Sankhyā : The Indian Journal of Statistics
Special issue in memory of D. Basu
2002, Volume 64, Series A, Pt. 3, pp 741-762

PHASE CHANGES WITH TIME FOR A CLASS OF AUTONOMOUS MULTISCALE DIFFUSIONS

By RABI BHATTACHARYA
Indiana University, Bloomington, USA

SUMMARY. We obtain precise time scales for different Gaussian phases of a class of multiscale diffusions with periodic coefficients, improving earlier results. The first Gaussian phase occurs for times $1 \ll t \ll a^{2/3}$ where 'a' is the large scale parameter, while the final Gaussian phase occurs for times $t \gg a^2$. Examples show that non-Gaussian phases appear between these two time zones. An application to a problem of solute transport in porous media is given, along with a survey of related literature.

1. Introduction

Multiscale phenomena occur commonly in nature and their analysis presents challenging problems (Glimm and Sharp, 1997). In the present article we provide a fairly complete analysis of the asymptotic behaviour of a class of multiscale diffusions which arise in the context of solute transport in porous media. To state the physical problem motivating this work, let $c(t, y)$ denote the concentration of a solute, say a chemical pollutant, in a saturated aquifer at time t and at a point y due to a point injection. The Fokker-Planck equation governing c is generally taken to be of the form (with $k = 3$)

$$\frac{\partial c}{\partial t} = -\sum_{i=1}^{k} \frac{\partial}{\partial y_i}\{v_i(y)c\} + \tfrac{1}{2}\sum_{i,j=1}^{k} \frac{\partial^2}{\partial y_i \partial y_j}\{D_{ij}(y)c\}(t > 0, y \in \mathbb{R}^k), \tag{1.1}$$
$$c(0, y) = c_0\delta_x(dy),$$

Paper received January 2001.
AMS (1991) *subject classification.* Primary 60J05; secondary 60H10, 60J70.
Keywords and phrases. Diffusion on a torus, speed of convergence to equilibrium, Gaussian and non-Gaussian phases, growth in dispersivity.

where $v(y) = (v_1(y), \cdots, v_k(y))$ is the velocity of water in the aquifer at a point y, $D(y)$ is the local (nonsingular) diffusion matrix, and the initial condition says that initially a mass c_0 of the solute is injected at a point x. One requires that the divergence of v be zero,

$$\sum_{i=1}^{k} \frac{\partial v_i(y)}{\partial y_i} = 0 \qquad (y \in \mathbb{R}^k), \tag{1.2}$$

corresponding to the incompressibility of water in the saturated aquifer. The main problem is to determine how the pollutant will spread as $t \to \infty$. Since some experiments with natural aquifers show Gaussian profiles of $c(t, y)$ for large times t (see Fried and Combarnous, 1971, and the articles in Cushman, 1990), one may impose further restrictions on $v(\cdot)$ and $D(\cdot)$. Gelhar and Axness (1983) and Winter et al. (1984) assume that $v(\cdot)$ and $D(\cdot)$ are ergodic random fields and use formal calculations which (i) indicate asymptotic Gaussian behaviour and (ii) provide an analysis of the dispersion matrix of this Gaussian.

A further intriguing phenomenon that has been observed is the increase in dispersivity, (i.e., the dispersion matrix per unit time of the fitted Gaussian) as time progresses (See Bhattacharya et al., 1989, Cushman, 1990, and the references therein). Referred to as the "scale effect in dispersion" in the engineering literature, such behaviour can only arise in the presence of multiple scales of heterogeneity in the geological formation of the aquifer (Bhattacharya and Gupta, 1983). For example, from high up one may observe evolving heterogeneities with distance in the topography of a region. One way to mathematically formulate this for two scales is to express the velocity field $v(\cdot)$ in (1.1) as

$$v(y) = b(y) + \beta(y/a), \tag{1.3}$$

where $b(\cdot)$ and $\beta(\cdot)$ are divergence-free vector fields and 'a' is a large parameter.

