(t, \zeta)$ by

$$b_0(t, \zeta)=\Theta^{\omega}(t)v(\Theta^{\omega}(t)^{-1}\zeta),$$

and define the matrix functions $\Gamma_0(t)$ and $\sigma_0(t, \zeta)$ by

$$\Gamma_0(t)=\Theta^{\omega}(t)\Gamma\Theta^{\omega}(t)^{-1},\qquad \sigma_0(t, \zeta)=\Theta^{\omega}(t)D(\Theta^{\omega}(t)^{-1}\zeta).$$

In order to get an approximation for the solution $z(t)$ of (4.1) we adopt the averaging method over the time interval $[0, 2\pi/\omega]$.

Definition 4.1. For $\zeta\in R^2$, define the vector function $\bar{b}(\zeta)$ and the matrix function $\bar{a}(\zeta)$ by

$$\bar{b}(\zeta)=\frac{\omega}{2\pi}\int_0^{2\pi/\omega}b_0(t, \zeta)dt,\qquad \bar{a}(\zeta)=\frac{\omega}{2\pi}\int_0^{2\pi/\omega}a_0(t, \zeta)dt,$$

where $a_0(t, \zeta)=\sigma_0(t, \zeta)(\sigma_0(t, \zeta))^*$. Let $\bar{\sigma}(\zeta)$ be the symmetric square root of

$\bar{a}(\zeta)$; i.e., $\bar{\sigma}(\zeta)\bar{\sigma}^*(\zeta)=\bar{a}(\zeta)$. Further define the matrix $\bar{\Gamma}$ by

$$\frac{1}{2}\bar{\Gamma}=\frac{\omega}{2\pi}\int_0^{2\pi/\omega}\Gamma_0(t)dt.$$

Accordingly we have derived the following equation;

$$d\bar{\zeta}(t)=\left[\bar{b}(\bar{\zeta}(t))-\frac{1}{2}\bar{\Gamma}\{\bar{\zeta}(t)-E(\bar{\zeta}(t))\}\right]dt+\bar{\sigma}(\bar{\zeta}(t))dB(t), \tag{4.2}$$

where $B(t)$ is a two-dimensional Brownian motion process.

Remark 4.1. By an elemetary calculation we get

$$\bar{\Gamma}=\begin{pmatrix} \gamma_{11}+\gamma_{22} & \omega\gamma_{12}-\dfrac{\gamma_{21}}{\omega} \\ -\left(\omega\gamma_{12}-\dfrac{\gamma_{21}}{\omega}\right) & \gamma_{11}+\gamma_{22} \end{pmatrix}.$$

The assumption on the components γ_{ij} of the matrix Γ implies that $\bar{\Gamma}$ is a nonzero matrix. When $D(z)$ has the form

$D(z)=\begin{pmatrix}\delta_{11} & \delta_{12} \\ \delta_{21} & \delta_{22}\end{pmatrix}$ with constants components δ_{ij}, we get

$$\bar{\sigma}(\zeta)\equiv\bar{\sigma}=\frac{1}{\sqrt{2}\,\omega}(\omega^2\delta_1^2+\delta_2^2)^{1/2}\begin{pmatrix}1 & 0 \\ 0 & 1\end{pmatrix},$$

where $\delta_1^2=\delta_{11}^2+\delta_{12}^2$ and $\delta_2^2=\delta_{21}^2+\delta_{22}^2$.

Remark 4.2. Eq (4.1) is a special case of (1.1) where

$$b[t, z, \mu]=Az+\varepsilon v(z)-\varepsilon\Gamma\int_{R^2}(z-\tilde{z})\mu(d\tilde{z}),$$

$$\sigma[t, z, \mu]=D(z) \quad \text{for } t\geqq 0,\ z\in R^d \text{ and } \mu\in\mathscr{P}(R^2).$$

Eq (4.2) is a special case of (1.1) where

$$b[t, \zeta, \mu]=\bar{b}(\zeta)-\frac{1}{2}\bar{\Gamma}\int_{R^2}(\zeta-\tilde{\zeta})\mu(d\tilde{\zeta}),$$

$$\sigma[t, \zeta, \mu]=\bar{\sigma}(\zeta) \quad \text{for } t\geqq 0,\ \zeta\in R^2 \text{ and } \mu\in\mathscr{P}(R^2).$$

Hereafter, for a two-dimensional random vector $\phi=(\phi^1, \phi^2)$, define ϕ^ω by $\phi^\omega=\Theta^\omega(0)\begin{pmatrix}\phi^1 \\ \phi^2\end{pmatrix}$; i.e. $\phi^\omega=\begin{pmatrix}\phi^1 \\ \phi^2/\omega\end{pmatrix}$.

