Commun. Math. Phys. 166, 317–336 (1994)

Communications in Mathematical Physics

A Four-Thirds law for phase randomization of stochastically perturbed oscillators and related phenomena

Robert Cogburn, James A. Ellison★

Department of Mathematics and Statistics, The University of New Mexico, Albuquerque, New Mexico 87131, USA. Email: ellison@math.unm.edu

Received: 19 May 1993/in revised form: 11 March 1994

Abstract: Let I be a set of invariants and θ be a set of angle variables for a system of differential equations with an $O(\varepsilon)$ vector field. When time dependent stochastic perturbations, also of $O(\varepsilon)$, are added to the system, we have shown that under suitable conditions I becomes a stochastic adiabatic invariant satisfying a diffusion equation on time scales of order $1/\varepsilon^2$, in the limit as $\varepsilon \to 0$. Here we show that the angle variables converge weakly to a Gaussian Markov process on an $O(\varepsilon^{-4/3})$ time scale, and thus the phase becomes randomized at these times. Application to nearly integrable Hamiltonian systems is considered.

0. Introduction

We consider the behavior of the stochastic differential equation in $\mathbb{R}^d$,

$$\dot{x} = \varepsilon f(x,t) + \varepsilon F(x,t,\omega) + o(\varepsilon^{5/3}) \tag{0.1}$$

as $\varepsilon \to 0$. We require that the expectation $EF(x,t) = 0$ and that the time average

$$\bar{f}(x) = \lim_{t \to \infty} \frac{1}{t} \int_0^t f(x,s)\,ds \tag{0.2}$$

exists for all x.

Making the change of scale $v = \varepsilon t$, (0.1) becomes

$$\frac{dx}{dv} = f\left(x, \frac{v}{\varepsilon}\right) + F\left(x, \frac{v}{\varepsilon}, \omega\right) + o(\varepsilon^{2/3})\,. \tag{0.3}$$

Then (0.2) and the law of large numbers applied to F suggest that the method of averaging may apply to (0.3), and for small ε the solution should be close to the "unperturbed equation"

$$\frac{dx}{dv} = \bar{f}(x)\,. \tag{0.4}$$

★ Supported by NSF grant DMR-8704348

In fact Khas'minskii [9] established that this was the case under suitable regularity conditions. Now suppose $y = I(x)$ taking values in $\mathbb{R}^p$ is an invariant for the unperturbed system:

$$\frac{\partial I}{\partial x}(x)\bar{f}(x) = 0 . \tag{0.5}$$

Letting X_ε denote a solution to (0.1) and $Y_\varepsilon(t) = I(X_\varepsilon(t))$, where I is a smooth function, Y_ε solves

$$\frac{dY_\varepsilon}{dt} = \varepsilon\frac{\partial I}{\partial x}(X_\varepsilon(t))f(X_\varepsilon(t),t) + \varepsilon G(X_\varepsilon(t),t,\omega) + o(\varepsilon^{5/3}) , \tag{0.6}$$

where $G = \left(\frac{\partial I}{\partial x}\right)F$. In [3] we show that, provided the unperturbed system (0.4) is ergodic on the surfaces $I(x) = \textit{constant}$ and certain regularity conditions apply, the $Y_\varepsilon(t)$ processes converge weakly to a diffusion on $O(\varepsilon^{-2})$ time scales as $\varepsilon \to 0$.

In this study we suppose that, in addition to $X_\varepsilon(t)$ and $Y_\varepsilon(t)$, there is a third process $Z_\varepsilon(t)$ with values in $\mathbb{R}^q$ solving

$$\frac{dZ_\varepsilon}{dt} = \varepsilon\{v(Y_\varepsilon(t)) + h(X_\varepsilon(t),t) + H(X_\varepsilon(t),t,\omega)\} + o(\varepsilon^{4/3}) , \tag{0.7}$$

where $EH(x,t) = 0$ and h has time average 0.

The corresponding unperturbed system for Z_ε is

$$\frac{dz}{dt} = \varepsilon v(y) , \tag{0.8}$$

where y is constant, hence the unperturbed $z(t)$ is a linear function of t.

The most obvious example fitting this description is an oscillator in phase space x, where y is the energy and z is the phase position. Or alternatively, y and z may be the "action" and "angle" variables of the system. This problem will be discussed in Sect. 2. In such cases there is a constant vector ζ whose i^{th} coordinate ζ_i is the period of the i^{th} coordinate z_i of z. Taking $z(\operatorname{mod}\zeta)$ to be the vector whose i^{th} coordinate is $z_i(\operatorname{mod}\zeta_i)$, from one point of view we should have $z(\operatorname{mod}\zeta) = \Phi(x)$ for some function $\Phi\colon \mathbb{R}^d \to \mathbb{R}^q$, however this implies discontinuities in the Φ function, so in the usual way we regard $\Phi(x)$ to be a multivalued function, the branch in effect at any given time being determined by continuity.

Another possibility would be where x is a laminar flow with shear, y indexes the layer of the unperturbed flow and z is the distance traveled in the layer. In both of these examples, using a suitable interpretation, we have $z = \Phi(x)$ and the unperturbed orbital derivative of Φ is

$$\frac{\partial \Phi}{\partial x}\bar{f}(x) = v(y) , \tag{0.9}$$

a function of y, and (0.7) becomes

$$\frac{dZ_\varepsilon}{dt} = \varepsilon\left\{v(Y_\varepsilon(t)) + \frac{\partial\Phi}{\partial x}(X_\varepsilon(t))((f(X_\varepsilon(t),t) - \bar{f}(X_\varepsilon(t)) + F(X_\varepsilon(t),t,\omega))\right\}$$
$$+ O(\varepsilon^{4/3}) . \tag{0.10}$$

Let x_0 be the initial value for the X_ε process and $y_0 = I(x_0)$ and z_0 be corresponding initial values for the Y_ε and Z_ε processes. We will show that in the scaled time $\tau = \varepsilon^{4/3}t$, the process

$$W_\varepsilon(\tau) = Z_\varepsilon(\tau\varepsilon^{-4/3}) - z_0 - \tau\varepsilon^{-1/3}\nu(y_0)\,, \tag{0.11}$$

representing the deviation of Z_ε from the unperturbed solution, converges weakly to a Gaussian Markov process as $\varepsilon \to 0$, under suitable conditions. From this we conclude, under a nondegeneracy condition, that $Z_\varepsilon(\bmod\,\zeta)$ becomes uniform on an $\varepsilon^{-4/3}$ time scale.

The following heuristic argument motivates the $\varepsilon^{-4/3}$ time scale. In a one degree of freedom nonlinear oscillator a point x_0 moves roughly on its phase plane oval with variance in action increasing like $E(J(t) - J_0)^2 = \bar{\bar{\mathfrak{F}}}(J_0)\varepsilon^2 t$. The earliest time to expect uniformity on a thin energy shell containing x_0 would be when adjacent points on J_0 and $J(t)$ have separated by one revolution, i.e.,

$$(\nu(J(t)) - \nu(J_0))\varepsilon t = 2\pi\,.$$

Thus $\nu'(J_0)(\bar{\bar{\mathfrak{F}}}(J_0)\varepsilon^2 t)^{1/2}\varepsilon t \simeq 2\pi$, which gives

$$t \simeq \left(\frac{2\pi}{\nu'(J_0)\sqrt{\bar{\bar{\mathfrak{F}}}(J_0)}}\right)^{2/3}\varepsilon^{-4/3}\,.$$

In Sect. 1 we introduce notation, discuss the assumptions and state the main results, and in Sect. 2 we discuss several examples that illustrate the theorems. Section 3 contains estimates, auxiliary results and proofs of the theorems.

1. Formulation of the Main Results

As in the introduction, the invariant $I\colon \mathbb{R}^d \to \mathbb{R}^p$. Since $y = I(x)$ varies only by small amounts on the time scales to be considered, it suffices to consider any open neighborhood D_1 of the initial value $y_0 = I(x_0)$ in $\mathbb{R}^p$. We let D_0 be the largest connected set in $I^{-1}(D_1)$ such that $x_0 \in D_0$.

In what follows $(\Omega, \mathscr{F}, P)$ is a probability space, and for each $x \in D_0$ and $t \geqq 0$, $F(x,t) = F(x,t,\omega)$ is an $\mathbb{R}^p$ valued random variable on Ω.

For an $m \times n$ vector or matrix $M = (M_{j,k})$, let $|M| = \sum_{j=1}^{m}\sum_{k=1}^{n}|M_{j,k}|$. When $M(x,s,t,\omega)$ is a vector or matrix valued function of $x_0 \in D_0$, $s \geqq 0$, $t \geqq 0$ and $\omega \in \Omega$, let

$$\|M\| = P - \operatorname*{ess\,sup}_{\omega}\ \sup_{x\in D_0}\ \sup_{s,t\geqq 0}|M(x,s,t,\omega)|\,,$$

$$\|M\|_1 = P - \operatorname*{ess\,sup}_{\omega}\ \sup_{\substack{x_1,x_2\in D_0\\ x\neq x_2}}\ \sup_{s,t\geqq 0}|M(x_1,s,t,\omega) - M(x_2,s,t,\omega)|/|x_1 - x_2| + \|M\|\,,$$

$$\|M\|_2 = \max_{1\leqq k\leqq d}\|\partial M/\partial x_k\|_1 + \|M\|\,,$$

and use the same conventions when M depends on a subset of these arguments. Note that

$$\|M\| \leqq \|M\|_1 \leqq \|M\|_2\,.$$

Now consider the assumptions:

(A1) Writing (0.1) as the initial value problem

$$\dot{x} = \varepsilon f(x,t) + \varepsilon F(x,t,\omega) + R(x,t,\omega,\varepsilon),\ x(0) = x_0\ , \tag{1.1}$$

we require F and R to be continuous in x and t, R to be locally x-Lipschitz and $\|F\|_1 < \infty$. We also require $EF(x,t) = 0$ for all x, t. All order statements for the limit as $\varepsilon \to 0$ are to be understood in terms of the first norm defined above, so the $o(\varepsilon^{5/3})$ quantity in (0.1) means $\|R\| = o(\varepsilon^{5/3})$.