To analyse the deterministic equation (1.1) using probability, note that the solution $c(t, y)$ is given by

$$c(t, y) = c_0 p(t; x, y) \tag{1.4}$$

where $p(t; x, y)$ is the transition probability density of the Markov process (diffusion) $X(t)$ on $\mathbb{R}^k$ governed by Itô's stochastic integral equation

$$X(t) = x + \int_0^t \{b(X(s)) + \beta(X(s)/a)\}ds + \int_0^t \sigma(X(s))dB(s). \tag{1.5}$$

Here $\sigma(y)$ is the positive definite square root of $D(y)$, and $\{B(s) : s \geq 0\}$ is the standard k-dimensional Brownian motion. Thus studying the asymptotics of $c(t, y)$ is equivalent to studying the asymptotic distribution of $X(t)$. Since $\beta(y/a)$ changes slowly with y for large 'a', one may expect that for a certain initial period of time $X(t)$ may be well approximated by replacing $\beta(\cdot)$ by its initial value $\beta(x/a)$, i.e., by the solution $X_0(t)$ of the Itô's equation

$$X_0(t) = x + \int_0^t \{b(X_0(s)) + \beta(x_0)\}ds + \int_0^t \sigma(X_0(s))dB(s), \qquad (1.6)$$

where the initial point x is scaled as

$$x = ax_0 \qquad (1.7)$$

for some fixed x_0 (not depending on 'a'). Such a scaling of the initial point is not essential for our results but appropriate, as it corresponds to keeping x relatively in the same position with respect to the rest of the space under the map $y \to ay$, as $a \to \infty$. Applying Girsanov's Theorem (Ikeda and Watanabe, 1981, pp.176-181) to compute the probability density of the distribution P^t of $\{X(s) : 0 \leq s \leq t\}$ with respect to the distribution P_0^t of $\{X_0(s) : 0 \leq s \leq t\}$, one arrives at the following result (Bhattacharya and Götze, 1995). We write $||\mu - v||_{TV}$ for the total variation distance between probability measures μ and ν.

THEOREM 1.1. *Assume that $b(\cdot)$ and $\beta(\cdot)$ are bounded and differentiable with bounded derivatives, and that $\beta(\cdot)$ has continuous and bounded second order derivatives. In addition, assume $D(\cdot)$ is Lipschitzian and has eigenvalues bounded away from zero and infinity. Let P^t, P_0^t denote the distributions on $C([0, T] \to \mathbb{R}^k)$ of the processes $\{X(s) : 0 \leq s \leq t\}$ and $\{X_0(s) : 0 \leq s \leq t\}$, respectively. Then there exist positive constants $c_i (i = 1, 2, 3)$ independent of 'a' such that*

$$||P^t - P_0^t||_{TV} \leq c_1 \frac{t^{3/2}}{a} + c_2 \frac{t}{a} + c_3 \frac{t^{3/2}}{a^2}. \qquad (1.8)$$

In particular, if $a \to \infty$ and $t/a^{2/3} \to 0$, then $||P^t - P_0^t||_{TV} \to 0$.

The following corollary is now immediate.

COROLLARY 1.2. *Assume, in addition to the hypothesis of Theorem 1.1, that $b(\cdot)$ and $D(\cdot)$ are such $X_0(t)$ is asymptotically Gaussian, then $X(t)$ (and $c(t, y)$) is asymptotically Gaussian, as $t \to \infty$ and $t/a^{2/3} \to 0$.*