Then as an application of Theorem 3.1 we get the following theorem.

Theorem 4.1. *Suppose that* (4.1) *satisfies Assumption* 1.0 *with the family* $\{K_M, K, c, p\}$ *of constants independent of* ε *and Assumption* 1.1 *with the family* $\{c_1, c_2, c_3\}$ *of constants replaced by* $\{\varepsilon c_1, \varepsilon c_2, \varepsilon c_3\}$,*and the function* $\beta(\cdot)$ *and the constant* l *independent of* ε. *By* $\Psi_\varepsilon(z)$ *denote the function* Ψ *given in Assumption* 1.1 *for* (4.1), *considering the dependence on* ε, *and suppose that there exists a scalar function* $\hat{\Psi}$ *on* R^2 *such that*

$$\sup_{0<\varepsilon\leqq 1}\Psi_\varepsilon(z)\leqq\hat{\Psi}(z) \qquad \textit{for all} \quad z\in R^2.$$

Further suppose that (4.2) *satisfies Assumption* 1.0 *and Assumption* 1.1. *By* $\bar{\Psi}(\zeta)$ *denote the function* Ψ *given in Assumption* 1.1 *for* (4.2).

Let ϕ *be any two-dimensional random vector independent of the two-dimensional Brownian motion process, satisfying*

$$E[\hat{\Psi}(\phi)^{2q}]<\infty, \quad E[\bar{\Psi}(\phi^\omega)^{2q}]<\infty, \quad q=\max\{2p, 2\},$$

where p *is the integer given in Assumption* 1.0 *for* (4.1). *Let* $\zeta^\varepsilon(t)$ *be the solution of* $(4, 1)^\varepsilon$ *with the initial state* $\zeta^\varepsilon(0)=\phi^\omega$.

Then $\zeta^\varepsilon(t)$ *converges weakly in* $C([0, T]; R^2)$, $T<\infty$, *but arbitrary, as* $\varepsilon\to 0$ *to the solution* $\bar{\zeta}(t)$ *of* (4.2), $t\geqq 0$, *with the initial state* $\bar{\zeta}(0)=\phi^\omega$.

Proof. Since (4.1) satisfies Assumption 1.0 and Assumption 1.1, Theorem 1.1 implies that (4.1) has the pathwise unique solution $z(t)$ with the initial state $z(0)=\phi$ and that the estimates (1.2) and (1.3) hold for $z(t)$. Put $U_\varepsilon(t)=\Psi_\varepsilon(z(t))$. Then, since $\Psi_\varepsilon(z)\leqq\hat{\Psi}(z)$ and since $E[\hat{\Psi}(\phi)^{2q}]<\infty$, by (1.3) we get the following estimate: For $2\leqq m\leqq 2q$,

$$E[(1+U_\varepsilon(t))^m]\leqq E[(1+\hat{\Psi}(\phi)+\hat{\Psi}(0))^m]\exp\left\{\int_0^t I_m^\varepsilon(s)ds\right\}, \tag{4.3}$$

where $I_m^\varepsilon(t)=\varepsilon m(m-1)\{c_1+c_2+c_3\beta(J^\varepsilon(t))\}$,

$$J^\varepsilon(t)=f^{-1}(f(\hat{r}_0)+\varepsilon\hat{c}t), \quad \hat{r}_0=E[\hat{\Psi}(\phi)]+\hat{\Psi}(0), \quad \hat{c}=\max\{c_1, c_2, c_3\}$$

and $f^{-1}(\cdot)$ is the inverse function of $f(s)=\int_0^s\frac{dr}{1+r+\beta(r)}$. Let $z(t)=(x(t), y(t))$ be the solution of (4.1) with the initial state $z(0)=\phi$. Then, since $\Theta^\omega(t)$ is a nonsingular matrix and since $\zeta^\varepsilon(t)=\Theta^\omega(t/\varepsilon)z(t/\varepsilon)$, $\zeta^\varepsilon(t)$ is the pathwise unique solution of $(4.1)^\varepsilon$ with the initial state $\zeta^\varepsilon(0)=\phi^\omega$. The definition of $\zeta^\varepsilon(t)$ implies that