We require $f(x,t)$ to be almost periodic, with the Fourier representation $f(x,t) = \sum a_k(x)e^{i\lambda_k t}$. Hence the λ_k are distnct, $\lambda_{-k} = -\lambda_k$ and $a_{-k} = a_k^*$ (the conjugate of a_k). Note that $a_0(x) = \bar{f}(x)$ as defined by (0.2). We assume

$$\sum \|a_k\| < \infty,\ \sum_{k \neq 0} \|a_k\|_1/|\lambda_k| < \infty\ \underline{\text{and}}\ \sum_{k \neq 0} \|a_k\|_2/|\lambda_k|^{1+\theta} < \infty \tag{1.2}$$

for some $0 < \theta \leqq 1$.

(A2) The mapping I must satisfy (0.5) and we require that $\|\frac{\partial I}{\partial x}\|_2 < \infty$. Note that this assumption and (A1) imply that $\|G\|_1 < \infty$.

(A3) The function $v\colon D_1 \to \mathbb{R}^q$ must satisfy the global Lipschitz condition

$$\left|\frac{\partial v}{\partial y}(y_2) - \frac{\partial v}{\partial y}(y_1)\right| \leqq C|y_2 - y_1|\ . \tag{1.3}$$

The equation (0.7) in z must have continuous, locally x-Lipschitz right-hand side and we require h to satisfy

$$\sup_{x,t}\left|\int_0^t h(x,s)\,ds\right| < \infty\ , \tag{1.4}$$

$\|h\|_1 < \infty$, $EH(x,t) = 0$ and $\|H\|_1 < \infty$.

(A4) For $0 \leqq s \leqq t \leqq \infty$ let $\mathscr{F}_s^t$ be sub-σ-fields of $\mathscr{F}$ such that for $t_1 \leqq t_2 \leqq t_3 \leqq t_4$, $\mathscr{F}_{t_2}^{t_3} \subset \mathscr{F}_{t_1}^{t_4}$ and such that $F(x,t)$ and $H(x,t)$ are $\mathscr{F}_t^t$ measurable for all $x \in D_0$ and $t \geqq 0$. If the initial state x_0 is random, then we also require x_0 to be $\mathscr{F}_0^0$ measurable. Let

$$\rho(t) = \sup_{s \geqq 0}\ \sup_{A \in \mathscr{F}_0^s : PA > 0}\ \sup_{B \in \mathscr{F}_{s+t}^\infty} |P(B\,|\,A) - PB|\ .$$

We require the mixing condition

$$\rho(t) = o\left(\frac{1}{t^2}\right) \tag{1.5}$$

as $t \to \infty$.

Before stating the next assumption we need some notation. Let

$$\Gamma(x,s,t) := E[F(x,s)F(x,t)^T] \tag{1.6}$$

and

$$\hat{\Gamma}(x,s) := \int_s^\infty (\Gamma(x,s,t) + \Gamma(x,t,s))\,dt\ . \tag{1.7}$$

An application of (A4) (see Lemma 3.1) shows that, for each fixed s and t, $\|\Gamma(x,s,t)\|_1 \leqq C\rho(|t - s|)$, and it follows easily that $\hat{\Gamma}$ is well defined and $\|\hat{\Gamma}\|_1 < \infty$. Let

$$\hat{\mathfrak{A}}(x,t) := \frac{\partial I}{\partial x}(x)\hat{\Gamma}(x,t)\frac{\partial I}{\partial x}(x)^T . \tag{18}$$

(A5) We assume there exists a function $\mathfrak{A}(x)$ such that

$$\sup_{x \in D_0, t \geq 0} \left| \frac{1}{l} \int_t^{t+l} \hat{\mathfrak{A}}(x,s)\,ds - \mathfrak{A}(x) \right| \to 0 \tag{1.9}$$

as $l \to \infty$. Note $\|\hat{\Gamma}\|_1 < \infty$ and (A2) imply $\|\hat{\mathfrak{A}}\|_1 < \infty$ and $\|\mathfrak{A}\|_1 < \infty$.

(A6) Let $\bar{x}(t, x_0)$ be the solution to the initial value problem

$$\dot{x} = \bar{f}(x), \quad x(0) = x_0 .$$

Assume there exists a continuous function $\bar{\mathfrak{A}}(y)$ of y on D_1 such that

$$\bar{\mathfrak{A}}(I(x)) = \lim_{l \to \infty} \frac{1}{l} \int_0^t \mathfrak{A}(\bar{x}(t,x))\,dt \tag{1.10}$$

exists uniformly in x in D_0.

(A6′) Assume there is a matrix M_0 such that

$$M_0 = \lim_{l \to \infty} \frac{1}{l} \int_0^l \mathfrak{A}(\bar{x}(t,x))\,dt \tag{1.11}$$

exists uniformly for $x \in I^{-1}\{y_0\}$, where y_0 is the initial value of the Y_ε process, and assume that for any sequence $x_n \in D_0$ such that $I(x_n) \to y_0$, there exist $\hat{x}_n \in D_0$ such that $I(\hat{x}_n) = y_0$ and $|x_n - \hat{x}_n| \to 0$. Note that under (A6), $M_0 = \bar{\mathfrak{A}}(y_0)$.

Our first result describes the behavior of the $Y_\varepsilon(t)$ processes on $\varepsilon^{-4/3}$ time scales. Let $C_k[0, \infty)$ denote the space of continuous functions on $[0, \infty)$ to $\mathbb{R}^k$ with supremum norm.

Let

$$V_\varepsilon(\tau) := \varepsilon^{-1/3}(Y_\varepsilon(\varepsilon^{-4/3}\tau) - y_0) . \tag{1.12}$$

Theorem 1.1. *If assumptions* (A1) *through* (A4) *hold then the processes* $\{V_\varepsilon(\tau)\}_{\tau \geq 0}$ *are relatively compact in* $C_p[0, \infty)$ *as* $\varepsilon \to 0$. *If, moreover,* (A5) and either (A6) *or* (A6′) *hold, then as* $\varepsilon \to 0$, $V_\varepsilon(\tau)$ *converges weakly to the Wiener process* $V_0(\tau)$ *having the representation*

$$V_0(\tau) = (\bar{\mathfrak{A}}(y_0))^{1/2} B(\tau) , \tag{1.13}$$

where $B(\tau)$ *is standard p-dimensional Brownian motion.*

Remark. In [3] we show that $Y_\varepsilon(\tau/\varepsilon^2)$ converges weakly to $Y_0(\tau)$, where $dY_0 = \bar{\mu}(Y_0)\,d\tau + \bar{\mathfrak{A}}(Y_0)^{1/2}\,dB(\tau)$. An heuristic perturbation argument at the $\varepsilon^{-4/3}$ time scale is consistent with (1.13).

Note. Here and in what follows, for a symmetric positive semidefinite (psd) matrix M, $(M)^{1/2}$ denotes its psd square root. Of course, the matrices $\hat{\Gamma}$, $\hat{\mathfrak{A}}$, $\mathfrak{A}$, $\bar{\mathfrak{A}}$ are psd.

Theorem 1.2. *If assumptions* (A1) *through* (A4) *hold then the processes* $\{W_\varepsilon(\tau)\}_{\tau \geqq 0}$ *defined in* (0.11) *are relatively compact in* $C_q[0, \infty)$ *as* $\varepsilon \to 0$. *If, moreover,* (A5) *and either* (A6) *or* (A6′) *hold, then as* $\varepsilon \to 0$, $W_\varepsilon(\tau)$ *converges weakly to the Gaussian Markov process* $W_0(\tau) = \frac{\partial v}{\partial y}(y_0)(\bar{\mathscr{Z}}(y_0))^{1/2} \int_0^\tau B(s)\,ds$, *where* $B(s)$ *is standard p-dimensional Brownian motion. This q-dimensional process has continuous sample functions, zero mean and covariance for* $0 \leqq \tau_1 \leqq \tau_2$,

$$EW_0(\tau_1)\,W_0(\tau_2)^T = \frac{\tau_1^2}{6}(3\tau_2 - \tau_1)\frac{\partial v}{\partial y}(y_0)\bar{\mathscr{Z}}(y_0)\frac{\partial v}{\partial y}(y_0)^T . \tag{1.14}$$

When $\frac{\partial v}{\partial y}(y_0)\bar{\mathscr{Z}}(y_0)\frac{\partial v}{\partial y}(y_0)^T$ is positive definite, it follows that $W_0(\tau)$ has a Gaussian distribution with large dispersion when τ is large, and this implies that $Z_\varepsilon(\tau\varepsilon^{-4/3})(\mathrm{mod}\,\zeta)$ is approximately uniform. In fact this "phase randomization" applies even without the mean stationarity and ergodic assumptions of (A5) and (A6) provided a minimal amount of stochastic perturbation is present, and this is the content of our third result.

For psd matrices A, B we write $A \geqq B$ when $A - B$ is psd. We need:

(A7) Assume there is a $q \times q$ psd matrix function $\Delta(x)$, $x \in D_0$, with $\|\Delta\|_1 < \infty$ and that there is a constant positive definite matrix Δ_0 and finite T_1 and T_2 such that

$$\int_t^{t+T_1} \frac{\partial v}{\partial y}(y_0)\hat{\mathscr{Z}}(x,s)\frac{\partial v}{\partial y}(y_0)^T\,ds \geqq \Delta(x) \tag{1.15}$$

for all $t \geqq 0$, and

$$\int_0^{T_2} \Delta(\bar{x}(t,x))\,dt \geqq \Delta_0 \tag{1.16}$$

for all $x \in D_0$. Note that the time scales in (1.15) and (1.16) are different.