From our point of view, a particularly important case satisfying the additional hypothesis in Corollary 1.2 is that of periodic $b(\cdot)$ and $D(\cdot)$ with a common period lattice, which we may take to be $\mathbb{Z}^k$ without any essential loss of generality. To prove the asymptotic Gaussian property of $X_0(t)$, write $\dot{X}_0(t) := X_0(t) \mod 1 \equiv (X_{01}(t) \mod 1, \ldots, X_{0k}(t) \mod 1)$, with $X_{0i}(t)$ denoting the i-th coordinate of $X_0(t)$. The process $\dot{X}_0$ is a Markov process on the torus $\mathcal{T}_1 = \{x \mod 1 : x \in \mathbb{R}^k\}$, which has a unique invariant probability π (See Bensoussan et al., 1978, Bhattacharya, 1985). One may write

$$X_0(t) = X_0(0) + \int_0^t \{b(\dot{X}_0(s) + \beta(x_0)\}ds + \int_0^t \sigma(\dot{X}_0(s))dB(s). \tag{1.9}$$

It follows from the general theory of ergodic Markov processes (Bhattacharya, 1982) that, for each j, there exists a unique mean-zero element ψ_j of $L^2(\mathcal{T}_1, \pi)$ which satisfies

$$L\psi_j(y) = b_j(y) - \overline{b}_j, \overline{b}_j := \int_{\mathcal{T}_1} b_j d\pi \ (1 \leq j \leq k), \tag{1.10}$$

where L is the infinitesimal generator of $\dot{X}(\cdot)$,

$$L = \sum_{j=1}^k \{b_j(y) + \beta_j(x_0)\} \frac{\partial}{\partial y_j} + \tfrac{1}{2} \sum_{i,j=1}^k D_{ij}(y) \frac{\partial^2}{\partial y_i \partial y_j}. \tag{1.11}$$

One may now use Itô's Lemma to get

$$\begin{aligned} \psi_j(\dot{X}_0(t)) - \psi_j(\dot{X}_0(0)) &= \int_0^t (b_j(\dot{X}_0(s)) - \overline{b}_j)ds \\ &+ \int_0^t \text{grad } \psi_j(\dot{X}_0(s)) \cdot \sigma(\dot{X}_0(s))dB(s)(1 \leq j \leq k). \end{aligned} \tag{1.12}$$

Writing $\psi(\cdot) = (\psi_1(\cdot), \ldots, \psi_k(\cdot))', \text{Grad } \psi = (\text{grad } \psi_1, \ldots, \text{grad } \psi_k)'$ one may express (1.9) as

$$\begin{aligned} \frac{X_0(t) - X_0(0) - t(\overline{b} + \beta(x_0))}{\sqrt{t}} &= \frac{\psi(\dot{X}_0(t)) - \psi(\dot{X}_0(0))}{\sqrt{t}} \\ &- \frac{1}{\sqrt{t}} \int_0^t (\text{Grad } \psi(\dot{X}_0(s)) - I_k)\sigma(\dot{X}_0(s))dB(s) \end{aligned} \tag{1.13}$$

where I_k is the $k \times k$ identity matrix. Assume first $\dot{X}_0(0)$ has distribution π. The first term on the right goes to zero in $L^2(\mathcal{T}, \pi)$, and hence in probability, as $t \to \infty$. For the second term, one may use the martingale central limit theorem or, more directly, characteristic functions and the ergodic theorem, to prove convergence of the right side to the Gaussian, or normal, distribution $N(0, K_0)$, where

$$K_0 = \int_{\mathcal{T}_1} (\text{Grad } \psi(y) - I_k) D(y) (\text{Grad } \psi(y) - I_k)' \pi(dy). \tag{1.14}$$

Finally, this Gaussian convergence is uniform with respect to the (initial) distribution of $\dot{X}_0(0)$. For, by applying Doeblin's minorization of the (transition) probability density $p(1; z, y)$ of $\dot{X}_0(1)$, given $\dot{X}_0(0) = z$, one shows that the distribution of $\dot{X}_0(t)$ converges exponentially fast in total variation distance to π, as $t \to \infty$, uniformly *w.r.t.* z. It follows that the left side of (1.13) converges in distribution to $N(0, K_0)$, uniformly w.r.t. $X_0(0) = x$. This derivation of the central limit theorem for diffusions with periodic coefficients is essentially due to Bensonssan et al. (1978); a different approach may be found in Bhattacharya (1985).