$$\omega^2|\zeta^\varepsilon(t)|^2=\omega^2x(t/\varepsilon)^2+y(t/\varepsilon)^2$$

and hence $|\zeta^\varepsilon(t)|^2\leqq\hat{c}_\omega|z(t/\varepsilon)|^2$, where $\hat{c}_\omega=\max\{1, 1/\omega^2\}$. By the condition (ii) of Assumption 1.1, since $|z|^2\leqq l\Psi_\varepsilon(z)$ for all $z\in R^2$, we have

$$\tilde{d}|\zeta^\varepsilon(t)|^2\leqq U_\varepsilon(t/\varepsilon), \quad \text{where } \tilde{d}=(\hat{c}_\omega l)^{-1}.$$

So, by substituting t/ε into t of (4.3) we obtain that for $2\leqq m\leqq 2q$,

$$E[(1+\tilde{d}\,|\zeta^{\varepsilon}(t)|^{2})^{m}]\leqq E[(1+\hat{\Psi}(\phi)+\hat{\Psi}(0))^{m}]\exp\left\{\int_{0}^{t}I_{m}(s)ds\right\}, \tag{4.4}$$

where $I_m(s)=I^{\varepsilon}_{m}(s)|_{\varepsilon=1}$.

Let $T<\infty$ be arbitrary and fixed. Then (4.4) yields

$$\sup_{0\leq t\leq T}E[|\zeta^{\varepsilon}(t)|^{4q}]<\infty \qquad \text{uniformly for} \quad 0<\varepsilon\leqq 1.$$

Namely the condition (0) of Theorem 3.1 is satisfied for $(4.1)^{\varepsilon}$. Since (4.2) satisfies Assumption 1.0 and Assumption 1.1 and since $E[\bar{\Psi}(\phi^{\omega})^{2q}]<\infty$, (4.2) has the pathwise unique solution $\bar{\zeta}(t)$ with the initial state $\bar{\zeta}(0)=\phi^{\omega}$. Thus the condition (I) of Theorem 3.1 is satisfied for (4.2). Here, by Definition 4.1 we note that the integer p given in Assumption 1.0 for (4.2) can be taken as the same integer with p given in Assumption 1.0 for (4.1). Eq $(4.1)^{\varepsilon}$ is a special case of $(2.1)^{\lambda}$ with $\lambda=\varepsilon$ where

$$b^{\varepsilon}[t,\zeta,\mu]=b_{0}\Big(\frac{t}{\varepsilon},\zeta\Big)-\Gamma_{0}\Big(\frac{t}{\varepsilon}\Big)\int_{R^{2}}(\zeta-\xi)\mu(d\xi),$$

$$\sigma^{\varepsilon}[t,\zeta,\mu]=\sigma_{0}\Big(\frac{t}{\varepsilon},\zeta\Big),\ t\geqq 0,\ \zeta\in R^{2}\ \text{and}\ \mu\in\mathcal{P}(R^{2}).$$

Eq (4.2) is a special case of $(2.1)^{\lambda_0}$ with $\lambda_0=0$ where

$$b^{0}[t,\zeta,\mu]=\bar{b}(\zeta)-\frac{1}{2}\bar{\Gamma}\int_{R^{2}}(\zeta-\xi)\mu(d\xi),$$

$$\sigma^{0}[t,\zeta,\mu]=\bar{\sigma}(\zeta),\ t\geqq 0,\ \zeta\in R^{2}\ \text{and}\ \mu\in\mathcal{P}(R^{2}).$$

Definition 4.1 implies that the condition (II) of Theorem 3.1 is satisfied for $(4.1)^{\varepsilon}$ and (4.2). Hence Theorem 3.1 applies for $(4.1)^{\varepsilon}$ and (4.2), and the proof is complete.

5. Oscillator with mean-field. Here we treat the oscillators strengthened by the fluctuation depending on the phase variable. Under suitable conditions on the coefficients, we can take the so-called energy function as a Lyapunov function satisfying Assumption 1.1.