Theorem 1.3. *If assumptions* (A1) *through* (A4) *and* (A7) *hold, then in the iterated limit as* $\varepsilon \to 0$ *followed by* $\tau \to \infty$, $Z_\varepsilon(\tau\varepsilon^{-4/3})(\mathrm{mod}\,\zeta)$ *converges in distribution to a uniform distribution on the rectangle* $\{z : 0 \leqq z_k < \zeta_k,\ k = 1, \ldots, q\}$, *where* ζ *is an arbitrary element of* $\mathbb{R}^q$ *such that* $\zeta_k > 0$, $k = 1, \ldots, q$.

Remarks. 1. The three theorems apply to the situation that $X_\varepsilon(0) = x_0$ fixed. If $X_\varepsilon(0) = X_0$ is a random variable and is $\mathscr{F}_0^0$ measurable, then Theorems 1.1 and 1.2 provide the conditional distributions of V_0 and W_0 given $X_0 = x_0$, provided the appropriate regularity conditions hold for the given value of x_0, since (A4) implies asymptotic independence of $\{X_\varepsilon(t), t \geqq \delta\}$ from X_0 as $\varepsilon \to 0$ for each fixed $\delta > 0$. Marginal distributions for the V_0 and W_0 processes are then obtained by integrating the conditional distributions with respect to the distribution of $X(0)$. In similar fashion Theorem 1.3 implies that the limit distribution of $Z_\varepsilon(\tau\varepsilon^{-4/3})(\mathrm{mod}\,\zeta)$ is still uniform.
2. Theorem 1.3 implies that an integrable Hamiltonian system with minimal stochastic perturbation (condition (A7)) asymptotically has a kind of ergodic averaging on constant energy surfaces, with this averaging taking place on time scales of order $\varepsilon^{-4/3}$. This result allows the extension of the adiabatic invariance results of [3] to this case under (A7), whether the unperturbed system is ergodic or not. We plan to provide the details in a subsequent paper.

2. Examples

In this section we discuss examples which illustrate the theory.

Example A. In this example we will consider perturbations of the one degree of freedom nonlinear oscillator defined by the Hamiltonian

$$H_0(x) := \frac{1}{2}x_2^2 + U(x_1), \tag{2.1}$$

where $U(x_1)$ is a symmetric bowl type potential so that all solutions are periodic. We write our example as

$$\frac{dx}{dv} = \bar{f}(x) + p\left(x, \frac{v}{\varepsilon}\right) + F\left(x, \frac{v}{\varepsilon}, \omega\right) \quad x(0) = x_0, \tag{2.2}$$

where $\bar{f}(x) = \begin{pmatrix} x_2 \\ -U'(x_1) \end{pmatrix}$, $p(x,t)$ has zero t-mean and $EF = 0$. If we let $\bar{x}(v, x_0)$ denote the solution of the unperturbed problem, then the transformation, $(x_1, x_2) \to (\theta, I)$, to the action angle variables of the unperturbed problem can be written

$$x = \bar{x}(\theta/\nu(I), \zeta(I)),$$

where $\nu(I)$ is the frequency as a function of action associated with H_0 and $\zeta(I)$ is an appropriately chosen initial condition, on the closed integral curve for H_0 associated with I. If we let $y(t) = I(x(t))$ and $z(t) = \theta(x(t))$, then

$$\dot{y} = \varepsilon I'(x)[p(x,t) + F(x,t,\omega)], \quad y(0) = y_0 = I(x_0),$$
$$\dot{z} = \varepsilon[\nu(y) + D\theta(x)\{p(x,t) + F(x,t,\omega)\}], \quad z(0) = z_0 = \theta(x_0). \tag{2.3}$$

We let D_1 be an open interval about $y_0 = I(x_0)$ and thus $D_0 = I^{-1}(D_1)$ is an energy shell about the initial energy oval, $\{x\colon H_0(x) = H_0(x_0)\}$. We assume the smoothness and almost periodicity conditions of (A1). For U smooth, (A2) and (1.3) of (A3) are satisfied. To apply the theorems, we need to calculate $\mathcal{Y}(x,t)$ and for illustration we assume

$$F(x,t,\omega) = Q(x)\xi(t,\omega), \tag{2.4}$$

where ξ is a scalar such that $E\xi = 0$ and that (A3) is satisfied. We define $\mathscr{F}_s^t = \sigma(\xi(\tau), t \leqq \tau \leqq s)$ and assume the mixing condition of (A4) is satisfied. Now recall that

$$V_\varepsilon(\tau) := \varepsilon^{-1/3}(Y_\varepsilon(\varepsilon^{-4/3}\tau) - y_0),$$
$$W_\varepsilon(\tau) := z_\varepsilon(\tau\varepsilon^{4/3}) - z_0 - \tau\varepsilon^{-1/3}\nu(y_0),$$

and at this point we know from Theorems 1.1 and 1.2 that $\{V_\varepsilon\}$ and $\{W_\varepsilon\}$ are relatively compact in their C-spaces. Condition (A6) is satisfied, so to proceed we have two options, depending on whether (A5) is satisfied or not. We look first at (A5).

Let $K(s,t) = E(\xi(s)\xi(t))$, then we have $|K(s+t,s| \leqq 2\|\xi(t)\|^2\rho(t) = 0(1/t^2)$, where the inequality follows from Proposition 2.2, p. 346 of Ref. [6] and the equality is as $t \to \infty$ and follows from (A4).

Clearly

$$\Gamma(x,s,t) = Q(x)Q^T(x)K(s,t)\,,$$

and since $K(s,t) = K(t,s)$,

$$\hat{\Sigma}(x,t) = 2I'(x)Q(x)Q^T(x)I'(x)\int_t^{\infty} K(\tau,t)\,d\tau\,,$$

which clearly exists because of our mixing condition. If

$$\lim_{l\to\infty}\frac{1}{l}\int_t^{t+l}\int_s^{\infty} K(\tau,s)\,d\tau\,ds$$

converges uniformly to a t-independent limit $\frac{1}{2}C$, then (A5) is satisfied with $\Sigma(x) = CI'(x)Q(x)Q^T(x)I'(x)^T$. A sufficient condition for this is the stationary of $\xi(t)$ in which case $C = 2\int_0^{\infty} K(s,0)\,ds$. (A6) is automatically satisfied because H_0 is ergodic and

$$\bar{\Sigma}(y) = 2\pi v(y)\frac{1}{\sqrt{2}}\int_{-a}^{a}[\Sigma(x_1,\sqrt{2}\sqrt{h-U(x_1)})$$

$$+\,\Sigma(-x_1,-\sqrt{2}\sqrt{h-U(x_1)})]\frac{dx_1}{\sqrt{h-U(x_1)}}\,, \qquad (2.5)$$

where $h = h(y)$ and $a = a(y)$ are the energy and oscillation amplitude as a function of action, respectively, for the unperturbed motion.

Theorems 1.1 and 1.2 now yield

$$V_\varepsilon \Rightarrow V_0\,, \quad \text{where } V_0(\tau) = \bar{\Sigma}(y_0)^{1/2}B(\tau)\,,$$

where $B(\tau)$ is standard Brownian motion, and

$$W_\varepsilon \Rightarrow W_0\,, \quad \text{where } W_0(\tau) = v'(y_0)^{1/2}\bar{\Sigma}(y_0)^{1/2}\int_0^{\tau} B(s)\,ds\,.$$

Thus, W_0 is Gauss–Markov with zero-mean and covariance given by (1.14).

Thus we obtain the phase randomization when $\bar{\Sigma}(y_0) > 0$, which requires noise, and $v'(y_0) \neq 0$, which requires H_0 to be a nonlinear oscillator. This type of phase randomization can also occur, in a coarse grained sense, without noise when the initial condition y_0 is not concentrated at a point [6].

In this case condition (A6) is automatic and (A5) holds, so there is no need to resort to condition (A7); nevertheless we discuss this condition to illustrate its use. Example B will illustrate its power.

Now assume there exists a T_1 such that

$$\inf_{t\geq 0}\int_0^{T_1}\int_0^{\infty} K(\tau+z+t, z+t)\,d\tau\,dz \geqq \alpha > 0\,,$$

then we can choose

$$\Delta(x) = 2\alpha I'(x)Q(x)Q(x)^T I'(x)^T = 2\alpha v(y)^{-2}(U'(x_1)Q_1 + x_2Q_2)^2 \geqq 0\,.$$

The second assumption in (A7) then amounts to $f(y) = \max_{H_0(x)=h(y)}|U'(x_1)Q_1(x) + x_2Q_2(x)|$ bounded away from zero on D_1 which is hardly any restriction at all. Theorem 1.3 then yields the phase randomization result.

Example B. Let $x = \binom{y}{z}$, where $y \in D_1 \subset \mathbb{R}^p$, $z \in \mathbb{R}^q$ and $d = p + q$. Consider the IVP

$$\frac{dy}{dt} = \varepsilon(g(x,t) + G(x,t,\omega)) \quad y(0) = y_0 ,$$

$$\frac{dz}{dt} = \varepsilon(v(y) + h(x,t) + H(x,t,\omega)) \quad x(0) = x_0 , \tag{2.6}$$

where D_1 is an open neighborhood of y_0 and all functions are 2π periodic in z, g and h have zero t-mean and $EG = EH = 0$. These equations fit our general framework with $f(x,t) = \begin{pmatrix} g(x,t) \\ v(y) + h(x,t) \end{pmatrix}$, $\bar{f}(x) = \begin{pmatrix} 0 \\ v(y) \end{pmatrix}$, $F = \begin{pmatrix} G \\ H \end{pmatrix}$, $R = 0$ and the invariant

$$I(x) = (x_1, \ldots, x_p)^T . \tag{2.7}$$

We assume the smoothness and almost periodicity conditions of (A1) and (A3). The fact that $D_0 = D_1 \times \mathbb{R}^q$ is unbounded is not a problem since all functions are 2π periodic in z. (A2) is trivially satisfied and we assume the measureability and mixing conditions of (A4). It is easy to check that

$$\hat{\mathscr{Y}}(x,t) = \int_s^\infty E(G(x,s)G^T(x,t))^Q \, dt , \tag{2.8}$$

and thus we see that our results are independent of whether $H = 0$ or not. That is, it is the noise in y that moves the system away from y_0 and allows the phase randomization due to the $v(y)$ term. In fact g and h do not affect the result either. If we assume that G is stationary with $E(G(x,s)G^T(x,t)) =: C(x,s-t)$, then

$$\hat{\mathscr{Y}}(x,t) = \int_{-\infty}^{\infty} C(x,s)\,ds =: \mathscr{Y}(x) , \tag{2.9}$$

and (A5) is satisfied.