In the context of Corollary 1.2, $X_0(t)$ may be shown to be asymptotically Gaussian if the generator of $X_0(\cdot)$ is in the divergence form

$$\frac{1}{2} \sum_{i,j=1}^{k} \frac{\partial}{\partial y_i} (D_{ij}(y)) \frac{\partial}{\partial y_j},$$

$\{D(y) : y \in \mathbb{R}^k\}$ is an ergodic random field and the eigenvalues of $D(y)$ are bounded away from zero and infinity (Papanicolaou and Varadhan, 1979). In the solute transport problem considered in this article, the generator is not in divergence form.

We will refer to the Gaussian limit in Corollary 1.2 as the *first phase*, which holds as $t \to \infty, t/a^{2/3} \to 0$. Example 1 below shows that this time scale for the first phase is precise, in the sense that for larger times (relative to 'a') it does break down in some instances. What happens at larger times of course depends on the nature of $\beta(\cdot)$. In order that a second Gaussian phase may arise at sufficiently large times we now assume that $b(\cdot), \beta(\cdot)$ and $D(\cdot)$ are all periodic with the same period lattice $\mathbb{Z}^k$ and that 'a' is an integer. It then follows from the central limit theorem for diffusions with periodic coefficients described above that, for each fixed 'a', $X(t)$ is asymptotically Gaussian with mean $t(\overline{b} + \overline{\beta})$ and dispersion matrix tK different from tK_0. For large 'a' when does this phase take hold? We will show under further assumptions that the time scale for this *final phase* is $t \gg a^2$, i.e., $t/a^2 \to \infty$. We may thus remove the $\log a, (\log a)^2$ factors appearing in the time scales for the final phase in Bhattacharya and Götze (1995, correction, ibid., 1996) and Bhattacharya (1999). To arrive at these precise time scales we have made use of both the spectral method of our earlier work (see Propositions 2.1, 3.3, 3.6) as well as a fascinating new estimate for a Doeblin type minorization due to Franke (2001) which is uniform over all divergence free vector fields (Proposition 2.2). In addition, some awkward technical conditions used earlier are either dispensed with or simplified.

We conclude this section with two examples which show that (1) these time scales are precise in general, (2) various other Gaussian non-Gaussian phases may occur in between the first and final Gaussian phases, and (3) the growth in dispersivity K as 'a'$\to \infty$ depends dramatically on the nature of $\beta(\cdot)$. These examples are taken from Bhattacharya and Götze (1995) and Bhattacharya (1999).

EXAMPLE 1. Let $k = 2, b_1(y) = c_0 + c_1 \sin 2\pi y_2, b_2(y) = \delta, \beta_1(y) = \cos 2\pi y_2, \beta_2(y) = 0$, with c_1, c_2, δ nonzero. Let the diffusion matrix be the identity matrix. Also for the initial point $x = ax_0$, take $x_0 = (x_{01}, x_{02})$ to be such that $\sin 2\pi x_{02} \neq 0$. Then one has the asymptotics given in Table 1.

TABLE 1. ASYMPTOTIC RESULTS FOR EXAMPLE 1

	Time scale	Asymptotic law
(1)	$1 \ll t \ll a^{2/3}$	$\dfrac{X_1(t) - X_1(0) - t(c_0 + c_2 \cos 2\pi x_{02})}{\sqrt{t}} \xrightarrow{\mathcal{L}} \mathcal{N}\Big(0, 1 + \dfrac{c_1^2}{2(\delta^2 + \pi^2)}\Big)$
(2)	$\dfrac{t}{a^{2/3}} \to r > 0$	$\dfrac{X_1(t) - X_1(0) - t(c_0 + c_2 \cos 2\pi x_{02})}{\sqrt{t}} \xrightarrow{\mathcal{L}} \mathcal{N}\Big(-c_2 \delta r^{3/2} \pi \sin 2\pi x_{02}, 1 + \dfrac{c_1^2}{2(\delta^2 + \pi^2)}\Big)$
(3)	$t \gg a^2$	$\dfrac{X_1(t) - X_1(0) - tc_0}{\sqrt{t}} \xrightarrow{\mathcal{L}} \mathcal{N}\Big(0, 1 + \dfrac{c_1^2}{2(\pi^2 + \delta^2)} + \dfrac{c_2^2}{2\delta^2}\Big)$