Example 5.1. (Liénard oscillator). We consider a response of an oscillator

$$\ddot{x}+\varepsilon f(x)\dot{x}+g_{\varepsilon}(x)+\varepsilon\gamma(\dot{x}-E(\dot{x}))=\sqrt{\varepsilon}\,\delta(x)\dot{w}$$

to a (formal) white noise $\dot{w}$, where the dotted notation stands for the symbolic derivative d/dt and $E(\ \ )$ denotes the symbol of the mathematical expectation. Here and below ε is a small parameter such that $0<\varepsilon\leqq 1$, γ is a positive constant and $\{f(x), g_\varepsilon(x), \delta(x)\}$ is a family of scalar functions, for which

$f(x)$ and $\delta(x)$ satisfy the local Lipschitz condition in $x \in R^1$,
$g_\varepsilon(x)$ is an odd polynomial in $x \in R^1$ such that
$g_\varepsilon(x) = \omega^2 x + \varepsilon h(x)$, where ω is a positive constant and
$h(x) = \sum_{k=1}^{n} (2k+2)\alpha_{2k+2} x^{2k+1}$ with a family $\{\alpha_{2k+2}\}$ of positive constants.

Introduce the function $F(x) = \int_0^x f(s)ds$, and then take the Liénard plane (x, y), where $y = \dot{x} + \varepsilon F(x) + \varepsilon\gamma(x - E(x))$. Then we consider the solution $z(t) = (x(t), y(t))$ of the following stochastic differential equation;

$$
(5.1) \qquad \begin{aligned} dx(t) &= [y(t) - \varepsilon F(x(t)) - \varepsilon\gamma\{x(t) - E(x(t))\}]dt, \\ dy(t) &= -g_\varepsilon(x(t))dt + \sqrt{\varepsilon}\,\delta(x(t))dw(t), \end{aligned}
$$

where $w(t)$ is a one-dimensional Brownian motion process. This is a special case of SDE(4.1), where

$$
A = \begin{pmatrix} 0 & 1 \\ -\omega^2 & 0 \end{pmatrix}, \quad v(z) = \begin{pmatrix} -F(x) \\ -h(x) \end{pmatrix}, \quad z = \begin{pmatrix} x \\ y \end{pmatrix} \in R^2,
$$

$$
\Gamma = \begin{pmatrix} \gamma & 0 \\ 0 & 0 \end{pmatrix}, \quad D(z) = \begin{pmatrix} 0 & 0 \\ 0 & \delta(x) \end{pmatrix} \quad \text{and} \quad W(t) = \begin{pmatrix} w_0(t) \\ w(t) \end{pmatrix}
$$

is a two-dimensional Brownian motion process.

For $z = (x, y) \in R^2$ and $0 < \varepsilon \leqq 1$, set

$$
G_\varepsilon(x) = \int_0^x g_\varepsilon(s)ds \quad \text{and} \quad V_\varepsilon(z) = G_\varepsilon(x) + y^2/2,
$$

and also set

$$
g(x) = g_1(x) \quad \text{and} \quad V(z) = V_1(z).
$$

Then we give the following assumption.

Assumption 5.1. (L_0) $|F(x)| + |g(x)| \leqq c(1 + |x|^p)$ for all $x \in R^1$ with a constant $c > 0$ and an integer $p \geqq 0$,

(L_0') $|\delta(x)| \leqq \bar{\delta}$ for all $x \in R^1$ with a constant $\bar{\delta} > 0$,

(L_1) $-g_\varepsilon(x)F(x) \leqq \alpha(1 + G_\varepsilon(x))$ for all $x \in R^1$ with a constant $\alpha > 0$ being independent of ε,

(L_2) $-xF(x) \leqq \bar{\alpha}(1 + x^2/2)$ for all $x \in R^1$ with a constant $\bar{\alpha} > 0$.