In this example $\bar{x}$ is particularly simple,

$$\bar{x}(t,x_0) = \begin{pmatrix} y_0 \\ v(y_0)t + z_0 \end{pmatrix},$$

and thus $\mathscr{Y}(\bar{x}(t,x_0))$ is quasi-periodic in t since $\mathscr{Y}(x)$ is a 2π periodic function of z. For $q \geqq 2$, (A6) is too restrictive since it will not be satisfied unless v is a constant with rationally independent components, in whic case $v' = 0$ and there will be no phase randomization, however (A6′) may be satisfied. Now $I^{-1}(y_0) = \{\binom{y_0}{z} | z \in \mathbb{R}^q|$ and the limit in (1.11) will exist uniformly in $I^{-1}(y_0)$ if the components of $v(y_0)$ are rationally independent (i.e. the rate of ergodization on the q-torus is independent of initial position on the torus). Let $x_n = \binom{y_n}{z_n}$ with $y_n \to y_0$ and $\hat{x}_n = \binom{y_0}{z_n}$, then $I(\hat{x}_n) = y_0$ and $|x_n - \hat{x}_n| = |y_n - y_0| \to 0$. Thus (A6′) is satisfied for y_0 such that the components of $v(y_0)$ are rationally independent and we can apply Theorems (1.1) and (1.2) to obtain the weak convergence of $\{V_\varepsilon\}$ and $\{W_\varepsilon\}$. The phase randomization follows if $v'(y_0)\bar{\mathscr{Y}}(y_0)v'(y_0)^T$ is positive definite (pd) and this is true if $v'(y_0)\mathscr{Y}(y_0)v'(y_0)^T$ is pd. The latter is true if $\mathscr{Y}(x)$ is positive definite, and the columns of $v'(y_0)^T$ are linearly independent (which requires $q \leqq p$).

However, since v is non-constant and continuous we generally expect y's arbitrarily close to y_0 such that the components of $v(y)$ are rationally dependent and yet it seems that the phase randomization should not be so sensitive to y_0 particularly in the presence of noise. If we take $T_1 = 1$ and $\Delta(x) = v'(y_0)\Sigma(x_0)v'(y_0)^T$ and assume as before that this is positive definite, then the left-hand side of (1.16) is positive definite for all $T_2 > 0$ and Theorem 1.3 entails the phase randomization without (A5) and (A6′).

Example C. Here we simply point out an important special case of Example B, namely a perturbation of an integral Hamiltonian system, H_0,

$$H(J,\theta,t) = \varepsilon(H_0(J) + H_d(J,\theta,t) + H_r(J,\theta,t,\omega)),$$

where $J \in \mathbb{R}^n$, $\theta \in T^n$, H_d has zero time mean and H_r zero expected value. The Hamiltonian equations of motion are now in the form of (2.6) with $y = J$, $z = \theta$, $\dot{y} = -\frac{\partial H}{\partial \theta}$ and $\dot{z} = \frac{\partial H}{\partial J}$.

3. Preliminary Estimates

We assume (A1) *through* (A4) *hold in all that follows. When* (A5), (A6), (A6′) *or* (A7) *is used, this will be explicitly stated.*

The following standard mixing result (e.g., see [10]) is needed:

Lemma 3.1. *Let $\Xi(x,\omega)$ be an $\mathscr{F}^{\infty}_{s+t}$ measurable random variable with values in $\mathbb{R}^p$ and be Borel measurable in x for each ω. Let $\|\Xi\| < \infty$ and let $\xi(x) = E\Xi(x)$. Then for any $\mathscr{F}^s_0$ measurable random variable Z with values in $\mathbb{R}^p$,*

$$\|E^{(s)}\Xi(Z) - \xi(Z)\| \leqq 2\|\Xi\|\rho(t). \tag{3.1}$$

In addition we need the following result (see Proposition 3.1 of [3]) where it is stated for a different time scaling):

Lemma 3.2. *For each $\varepsilon_0 > 0$ there is a $C < \infty$ such that for all $0 < \varepsilon \leqq \varepsilon_0$ and $t \geqq 0$,*

$$\left\|\varepsilon\int_0^t \frac{\partial I}{\partial x}(X_\varepsilon(s))f(X_\varepsilon(s),s)\,ds\right\| \leqq C\varepsilon(\varepsilon t + 1). \tag{3.2}$$

Lemma 3.3. *For $t \geqq \varepsilon^{-1/2}$ there is a $C < \infty$ such that*

$$\left\|\int_0^t h(X_\varepsilon(s),s)\,ds\right\| \leqq Ct\varepsilon^{1/2}. \tag{3.3}$$

Proof. By (A3), $\|h\|_1 < \infty$, and by (A1),

$$|X_\varepsilon(r) - X_\varepsilon(s)| \leqq C_0\varepsilon|r-s| \tag{3.4}$$

for some $C_0 < \infty$. Hence

$$\int_0^t h(X_\varepsilon(s),s)\,ds = \int_{\varepsilon^{-1/2}}^t \varepsilon^{1/2}\int_{s-\varepsilon^{-1/2}}^s h(X_\varepsilon(r),s)\,dr\,ds + tO(\varepsilon^{1/2})$$

$$= \varepsilon^{1/2}\int_0^t \int_{\max\{r,\varepsilon^{-1/2}\}}^{\min\{r+\varepsilon^{-1/2},t\}} h(X_\varepsilon(r),s)\,ds\,dr + tO(\varepsilon^{1/2}) = tO(\varepsilon^{1/2})$$

by (1.4). □

Now let $L = L(\varepsilon)$ be chosen so $L(\varepsilon) = o(\varepsilon^{-1/3})$ and $\rho(L(\varepsilon)) = o(\varepsilon^{2/3})$. This is possible since $\rho(t) = o(t^{-2})$ by (A4).

Lemma 3.4. *For* $0 \leqq t_1 < t_2 \leqq \tau\varepsilon^{-4/3}$ *and each fixed* τ, *as* $\varepsilon \to 0$,

$$Y_\varepsilon(t_2) - Y_\varepsilon(t_1) = \varepsilon \int_{t_1}^{t_2} G(X_\varepsilon(t), t + L)\,dt + o(\varepsilon^{1/3}) \tag{3.5}$$

and

$$Z_\varepsilon(t_2) - Z_\varepsilon(t_1) = \varepsilon \int_{t_1}^{t_2} \nu(Y_\varepsilon(t))dt + \varepsilon \int_{t_1}^{t_2} H(X_\varepsilon(t), t + L)\,dt + o(1)\,. \tag{3.6}$$

Proof. Taking Lemma 3.2 into account, we have from (0.6) that

$$Y_\varepsilon(t_2) - Y_\varepsilon(t_1) = \varepsilon \int_{t_1}^{t_2} G(X_\varepsilon(t), t)\,dt + o(\varepsilon^{1/3})\,,$$

and

$$\left|\varepsilon \int_{t_1}^{t_2} G(X_\varepsilon(t), t + L)\,dt - \varepsilon \int_{t_1}^{t_2} G(X_\varepsilon(t), t)\,dt\right|$$

$$\leqq \varepsilon \int_{t_1+L}^{t_2+L} |G(X_\varepsilon(t-L), t) - G(X_\varepsilon(t), t)\,dt + 0(\varepsilon L) = o(\varepsilon^{1/3})$$

by (3.4) and since $\|G\|_1 < \infty$. This proves (3.5) and (3.6) follows by Lemma 3.3 and a similar argument, the $o(1)$ term arising from the defining Z equation (0.7). □

Since $\|E^{(t)}G(X_\varepsilon(t), t + L)\| \leq \|G\|\rho(L) = o(\varepsilon^{2/3})$ by Lemma 3.1, an immediate consequence of (3.5) is

Lemma 3.5. *For* $0 \leqq t_1 < t_2 \leqq \tau\varepsilon^{-4/3}$ *and each fixed* τ,

$$\| E^{(t_1)} Y_\varepsilon(t_2) - Y_\varepsilon(t_1))\| = o(\varepsilon^{1/3})\,, \tag{3.7}$$

and, for $0 \leqq \tau_1 < \tau_2 \leqq \tau_0$ *and each fixed* τ_0,

$$\| E^{(\tau_1\varepsilon^{-4/3})}(V_\varepsilon(\tau_2) - V_\varepsilon(\tau_1))\| \to 0 \tag{3.8}$$

as $\varepsilon \to 0$.

Formula (3.5) also implies that $V_\varepsilon(\tau) = \hat{V}_\varepsilon(\tau) + o(1)$, where

$$\hat{V}_\varepsilon(\tau) := \varepsilon^{2/3} \int_0^{\tau\varepsilon^{-4/3}} G(X_\varepsilon(t), t + L)\,dt\,, \tag{3.9}$$

and we will establish Theorem 1.1 for $\hat{V}_\varepsilon(\tau)$, since the necessary estimates can be provided directly for this quantity.