EXAMPLE 2. Same as in Example 1, except that $\delta = 0$ and $\sin 2\pi x_{02} = 0$. The asymptotic results are given in Table 2. The sign $\pm$ is $+$ if $\cos 2\pi x_{02} = -1$ and $-$ if $\cos 2\pi x_{02} = +1$.

2. Speed of Convergence to Equilibrium for Diffusions on a Big Torus

Consider a diffusion $X(\cdot)$ on $\mathbb{R}^k (k > 1)$ governed by Itô's stochastic integral equation

$$X(t) = X(0) + \int_0^t \{b(X(s)) + \beta(X(s)/a)\} ds + \sigma B(t) \tag{2.1}$$

TABLE 2. ASYMPTOTIC RESULTS FOR EXAMPLE 2

	Time scale	Asymptotic law
(1)	$1 \ll t \ll a^{4/3}$	$\frac{X_1(t)-X_1(0)-t(c_0+c_2\cos 2\pi x_{02})}{\sqrt{t}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1+c_1^2/2\pi^2)$
(2)	$a^{4/3} \ll t \ll a^2$	$\frac{X_1(t)-X_1(0)-t(c_0+c_2\cos 2\pi x_{02})}{t^2/a^2} \xrightarrow{\mathcal{L}} \mathcal{L}\big(\pm 2c_2\pi^2 \int_0^1 B_2^2(s)ds\big)$
(3)	$t/a^2 \to r > 0$	$\frac{X_1(t)-X_1(0)-tc_0}{\sqrt{t}} \xrightarrow{\mathcal{L}} \mathcal{L}\Big(\frac{c_2}{r} + \int_0^r \cos(2\pi(x_{02}+B_2(s)))ds\Big)$
(4)	$t \gg a^2$	$\frac{X_1(t)-X_1(0)-tc_0}{a\sqrt{t}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,c_2^2/2\pi^2)$

where it is assumed that

(A1) $b(\cdot)$, $\beta(\cdot)$ *are infinitely differentiable vector fields on* $\mathbb{R}^k$ *which are periodic with period lattice* $\mathbb{Z}^k$ *and are divergence-free, i.e.,*

$$\sum_{i=1}^{k} \frac{\partial}{\partial x_i} b_i(x) = 0 = \sum_{i=1}^{k} \frac{\partial}{\partial x_i} \beta_i(x); \tag{2.2}$$

(A2) σ *is a nonsingular constant* $k \times k$ *matrix;*

(A3) *'a' is a positive integer.*

Note that $b(\cdot)$ and $\beta(\cdot/a)$ are periodic with period lattice $a\mathbb{Z}^k$ and

$$\dot{X}(t) := X(t) \mod a \quad \equiv (X_1(t) \mod a, \ldots, X_k(t) \mod a) \tag{2.3}$$

is a diffusion on the torus $\mathcal{T}_a = \{x \mod a : x \in \mathbb{R}^k\}$. Let D be the diffusion matrix, $D = \sigma\sigma'$. Then the infinitesimal generator of $\dot{X}(t)$ is

$$L_a = \frac{1}{2}\sum_{i,j=1}^{k} D_{ij} \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{k} (b_i(x) + \beta_i(x/a)) \frac{\partial}{\partial x_i} \tag{2.4}$$