The conditions (L_0) and (L_0') together with the local Lipschitz condition on the coefficients imply that (5.1) satisfies Assumption 1.0 with the family

$\{K_M, K, c, p\}$ of constants independent of ε. The conditions (L_0') and (L_1) imply that (5.1) satisfies Assumption 1.1 for $\Psi(t, z)=V_\varepsilon(z)$, $\beta(r)=r$ and $\{c_1, c_2, c_3\}$ replaced by $\{\varepsilon(\alpha+\bar{\delta}^2/2), \varepsilon(\alpha+\bar{\delta}^2), \varepsilon\gamma\}$. Further (L_0') and (L_2) imply that (4.2) derived from (5.1) satisfies Assumption 1.1 for $\Psi(t, z)=|z|^2/2$ and $\beta(r)=r$. Evidently, (L_0) and (L_0') imply that (4.2) derived from (5.1) satisfies Assumption 1.0 for the same integer p. We note that $V_\varepsilon(z)\leqq V(z)$ for $z\in R^2$ uniformly for $0<\varepsilon\leqq 1$. Let ϕ be any two-dimensional random vector independent of the Brownian motion process, such that

$$E[V(\phi)^{2q}]<\infty, \qquad q=\max\{2p, 2\}.$$

For $z\in R^2$, set $\bar{V}(z)=|z|^2/2$. For $\phi=(\phi^1, \phi^2)$, put $\phi^\omega=(\phi^1, \phi^2/\omega)$. Then, since $\omega^2\bar{V}(\phi^\omega)\leqq V(\phi)$, the assumption on ϕ implies

$$E[\bar{V}(\phi^\omega)^{2q}]<\infty.$$

Let $\zeta^\varepsilon(t)$ be the solution of $(4.1)^\varepsilon$ with the initial state $\zeta^\varepsilon(0)=\phi^\omega$ derived from (5.1). Then, in Theorem 4.1 we can take the functions $\Psi_\varepsilon(z)$, $\hat{\Psi}(z)$ and $\bar{\Psi}(z)$ by $\Psi_\varepsilon(z)=V_\varepsilon(z)$, $\hat{\Psi}(z)=V(z)$ and $\bar{\Psi}(z)=\bar{V}(z)$, respectively. Hence, under Assumption 5.1, we can apply Theorem 4.1 for $\zeta^\varepsilon(t)$ (for example see Narita [5]).

Example 5.2. (Quasiharmonic oscillator). We consider a response of an oscillator

$$\ddot{x}+\omega^2 x+\varepsilon\gamma(\dot{x}-E(\dot{x}))=\varepsilon f(x, \dot{x})+\sqrt{\varepsilon}\,\delta(x, \dot{x})\dot{w}$$

to a (formal) white noise $\dot{w}$, where ε is a small parameter such that $0<\varepsilon\leqq 1$, and ω and γ are positive constants. Here $f(x, y)$ and $\delta(x, y)$ are scalar functions satisfying the local Lipschitz condition in $(x, y)\in R^2$. Take the usual position and velocity variable; $y=\dot{x}$. Then we consider the solution $z(t)=(x(t), y(t))$ of the following stochastic differential equation;

$$(5.2)\qquad \begin{aligned} &dx(t)=y(t)dt, \\ &dy(t)=[-\omega^2 x(t)+\varepsilon f(x(t), y(t))-\varepsilon\gamma\{y(t)-E(y(t))\}]dt, \\ &\qquad +\sqrt{\varepsilon}\,\delta(x(t), y(t))dw(t), \end{aligned}$$

where $w(t)$ is a one-dimensional Brownian motion process. This is a special case of SDE(4.1), where

$$A=\begin{pmatrix} 0 & 1 \\ -\omega^2 & 0 \end{pmatrix}, \quad v(z)=\begin{pmatrix} 0 \\ f(z) \end{pmatrix}, \quad z=\begin{pmatrix} x \\ y \end{pmatrix}\in R^2,$$

$$\Gamma=\begin{pmatrix} 0 & 0 \\ 0 & \gamma \end{pmatrix}, \quad D(z)=\begin{pmatrix} 0 & 0 \\ 0 & \delta(z) \end{pmatrix} \text{ and } W(t)=\begin{pmatrix} w_0(t) \\ w(t) \end{pmatrix}.$$

is a two-dimensional Brownian motion process.

For $z=(x, y)\in R^2$, define $V(z)$ by $V(z)=(\omega^2x^2+y^2)/2$. Then we give the following assumption.