To simplify notation in what follows we will use the following special symbols: for a vector or matrix M, $M^Q := MM^T$ and for a square matrix M, $M^S := M + M^T$. Note that if M has m rows and n columns then $|M^Q| \leqq |M|^2 \leqq mn|M^Q|$.

Lemma 3.6. *For* $0 \leqq t_1 < t_2 \leqq \tau\varepsilon^{-4/3}$ *and any fixed* τ, *there is a* $C < \infty$ *such that*

$$\left\| E^{(t_1)}\left(\int_{t_1}^{t_2} G(X_\varepsilon(t), t + L)\,dt\right)^Q \right\| \leqq C(t_2 - t_1) \tag{3.10}$$

and

$$\left\| E^{(t_1)}\left(\int_{t_1}^{t_2} H(X_\varepsilon(t), t+L)\,dt\right)^Q \right\| \leqq C(t_2 - t_1) \tag{3.11}$$

Proof. Let

$$m_\varepsilon(t) = \max_{0 \leqq t_2 - t_1 \leqq t} \left\| E^{(t_1)}\left(\int_{t_1}^{t_2} G(X_\varepsilon(s), s+L)\,ds\right)^Q \right\| .$$

We use the decomposition

$$\begin{aligned} E^{(t_1)}&\left(\int_{t_1}^{t_2} G(X_\varepsilon(t), t+L)\,dt\right)^Q \\ &= E^{(t_1)}\left(\int_{t_1}^{t_2} G(X_\varepsilon(t), t+L)\int_t^{t+L} G(X_\varepsilon(s), s+L)^T\,ds\,dt\right)^S \\ &\quad + E^{(t_1)}\left(\int_{t_1}^{t_2} G(X_\varepsilon(t), t+L)\int_{t+L}^{t_2} G(X_\varepsilon(s), s+L)^T\,ds\,dt\right)^S \\ &=: I_1 + I_2 . \end{aligned} \tag{3.12}$$

Applying Lemma 3.1,

$$\|I_1\| \leqq \left\| E^{(t_1)}\left(\int_{t_1}^{t_2} G(X_\varepsilon(t), t+L)\int_t^{t+L} E^{(t+L)}G(X_\varepsilon(s), s+L)^T\,ds\,dt\right)^S \right\| \leqq C(t_2 - t_1)$$

and

$$\begin{aligned} \|I_2\| &\leqq 2\left\| E^{(t_1)}\int_{t_1+L}^{t_2+L} |E^{(s)}G(X_\varepsilon(s), s+L)| \left|\int_{t_1}^{s-L} G(X_\varepsilon(t), t+L)^T\,dt\right| ds \right\| \\ &\leqq C_2 \int_{t_1+L}^{t_2+L} \rho(L) \left| E^{(t_1)}\left(\int_{t_1}^{s-L} G(X_\varepsilon(t), t+L)\,dt\right)^Q \right|^{1/2} ds \\ &\leqq C_2(t_2 - t_1)\, m_\varepsilon^{1/2}(t_2 - t_1)\, o(\varepsilon^{2/3}) . \end{aligned} \tag{3.13}$$

Thus

$$m_\varepsilon(t) \leqq C_1 t + C_2 t\, m_\varepsilon^{1/2}(t)\, o(\varepsilon^{2/3}) ,$$

and for $t \leqq \tau\varepsilon^{-4/3}$ it follows that $m_\varepsilon t \leqq Ct$.
The second inequality is established by a similar argument. □

Proposition 3.1. *For* $0 \leqq t_1 < t_2 \leqq \tau\varepsilon^{-4/3}$ *and any fixed* τ, *as* $\varepsilon \to 0$,

$$\left\| E^{(t_1)}\left(\int_{t_1}^{t_2} G(X_\varepsilon(t), t+L)\,dt\right)^Q - E^{(t_1)}\int_{t_1}^{t_2} \hat{\mathscr{Q}}(X_\varepsilon(t), t+L)\,dt \right\| = (t_2 - t_1)\,o(1) \tag{3.14}$$

and

$$\left\| E^{(t_1)}(Y_\varepsilon(t_2) - Y_\varepsilon(t_1))^Q - \varepsilon^2 E^{(t_1)}\int_{t_1}^{t_2} \hat{\mathscr{Q}}(X_\varepsilon(t), t)\,dt \right\| = o(\varepsilon^{2/3}) . \tag{3.15}$$

Moreover, if (A5) *holds then*

$$\left\| E^{(t_1)}(Y_\varepsilon(t_2) - Y_\varepsilon(t_1))^Q - \varepsilon^2 E^{(t_1)} \int_{t_1}^{t_2} \mathcal{Z}(X_\varepsilon(t))\, dt \right\| = o(\varepsilon^{2/3}) . \tag{3.16}$$

Proof. Apply the decomposition in (3.12). The term I_2 is $(t_2 - t_1)o(1)$ by (3.13), and applying Lemma 3.1,

$$\begin{aligned} I_1 &= E^{(t_1)}\left(\int_{t_1}^{t_2} \int_{t}^{t+L} E^{(t)}[G(X_\varepsilon(t), t + L)\, G(X_\varepsilon(s), s + L)^T]\, ds\, dt \right)^S \\ &\quad + (t_2 - t_1) O(\varepsilon L^2) \\ &= E^{(t_1)}\left(\int_{t_1}^{t_2} \int_{t}^{\infty} \frac{\partial I}{\partial x}(X_\varepsilon(t))\, \Gamma(X_\varepsilon(t), t + L, s + L)^S \frac{\partial I}{\partial x}(X_\varepsilon(t))^T\, ds\, dt \right. \\ &\quad + (t_2 - t_1) o(1) \\ &= E^{(t_1)} \int_{t_1}^{t_2} \hat{\mathcal{Z}}(X_\varepsilon(t), t + L)\, dt + (t_1 - t_2) o(1) \end{aligned}$$

proving (3.14). It follows from (3.5) and (3.7) that

$$E^{(t_1)}(Y_\varepsilon(t_2) - Y_\varepsilon(t_1))^Q = \varepsilon^2 E^{(t_1)} \left(\int_{t_1}^{t_2} G(X_\varepsilon(t), t + L)\, dt \right)^Q + o(\varepsilon^{2/3}) ,$$

and (3.15) follows from this, (3.14) and since

$$|\hat{\mathcal{Z}}(X_\varepsilon(t + L), t + L)) - \hat{\mathcal{Z}}(X_\varepsilon(t), t + L)| \leqq \varepsilon C L .$$

For (3.16) note that

$$\begin{aligned} \int_{t_1}^{t_2} \hat{\mathcal{Z}}(X_\varepsilon(t), t)\, dt &= \int_{t_1}^{t_2} \frac{1}{L} \int_{t}^{t+L} \hat{\mathcal{Z}}(X_\varepsilon(s), t)\, ds\, dt + (t_2 - t_1) o(1) \\ &= \int_{t_1+L}^{t_2} \frac{1}{L} \int_{s-L}^{s} \hat{\mathcal{Z}}(X_\varepsilon(s), t)\, dt\, ds + (t_2 - t_1) o(1) + O(L) \\ &= \int_{t_1}^{t_2} \mathcal{Z}(X_\varepsilon(s))\, ds + o(\varepsilon^{-4/3}) \end{aligned}$$

by (A5). □

Lemma 3.7. *For each fixed* τ, *as* $\varepsilon \to 0$,

$$E\left\{ \max_{0 \leqq t \leqq \tau\varepsilon^{-4/3}} \left| Z_\varepsilon(t) - z_0 - \varepsilon \int_0^t v(Y_\varepsilon(s))\, ds \right| \right\} \to 0 . \tag{3.17}$$

Proof. By (3.6), it suffices to show that

$$B_\varepsilon := E\left\{ \max_{0 \leqq t \leqq \tau\varepsilon^{-4/3}} \left| \varepsilon \int_0^t H(X_\varepsilon(s), s + L)\, ds \right| \right\} \to 0 .$$

as $\varepsilon \to 0$. Let $\lambda = \varepsilon^{-8/9}$. By (A1) and (A3), $\|H\|_1 < \infty$ and, with k an integer,

$$B_\varepsilon \leqq E\left\{\max_{1 \leqq k \leqq \tau\varepsilon^{-4/3}/\lambda}\left|\varepsilon\int_0^{k\lambda} H(X_\varepsilon(s), s+L)\,ds\right|\right\} + O(\varepsilon^{1/9})$$

$$\leqq \sum_{k=1}^{[\tau\varepsilon^{-4/3}/\lambda]} E\left|\varepsilon\int_{(k-1)\lambda}^{k\lambda} H(X_\varepsilon(s), s+L)\,ds\right| + O(\varepsilon^{1/9})$$

$$\leqq C\sum_{k=1}^{[\tau\varepsilon^{-4/3}/\lambda]} \varepsilon\left|E\left(\int_{(k-1)\lambda}^{k\lambda} H(X_\varepsilon(s), s+L)\,ds\right)^Q\right|^{1/2} + O(\varepsilon^{1/9})$$

$$= O(\varepsilon^{1/9})$$

by (3.11). □

Lemma 3.8. *For each fixed τ, as $\varepsilon \to 0$,*

$$E\left(\max_{0 \leqq t \leqq \tau\varepsilon^{-4/3}}\left|\varepsilon\int_0^t v(Y_\varepsilon(s))\,ds - \varepsilon\left(tv(y_0) + \frac{\partial v}{\partial y}(y_0)\int_0^t (Y_\varepsilon(s) - y_0)\,ds\right)\right|\right\} \to 0\,. \tag{3.18}$$

Proof. By Taylor's theorem,

$$v(Y_\varepsilon(s)) - v(y_0) - \frac{\partial v}{\partial y}(y_0)(Y_\varepsilon(s) - y_0),$$

$$= \int_0^1\left[\frac{\partial v}{\partial y}(y_0 + \eta(Y_\varepsilon(s) - y_0)) - \frac{\partial v}{\partial y}(y_0)\right]d\eta(Y_\varepsilon(s) - y_0)\,.$$