operating on smooth functions on $\mathcal{T}_a$. In view of the divergence-free property (2.2), it is simple to check that the uniform distribution π_a on $\mathcal{T}_a$ is the unique invariant probability on $\mathcal{T}_a$. By identifying $\mathcal{T}_a$ with $[0,a)^k, \pi_a$ has density $1/a^k$ w.r.t. Lebesgue measure on $[0,a)^k$. Let $L^1(\mathcal{T}_a), L^2(\mathcal{T}_a)$ denote the space of integrable and square integrable complex-valued functions on $\mathcal{T}_a$, respectively. We will denote by $\| \cdot \|$ and $\langle \ , \ \rangle$ the norm and inner product on $L^2(\mathcal{T}_a) : \langle f, g\rangle = \int_{[0,a)^k} f(x)g^-(x)dx/a^k$. Let $\{T_{t,a} : t \geq 0\}$ be the semigroup of transition operators: $T_{t,a}f(x) = \int_{\mathcal{T}_a} f(y)p_a(t;x,y)dy \ \forall f \in L^1(\mathcal{T}_a)$, where $p_a(t;x,y)$ is the transition probability density of $\dot{X}(t)$ w.r.t. Lebesgue measure on $\mathcal{T}_a$. Let $1^\perp$ denote the subspace of $L^2(\mathcal{T}_a)$ comprising functions of mean zero under π_a.

PROPOSITION 2.1. *Let $\dot{X}(0)$ have the equilibrium distribution π_a. Then, for all $f, g, \in 1^\perp$ and all $s \geq 0, t \geq 0$, one has*

$$|Ef(\dot{X}(s))g(\dot{X}(s+t))| \leq \|f\| \ \|g\|e^{-\lambda t} \tag{2.5}$$

where λ satisfies

$$\frac{2\pi^2\alpha_1}{a^2} \leq \lambda \leq \frac{2\pi^2 d_m}{a^2} \tag{2.6}$$

with α_1 as the smallest eigenvalue of the diffusion matrix D, and $d_m =$ $\min\{D_{jj} : 1 \leq j \leq k\}$.

PROOF. Note that

$$|Ef(\dot{X}(s))g(\dot{X}(s+t))| = |\langle f, T_{t,a}g\rangle| \leq \|f\| \ \|T_{t,a}g\|. \tag{2.7}$$

Now, as shown in Bhattacharya (1999, Proposition 4.1 and relation (4.9)),

$$\|T_{t,a}g\|^2 \leq \|g\|^2 e^{-2\lambda t}$$

where λ is the spectral gap of $1/2\sum D_{ij}\partial^2/\partial x_i \partial x_j$ on $L^2(\mathcal{T}_1)$. The inequalities (2.5) are derived in Bhattacharya (1999, Corollary 4.4). □

The spectral method may also be used to provide an estimation of the total variation distance between $p(t;x,y)dy$ and π_a of the order $O(a^{k/2}e^{-\lambda_a t})$ (Bhattacharya, 1999, Theorem 4.5). A better estimate for this however has been recently obtained by Franke (2001). To state this results consider a diffusion $Z(t)$ on $\mathbb{R}^k$ with an arbitrary continuously differentiable and divergence-free periodic drift $b(\cdot)$ with period lattice $\mathbb{Z}^k$, and a fixed positive definite diffusion matrix D. Let $q_b(t;x,y)$ denote the transition probability density of the diffusion $\dot{Z}(t) := Z(t) \mod 1$ on the unit torus $\mathcal{T}_1$. Then one has the following uniform Doeblin minorization:

$$\inf_b \min\{q_b(1;x,y) : x, y \in \mathcal{T}_1\} > 0 \ , \tag{2.8}$$

where the infimum is over all divergence-free periodic vector fields. Applying this to the diffusion $\dot{Y}(t)$ on $\mathbb{R}^k$ defined by

$$Y(t) = \frac{X(a^2 t)}{a}, \ \dot{Y}(t) = Y(t) \mod 1, \tag{2.9}$$

one obtains

PROPOSITION 2.2. (Franke, 2001). *Under the assumptions (A1)-(A3), there exist positive constants c_1, c_2 independent of 'a' such that*

$$\int_{\mathcal{T}_a} |p_a(t;x,y) - \frac{1}{a^k}|dy \le c_1 \exp\{-c_2 t/a^2\}. \tag{2.10}$$