Assumption 5.2. (Q_0) $|f(z)|\leqq c(1+|z|^p)$ for all $z\in R^2$ with a constant $c>0$ and an integer $p\geqq 0$,

(Q_0') $|\delta(z)|\leqq\bar{\delta}$ for all $z\in R^2$ with a constant $\bar{\delta}>0$,

(Q_1) $yf(z)\leqq\alpha(1+V(z))$ for all $z\in R^2$ with a constant $\alpha>0$.

The conditions (Q_0) and (Q_0') together with the local Lipschitz condition on the coefficients imply that (5.2) satisfies Assumption 1.0 with the family $\{K_M, K, c, p\}$ of constants independent of ε. The conditions (Q_0') and (Q_1) imply that (5.2) satisfies Assumption 1.1 for $\Psi(t, z)=V(z)$, $\beta(r)=r$ and $\{c_1, c_2, c_3\}$ replaced by $\{\varepsilon(\alpha+\bar{\delta}^2/2), \varepsilon(\alpha+\gamma+\bar{\delta}^2), \varepsilon\gamma\}$. Further (Q_0') and (Q_1) imply that (4.2) derived from (5.2) satisfies Assumption 1.1 for $\Psi(t, z)=|z|^2/2$ and $\beta(r)=r$. Evidently, (Q_0) and (Q_0') imply that (4.2) derived from (5.2) satisfies Assumption 1.0 for the same integer p. Let ϕ be any two-dimensional random vector independent of the Brownian motion process, such that

$$E[V(\phi)^{2q}]<\infty, \qquad q=\max\{2p, 2\}.$$

For $z\in R^2$, set $\bar{V}(z)=|z|^2/2$. Then the condition on ϕ yields

$$E[\bar{V}(\phi^{\omega})^{2q}]<\infty .$$

Let $\zeta^{\varepsilon}(t)$ be the solution of $(4.1)^{\varepsilon}$ with the initial state $\zeta^{\varepsilon}(0)=\phi^{\omega}$ derived from (5.2). Then, in Theorem 4.1 we can take the functions $\Psi_{\varepsilon}(z)$, $\hat{\Psi}(z)$ and $\bar{\Psi}(z)$ by $\Psi_{\varepsilon}(z)=V(z)$, $\hat{\Psi}(z)=V(z)$ and $\bar{\Psi}(z)=\bar{V}(z)$, respectively. Hence, under Assumption 5.2, we can apply Theorem 4.1 for $\zeta^{\varepsilon}(t)$.

Example 5.3. For the van der Pol oscillator

$$\ddot{x}+\varepsilon\kappa(x^2-1)\dot{x}+\omega^2x+\varepsilon\gamma(\dot{x}-E(\dot{x}))=\sqrt{\varepsilon}\,\delta\dot{w}$$

with a family $\{\varepsilon, \kappa, \omega, \gamma, \delta\}$ of positive constants, we can get the same limit diffusion process governed by (4.2) as in Narita [4] whenever we start from the formulations (5.1) and (5.2).

References

[1] I.I. Gikhman and A.V. Skorokhod: *Stochastic Differential Equations*, Springer-Verlag, New York, 1972.

[2] R.Z. Hasminskii: Principle of averaging for parabolic and elliptic differential equations and for Markov processes with small diffusion, *Theory Probab. Appl.*, **8** (1963), 1-21.

[3] H. J. Kushner: *Approximation and Weak Convergence Methods for Random Processes, with Applications to Stochastic Systems Theory*, The MIT Press, Cambridge, Massachusetts, 1984.

[4] K. Narita: Asymptotic analysis for interactive oscillators of the van der Pol type, *Adv. in Appl. Probab.* **19** (1987), 44-80.

[5] K. Narita: Stochastic Liénard equation with mean-field interaction, *SIAM J. Appl. Math.* **49** (1989), 888-905.

[6] G.C. Papanicolaou, D. Stroock and S.R.S. Varadhan: Martingale Approach to Some Limit Theorems, *Proceedings of the* 1976 *Duke Univ. Conference on Turbulence*, 1977.

[7] Yu.V. Prokhorov: Convergence of random processes and limit theorems in probability theory, *Theory Probab. Appl.*, **1** (1956), 157-214.

[8] A.V. Skorokhod: Limit theorems for stochastic processes, *Theory Probab. Appl.*, **1** (1956), 261-290.

Department of Mathematics
Faculty of Technology
Kanagawa University
Rokkakubashi Kanagawa-ku
Yokohama 221, Japan.