Integrating this over $[0, t]$ and using (1.3) gives that the left-hand side of (3.18) is at most

$$C\varepsilon E\int_0^{\tau\varepsilon^{-4/3}} |Y_\varepsilon(s) - y_0|^2\,ds \leqq C\varepsilon\int_0^{\tau\varepsilon^{-4/3}} E|(Y_\varepsilon(s) - y_0)^Q|\,ds$$

for some constant C. However, there exists a constant $C_1 > 0$ such that $|EM^Q| \geqq C_1 E|M^Q|$, and thus using (3.15) and the boundedness of $\hat{\mathscr{Y}}$ gives that the last quantity is at most

$$C_2\varepsilon\int_0^{\tau\varepsilon^{-4/3}} |E(Y_\varepsilon(s) - y_0)^Q|\,ds = o(\varepsilon^{1/3})\,. \quad \square$$

An immediate consequence of Lemma 3.7 and 3.8 is

Proposition 3.2. *For each fixed τ_0, as $\varepsilon \to 0$,*

$$E\left\{\max_{0 \leqq \tau \leqq \tau_0}\left|W_\varepsilon(\tau) - \varepsilon\frac{\partial v}{\partial y}(y_0)\int_0^{\tau\varepsilon^{-4/3}} (Y_\varepsilon(t) - y_0)\,dt\right|\right\} \to 0\,. \tag{3.19}$$

Based on this result and (3.9) it suffices to consider

$$\hat{W}_\varepsilon(\tau) := \frac{\partial v}{\partial y}(y_0)\int_0^\tau \hat{V}_\varepsilon(s)\,ds\,. \tag{3.20}$$

In what follows we establish Theorem 1.2 for $\hat{W}_\varepsilon(\tau)$.

Lemma 3.9. *For $0 \leqq \tau < \tau + \delta \leqq \tau_0$ and any fixed τ_0, there is a constant C depending only on τ_0 such that, as $\varepsilon \to 0$,*

$$E|\hat{V}_\varepsilon(\tau + \delta) - \hat{V}_\varepsilon(\tau)|^4 \leqq C\delta^2 \tag{3.21}$$

and

$$E|\hat{W}_\varepsilon(\tau + \delta) - \hat{W}_\varepsilon(\tau)|^4 \leqq C\delta^2 \tag{3.22}$$

Proof. 1. By (3.9),

$$|\hat{V}_\varepsilon(\tau + \delta) - \hat{V}_\varepsilon(\tau)|^4 \leqq C\left|\varepsilon^{2/3} \int_{\tau\varepsilon^{-4/3}}^{(\tau+\delta)\varepsilon^{-4/3}} G(X_\varepsilon(t), t + L)\,dt\right|^4 . \tag{3.23}$$

Let

$$U_{\varepsilon,t}(l) = \int_t^{t+l} G(X_\varepsilon(s), s + L)\,ds$$

and

$$\beta(l) = \sup_{0 \leqq s \leqq l,\, 0 < \varepsilon \leqq \varepsilon_0,\, t \geqq 0} E|U_{\varepsilon,t}(s)|^4 .$$

To simplify notation we treat $U = U_{\varepsilon,t}$ as one dimensional. Improvising on a method of Borodin [2],

$$EU^4(l) = 4\int_t^{t+l} EG(X_\varepsilon(t+s), t + s + L)\,U^3(s)\,ds = 4\int_t^{t+l} \sum_{k=0}^{3} EI_k(s)\,ds ,$$

where

$$I_k(s) = \binom{3}{k} G(X_\varepsilon(t+s), t + s + L)(U(s) - U(s-L))^k U^{3-k}(s-L) .$$

For $s \leqq l$,

$$|EI_0(s)| = |EU^3(s-L)E^{(t+s)}G(X_\varepsilon(t+s), t + s + L)| \leqq C_0\beta^{3/4}(l)\varepsilon^{2/3} ,$$

and for $k = 1, 2, 3$,

$$\begin{aligned}|EI_k(s)| &\leqq k!\binom{3}{k} \int_{t+s-L}^{t+s} du_1 \ldots \int_{t+s-L}^{u_{k-1}} du_k \\ &\quad \cdot \left|EU^{3-k}(s-L) \prod_{j=1}^{k} G(X_\varepsilon(u_j), u_j + L)E^{(u_1+L)}G(X_\varepsilon(t+s), t + s + L)\right| \\ &\leqq C_k\beta^{(3-k)/4}(l)L^{k-1} .\end{aligned}$$

Combining these estimates yields

$$\beta(l) \leqq C_4 l(\beta^{3/4}(l)\varepsilon^{2/3} + \beta^{1/2}(l) + \beta^{1/4}(l)\varepsilon^{1/3} + \varepsilon^{2/3}) .$$

Now suppose $1 \leqq l \leqq \tau_0\varepsilon^{-4/3}$ and set $B = B(l) = B(l)/l^2$. Then

$$Bl^2 \leqq C_5 l^2(B^{3/4} + B^{1/2} + B^{1/4} + 1) .$$

It follows that $B(l) \leqq C_6 < \infty$, and applying this estimate to (3.23) yields the first assertion.

2. For the second assertion use

$$E|\hat{W}_\varepsilon(\tau+\delta)-\hat{W}_\varepsilon(\tau)|^4 \leqq C_1 E\left|\int_\tau^{\tau+\delta}\hat{V}_\varepsilon(s)\,ds\right|^4 \leqq C_1\delta^4\varepsilon^{8/3}\beta(\tau_0\varepsilon^{-4/3}) \leqq C_2\delta^2 . \quad \square$$

In [3] we use the following variation on the first order stochastic averaging result of Khas'minskii (see [9, 8]):

Proposition 3.3. *For each fixed* l,

$$\sup_t \left\| E^{(t)} \max_{0\leqq s\leqq l} |X_\varepsilon(t+s\varepsilon^{-1})-\bar{x}(s,X_\varepsilon(t))| \right\| \to 0 , \tag{3.24}$$

as $\varepsilon\to 0$.

The result in [3] is stated for times scaled by $1/\varepsilon^2$. As noted there, the uniformity in t and the use of $E^{(t)}$ in place of E are justified by uniformity of conditions (A1), (A2) and (A4).

Lemma 3.10. *Let* (A5) *hold and let* (A6) *or* (A6′) *hold. Then for each* $\delta>0$, *as* $\varepsilon\to 0$,

$$\| E^{(\tau\varepsilon^{-4/3})}(\hat{V}_\varepsilon(\tau+\delta)-\hat{V}_\varepsilon(\tau))^Q-\delta\bar{\mathcal{Q}}(y_0)\| \to 0 . \tag{3.25}$$

(To simplify notation we use $\bar{\mathcal{Q}}(y_0)$ for M_0 under (A6′) as well as (A6) since as noted under (A6′), M_0 and $\bar{\mathcal{Q}}(y_0)$ coincide when (A6) and (A6′) both hold.

Proof. For fixed τ and l, using Proposition 3.3 at the second equality,

$$E^{(\tau\varepsilon^{-4/3})}\frac{1}{\varepsilon^{1/3}l}\int_\tau^{\tau+\varepsilon^{1/3}l}\mathcal{Q}(X_\varepsilon(s\varepsilon^{-4/3}))\,ds=\frac{1}{l}\int_0^l E^{(\tau\varepsilon^{-4/3})}\mathcal{Q}(X_\varepsilon(\tau\varepsilon^{-4/3}+r\varepsilon^{-1}))\,dr$$

$$=\frac{1}{l}\int_0^l E^{(\tau\varepsilon^{-4/3})}\mathcal{Q}(\bar{x}(r,X_\varepsilon(\tau\varepsilon^{-4/3})))\,dr+o(\varepsilon^{1/3}) .$$

Under (A6) the last integral is $E^{(\tau\varepsilon^{-4/3})}\bar{\mathcal{Q}}(Y_\varepsilon(\tau\varepsilon^{-4/3}))+o(1)$ in the iterated limit as $\varepsilon\to 0$ then $l\to\infty$, which is $\bar{\mathcal{Q}}(y_0)+o(1)$ since $\|E^{(\tau\varepsilon^{-4/3})}(Y_\varepsilon(\tau\varepsilon^{-4/3})-y_0)^Q\|\to 0$ as $\varepsilon\to 0$ by (3.15) and since $\|\bar{\mathcal{Q}}\|_1\leqq\|\mathcal{Q}\|_1<\infty$.

Under (A6′), the above convergence of $Y_\varepsilon(\tau\varepsilon^{-4/3})$ implies there exist $\hat{X}_\varepsilon$ such that $I(\hat{X}_\varepsilon)=y_0$ and $X_\varepsilon(\tau\varepsilon^{-4/3})-\hat{X}_\varepsilon\to 0$ uniformly in $P^{(\tau\varepsilon^{-4/3})}$ probability. Using the continuous dependence of $\bar{x}(r,x)$ on x and the dominated convergence theorem, it again follows that

$$E^{(\tau\varepsilon^{-4/3})}\int_\tau^{\tau+\varepsilon^{1/3}l}\mathcal{Q}(X_\varepsilon(s\varepsilon^{-4/3}))\,ds=\varepsilon^{1/3}l(\bar{\mathcal{Q}}(y_0)+o(1))$$

as $\varepsilon\to 0$ then $l\to\infty$.

Applying (1.12), (3.7), (3.16) and a change of the integration variable,

$$E^{(\tau\varepsilon^{-4/3})}(\hat{V}_\varepsilon(\tau+\delta)-\hat{V}_\varepsilon(\tau))^Q=\varepsilon^{-2/3}E^{(\tau\varepsilon^{-4/3})}(Y_\varepsilon((\tau+\delta)\varepsilon^{-4/3})-Y_\varepsilon(\tau\varepsilon^{-4/3}))^Q+o(1)$$

$$=E^{(\tau\varepsilon^{-4/3})}\int_\tau^{\tau+\delta}\mathcal{Q}(X_\varepsilon(\sigma\varepsilon^{-4/3}))\,d\sigma+o(1)$$

$$= \int_{\tau}^{\tau+\delta} \frac{1}{\varepsilon^{1/3} l} \int_{\sigma}^{\sigma+\varepsilon^{1/3} l} E^{(\tau\varepsilon^{-4/3})} \bar{\mathfrak{Z}}(X_\varepsilon(s\varepsilon^{-4/3}))\, ds\, d\sigma + o(1)$$

$$= \int_{\tau}^{\tau+\delta} (\bar{\mathfrak{Z}}(y_0) + o(1))\, d\sigma + o(1)$$

$$= \delta \bar{\mathfrak{Z}}(y_0) + o(1),$$

as $\varepsilon \to 0$ then $l \to \infty$. □

Proof of Theorem 1.1. By (3.9) it suffices to prove the asserted limit for the processes $\{\hat{V}_\varepsilon(\tau), \tau \geqq 0\}$ as $\varepsilon \to 0$. Lemma 3.9 implies that these processes are relatively compact in $C_p[0, \tau_0)$ for each τ_0, and this implies relative compactness in $C_p[0, \infty)$ ([1] [7]). Moreover fourth moments are bounded on each $[0, \tau_0)$, so the $\hat{V}_\varepsilon(\tau)$ are uniformly integrable as $\varepsilon \to 0$.

To prove the asserted weak convergence when (A5) and (A6) hold, it suffices to show that for any weakly convergent sequence $\hat{V}_{\varepsilon_n} \to V_0$ the limit process has the asserted distribution. Using the Skorohod representation theorem, we can assume without loss of generality that $\hat{V}_{\varepsilon_n} \to V_0$ a.s. Now let $\tau_1 < \tau_2 < \cdots < \tau_k \leqq \tau < \tau + \delta$ and η: $\mathbb{R}^k \to \mathbb{R}^1$ be bounded and continuous. Then, using (3.25) at the third equality below,

$$E[\eta(V_0(\tau_1), \ldots, V_0(\tau_k))(V_0(\tau+\delta) - V_0(\tau))^Q]$$

$$= \lim_{n\to\infty} E[\eta(\hat{V}_{\varepsilon_n}(\tau_1), \ldots, \hat{V}_{\varepsilon_n}(\tau_k))(\hat{V}_{\varepsilon_n}(\tau+\delta) - \hat{V}_{\varepsilon_n}(\tau))^Q]$$

$$= \lim_{n\to\infty} E[\eta(\hat{V}_{\varepsilon_n}(\tau_1), \ldots, \hat{V}_{\varepsilon_n}(\tau_k)) E^{(\tau\varepsilon^{-4/3})}(\hat{V}_{\varepsilon_n}(\tau+\delta) - \hat{V}_{\varepsilon_n}(\tau))^Q]$$

$$= \lim_{n\to\infty} E[\eta(\hat{V}_{\varepsilon_n}(\tau_1), \ldots, \hat{V}_{\varepsilon_n}(\tau_k)) \delta \bar{\mathfrak{Z}}(y_0)]$$

$$= E[\eta(V_0(\tau_1), \ldots, V_0(\tau_k)) \delta \bar{\mathfrak{Z}}(y_0)] .$$

Letting $\mathscr{A}_\tau = \sigma(V_0(s), s < \tau)$, it follows that

$$E((V_0(\tau+\delta) - V_0(\tau))^Q | \mathscr{A}_\tau) = \delta \bar{\mathfrak{Z}}(y_0) \text{ a.s.},$$

and a similar argument using (3.8) in place of (3.25) shows that $V_0(\tau)$ is a martingle. It then follows that $V_0(\tau) = (\bar{\mathfrak{Z}}(y_0))^{1/2} B(\tau)$, where $B(\tau)$ is standard Browninan motion in $\mathbb{R}^q$. (The essential ideas are in [4] and a modern treatment of the multivariate case can be found in [7].) □

Proof of Theorem 1.2. The relative compactness in $C_q[0, \tau_0]$ for each τ_0 follows from (3.22).

Now assume (A5) and (A6), so $\hat{V}_\varepsilon \to V_0$ weakly as $\varepsilon \to 0$. Since by (3.20), each $\hat{W}_\varepsilon(\tau)$ is a continuous functional of $\{\hat{V}_\varepsilon(s): s \leqq \tau\}$, it follows that $\hat{W}_\varepsilon$ converges weakly to W_0, where the process $\{W_0(\tau): \tau \geqq 0\}$ has the distribution of

$$\frac{\partial v}{\partial y}(y_0) \int_0^\tau V_0(s)\, ds, \; \tau > 0 . \tag{3.26}$$

Since V_0 is Gaussian, it follows that W_0 is Gaussian, too, and the independence of the increments of V_0 implies W_0 is a Markov process. Moreover

$$EW_0(\tau) = \frac{\partial v}{\partial y}(y_0) \int_0^\tau EV_0(s)\,ds = 0\,,$$

and for $0 \leqq \tau_1 \leqq \tau_2$

$$\begin{aligned} E[W_0(\tau_1)W_0(\tau_2)^T] &= \frac{\partial v}{\partial y}(y_0) \int_0^{\tau_1}\int_0^{\tau_2} E[V_0(s)V_0(t)^T]\,dt\,ds \frac{\partial v}{\partial y}(y_0)^T \\ &= \frac{\partial v}{\partial y}(y_0)\bar{\mathscr{Q}}(y_0)\frac{\partial v}{\partial y}(y_0)^T \int_0^{\tau_1}\int_0^{\tau_2} \min\{s,t\}\,dt\,ds \\ &= \frac{\tau_1^2}{6}(3\tau_2 - \tau_1)\frac{\partial v}{\partial y}(y_0)\bar{\mathscr{Q}}(y_0)\frac{\partial v}{\partial y}(y_0)^T\,. \end{aligned}$$

□

Proof of Theorem 1.3. Let

$$\Psi_\varepsilon(\tau,\omega) = \hat{\mathscr{Q}}(X_\varepsilon(\tau\varepsilon^{-4/3},\omega),\tau\varepsilon^{-4/3})\,.$$

Since $\|\hat{\mathscr{Q}}\| < \infty$, $\{\Psi_\varepsilon(\tau,\omega)\}_{0<t\leqq\varepsilon_0}$ is a weakly sequentially compact set in $L_1([0,\tau_0]\times\Omega,\ m\times P)$ for each τ_0, where m is Lebesgue measure [5]. Using a diagonalization argument, given any sequence of the Ψ_ε's, there exists a subsequence that converges weakly in $L_1([0,\tau]\times\Omega,\ m\times P)$ for every finite τ. By Theorem 1.1, the $\hat{V}_\varepsilon$ processes are relatively compact, and from any weakly convergent sequence of $\hat{V}_\varepsilon$ we can extract an L_1-weakly convergent subsequence $\hat{V}_{\varepsilon_n}$ such that $\Psi_{\varepsilon_n} \to \Psi_0$ for some Ψ_0, the convergence holding for $L_1([0,\tau]\times\Omega,\ m\times P)$ and each $\tau < \infty$. It suffices to establish that the asserted uniform limit distribution of $Z_\varepsilon(\tau\varepsilon^{-4/3})(\operatorname{mod}\xi)$ is approached by all such sequences in the iterated limit as $\varepsilon_n \to 0$ then $\tau \to \infty$.

Using the Skorohod representation theorem, letting $\tau_1 < \tau_2 < \cdots < \tau_k \leqq \tau < \tau + \delta$ and $\eta\colon \mathbb{R}^k \to \mathbb{R}^1$ be bounded and continuous,

$$\begin{aligned} &E[\eta(V_0(\tau_1),\ldots,V_0(\tau_k))(V_0(\tau+\delta)-V_0(\tau))^{Q}] \\ &\quad= \lim_{n\to\infty} E[\eta(\hat{V}_{\varepsilon_n}(\tau_1),\ldots,\hat{V}_{\varepsilon_n}(\tau_k))(\hat{V}_{\varepsilon_n}(\tau+\delta)-\hat{V}_{\varepsilon_n}(\tau))^{Q}] \\ &\quad= \lim_{n\to\infty} E\int_\tau^{\tau+\delta} \eta(\hat{V}_{\varepsilon_n}(\tau_1),\ldots,\hat{V}_{\varepsilon_n}(\tau_k))\,\Psi_{\varepsilon_n}(s)\,ds \\ &\quad= \lim_{n\to\infty} E\int_\tau^{\tau+\delta} \eta(V_0(\tau_1),\ldots,V_0(\tau_k))\,\Psi_{\varepsilon_n}(s)\,ds \\ &\quad= E\int_\tau^{\tau+\delta} \eta(V_0(\tau_1),\ldots,V_0(\tau_k))\Psi_0(s)\,ds\,, \end{aligned} \tag{3.27}$$

where at the second equality we use (3.15) and at the third equality we use the a.e. convergence of $V_{\varepsilon_n}(\tau_j)$ to $V_0(\tau_j)$, which implies

$$\begin{aligned} &\left|E\int_\tau^{\tau+\delta} (\eta(\hat{V}_{\varepsilon_n}(\tau_1),\ldots,\hat{V}_{\varepsilon_n}(\tau_k)) - \eta(V_0(\tau_1),\ldots,V_0(\tau_k)))\,\Psi_{\varepsilon_n}(s)\,ds\right| \\ &\quad\leq \|\hat{\mathscr{Q}}\| \int_\tau^{\tau+\delta} E|\eta(\hat{V}_{\varepsilon_n}(\tau_1),\ldots,\hat{V}_{\varepsilon_n}(\tau_k)) - \eta(V_0(\tau_1),\ldots,V_0(\tau_k))|\,ds \to 0\,, \end{aligned}$$

as $n \to \infty$ by the dominated convergence theorem. Again we set $\mathscr{A}_\tau = \sigma(V_0(s), s \leqq \tau)$. Then (3.27) implies

$$E((V_0(\tau+\delta) - V_0(\tau))^Q \mid \mathscr{A}_\tau) = E\left(\int_\tau^{\tau+\delta} \Psi_0(s)\,ds \mid \mathscr{A}_\tau\right) \text{ a.s.}\,.$$

Let $\Psi(\tau) = E(\Psi_0(\tau) \mid \mathscr{A}_\tau)$. Then

$$E\left(\int_\tau^{\tau+\delta} \Psi_0(s)\,ds \mid \mathscr{A}_\tau\right) = E\left(\int_\tau^{\tau+\delta} E(\Psi_0(s) \mid \mathscr{A}_s)\,ds \mid \mathscr{A}_\tau\right) = E\left(\int_\tau^{\tau+\delta} \Psi_0(s)\,ds \mid \mathscr{A}_\tau\right) \text{ a.s.}$$

and

$$E\left(\frac{\partial v}{\partial y}(y_0)(V_0(\tau+\delta) - V_0(\tau))^Q \frac{\partial v}{\partial y}(y_0)^T \mid \mathscr{A}_\tau\right)$$
$$= E\left(\int_\tau^{\tau+\delta} \frac{\partial v}{\partial y}(y_0)\,\Psi(s)\frac{\partial v}{\partial y}(y_0)^T\,ds \mid \mathscr{A}_\tau\right) \text{ a.s.}$$

As in Theorem 1.1, $\frac{\partial v}{\partial y}(y_0)V_0(\tau)$ is a martingale in $C_q[0,\tau_0]$ with $\frac{\partial v}{\partial y}(y_0)V_0(0) = 0$. By [4] $\frac{\partial v}{\partial y}(y_0)V_0(\tau)$ has the representation

$$\frac{\partial v}{\partial y}(y_0)\,V_0(\tau) = \int_0^\tau \left(\frac{\partial v}{\partial y}(y_0)\,\Psi(s)\frac{\partial v}{\partial y}(y_0)^T\right)^{1/2} dB(s)\,, \tag{3.28}$$

where $B(s)$ is a Brownian motion process in $\mathbb{R}^q$ with $B(\tau+\delta) - B(\tau)$ independent of $\mathscr{A}_\tau$ for each τ.

Next we establish a lower bound on $\frac{\partial v}{\partial y}(y_0)\,\Psi\,\frac{\partial v}{\partial y}(y_0)^T$ using (A7). Let $A \in \mathscr{F}$ and $\tau, \delta \geqq 0$ and let $K = T_1 T_2$:

$$\int_A \int_\tau^{\tau+\delta} \frac{\partial v}{\partial y}(y_0)\,\Psi_\varepsilon(s)\frac{\partial v}{\partial y}(y_0)^T\,ds\,dP$$
$$= \int_A \int_\tau^{\tau+\delta} \frac{1}{T_1}\int_0^{T_1} \frac{\partial v}{\partial y}(y_0)\,\mathfrak{Z}(X_\varepsilon(s\varepsilon^{-4/3} - t), s\varepsilon^{-4/3})\frac{\partial v}{\partial y}(y_0)^T\,dt\,ds\,dP + O(\varepsilon)$$
$$= \int_A \int_\tau^{\tau+\delta} \frac{1}{T_1}\int_0^{T_1} \frac{\partial v}{\partial y}(y_0)\,\mathfrak{Z}(X_\varepsilon(s\varepsilon^{-4/3}), s\varepsilon^{-4/3} + t)\frac{\partial v}{\partial y}(y_0)^T\,dt\,ds\,dP + O(\varepsilon)$$
$$\geqq \int_A \int_\tau^{\tau+\delta} \frac{1}{T_1}\Delta(X_\varepsilon(s\varepsilon^{-4/3}))\,ds\,dP + O(\varepsilon)$$
$$= \int_A \int_0^{T_2} \int_{\tau+t\varepsilon^{1/3}}^{\tau+\delta+t\varepsilon^{1/3}} \frac{1}{K}\Delta(X_\varepsilon(s\varepsilon^{-4/3}))\,ds\,dt\,dP + O(\varepsilon^{1/3})$$
$$= \int_A \int_0^{T_2} \int_\tau^{\tau+\delta} \frac{1}{K}\Delta(X_\varepsilon(s\varepsilon^{-4/3} + t\varepsilon^{-1}))\,ds\,dt\,dP + O(\varepsilon^{1/3})$$
$$= \int_A \int_\tau^{\tau+\delta} \int_0^{T_2} \frac{1}{K}\Delta(X_\varepsilon(s\varepsilon^{-4/3} + t\varepsilon^{-1}))\,ds\,dt\,dP + O(\varepsilon^{1/3})$$
$$= \int_A \int_\tau^{\tau+\delta} \int_0^{T_2} \frac{1}{K}\Delta(\bar{x}(t, X_\varepsilon(s\varepsilon^{-4/3})))\,ds\,dt\,dP + o(1)$$
$$\geqq \int_A \int_\tau^{\tau+\delta} \frac{1}{K}\Delta_0\,ds\,dP + o(1)\,,$$

where Proposition 3.3 is applied at the next to last line.

Thus

$$\int_A \int_\tau^{\tau+\delta} \frac{\partial v}{\partial y}(y_0)\,\Psi_\varepsilon(s)\frac{\partial v}{\partial y}(y_0)^T\,ds\,dP \geqq \int_A \int_\tau^{\tau+\delta} \frac{1}{K}\Delta_0\,ds\,dP\,,$$

and it follows that $\frac{\partial v}{\partial y}(y_0)\Psi_0\frac{\partial v}{\partial y}(y_0)^T \geqq K^{-1}\Delta_0$ a.e. and, since conditional expectation is a positive operator, $\frac{\partial v}{\partial y}(y_0)\Psi\frac{\partial v}{\partial y}(y_0)^T \geqq K^{-1}\Delta_0$ a.e. Let $B^{(1)}$ be Brownian motion in $\mathbb{R}^q$ with $B^{(1)}(\tau+\delta) - B^{(1)}(\tau)$ independent of $\mathscr{A}_\tau$, for each τ, and let $B^{(2)}$ be a second Brownian motion in $\mathbb{R}^q$ independent of $B^{(1)}$ and $\sigma(\cup\mathscr{A}_\tau)$. Then

$$M(\tau) := \int_0^\tau \left(\frac{\partial v}{\partial y}(y_0)\Psi(s)\frac{\partial v}{\partial y}(y_0)^T - \frac{1}{K}\Delta_0\right)^{1/2} dB^{(1)}(s) + \left(\frac{1}{K}\Delta_0\right)^{1/2} B^{(2)}(\tau)\,,$$

is a martingale with respect to $\mathscr{B}_\tau := \sigma(\mathscr{A}_\tau \cup \sigma(M(s), 0 \leqq s \leqq \tau))$ and satisfies

$$E((M(\tau+\delta) - M(\tau))^Q \,|\, \mathscr{B}_\tau) = E\left(\int_\tau^{\tau+\delta} \frac{\partial v}{\partial y}(y_0)\Psi(s)\frac{\partial v}{\partial y}(y_0)^T\,ds \,|\, \mathscr{B}_\tau\right) \text{ a.s.}$$

But then $M(\tau)$ has a representation like that in (3.28) with $\mathscr{A}_\tau$ replaced by $\mathscr{B}_\tau$, so this process is equal in distribution to the process $\frac{\partial v}{\partial y}(y_0)V_0(\tau)$. Using (3.26), $W_0(\tau)$ is equal in distribution to

$$\int_0^\tau \int_0^s \left(\frac{\partial v}{\partial y}(y_0)\Psi(s)\frac{\partial v}{\partial y}(y_0)^T - \frac{1}{K}\Delta_0\right)^{1/2} dB^{(1)}(t)\,ds + \left(\frac{1}{K}\Delta_0\right)^{1/2} \int_0^\tau B^{(2)}(s)\,ds\,.$$

The second term is independent of the first and Gaussian with covariance matrix $(\tau^3/3K)\Delta_0$. The density for $W_0(\tau)$ is the convolution of this Gaussian density with that of the first term, and has derivative bounded by the derivative of the Gaussian density. It follows that $W_0(\tau)$ (mod ζ) is approximately uniform for large τ. □

References

1. Billingsley, P.: Convergence of Probability Measures, New York: Wiley, 1968
2. Borodin, A.N.: A limit theorem for solutions of differential equations with random right-hand side. Theory Prob. Appl. **22**, 482–497 (1977)
3. Cogburn, R., Ellison, J.A.: A stochastic theory of adiabatic invariance. Commun. Math. Phys. **148**, 97–126 (1992)
4. Doob, J.L.: Stochastic Processes. New York: Wiley, 1953
5. Dunford, N., Schwartz, J.: Linear Operators Part I: General Theory. New York: Wiley-Interscience, 1958
6. Ellison, J.A., Guinn, T.: Statistical equilibrium, planar channeling and the continuum model. Phys. Rev. B **13**, 1880–1883 (1976)
7. Ethier, S.N., Kurtz, T.G.: Markov Processes: Characterization and Convergence. New York: Wiley-Interscience, 1986
8. Freidlin, M.I., Wentzell, A.D.: Random Perturbations of Dynamical Systems. Berlin, Heidelberg, New York: Springer, 1984
9. Khas'minskii, R.Z.: On processes defined by differential equations with a small parameter. Theory Prob. Appl. **11**, 211–228 (1966)
10. Khas'minskii, R.Z.: A limit theorem for solutions of differential equations with random right-hand side. Theory Prob. Appl. **11**, 390–406 (1966)

Communicated by J.L. Lebowitz