3. Time Scales for the Initial and Final Gaussian Phases

It follows as a special case of Theorem 1.1 that, for times $1 \ll t \ll a^{2/3}$, the solution of

$$X(t) = ax_0 + \int_0^t \{b(X(s)) + \beta(X(s)/a)\}ds + \sigma B(t) \tag{3.1}$$

is close (in total variation distance) to that of

$$X_0(t) = ax_0 + \int_0^t \{b(X_0(s)) + \beta(x_0)\}ds + \sigma B(t), \tag{3.2}$$

under broad assumptions on $b(\cdot)$, $\beta(\cdot)$, provided the constant matrix σ is nonsingular. Assume in addition that $b(\cdot)$ is periodic with period lattice $\mathbb{Z}^k$ and is divergence-free. Then the invariant distribution π is uniform and

$$\frac{X_0(t) - ax_0 - t(\overline{b} + \beta(x_0))}{\sqrt{t}} \stackrel{\mathcal{L}}{\to} N(0, K_0) \tag{3.3}$$

where

$$\begin{aligned} \overline{b} &= \int_{[0,1)^k} b(x)dx, \\ K_0 &= \int_{[0,1)^k} [\text{Grad } \Psi_0(y) - I_k]D[\text{Grad } \Psi_0(y) - I_k]'dy. \end{aligned} \tag{3.4}$$

Here $\Psi_0 = (\Psi_{01}, \Psi_{02}, \ldots, \Psi_{0k})'$, Grad $\Psi_0 = (\text{grad}\Psi_{01}, \ldots, \text{grad}\Psi_{0k})'$, and Ψ_{0j} is the unique mean-zero periodic solution of

$$\begin{aligned} L^0\Psi_{0j}(y) &= b_j(y) - \overline{b}_j, \qquad (1 \le j \le k), \\ L^0 &:= \frac{1}{2}\sum_{i,j=1}^k D_{ij}\frac{\partial^2}{\partial y_i \partial y_j} + \sum_{i=1}^k \{b_i(y) + \beta_i(x_0)\}\frac{\partial}{\partial y_i} \end{aligned} \tag{3.5}$$

Since the integral of Grad $\Psi_0(y)D$ in (3.4) vanishes due to periodic boundary conditions, K_0 reduces to

$$K_0 = \int_{[0,1)^k} \text{Grad } \Psi_0(y)D(\text{Grad } \Psi_0(y))'dy + D. \tag{3.6}$$

We are then led to

THEOREM 3.1. *Assume (i) $b(\cdot)$ is continuously differentiable, periodic with period lattice $\mathbb{Z}^k$, and divergence-free, (ii) $\beta(\cdot)$ and its first and second order derivatives are bounded, and (iii) σ is a nonsingular constant matrix. Then for times $1 \ll t \ll a^{2/3}$, i.e., as*

$$t \to \infty, \quad \frac{t}{a^{2/3}} \to 0, \tag{3.7}$$

one has

$$\frac{X(t) - ax_0 - t(\overline{b} + \beta(x_0))}{\sqrt{t}} \xrightarrow{\mathcal{L}} N(0, K_0). \tag{3.8}$$

Example 1 in Section 1 shows that the time scale (3.7) for this initial Gaussian phase is in general precise, in the absence of further assumptions (in addition to (A1)-(A3)).

We now turn to our main goal, namely, that of determining the nature and time scale of the final Gaussian phase, under the assumptions (A1)-(A3). For each fixed 'a', the solution $X(t)$ of (3.1) is asymptotically Gaussian (Bensoussan et al., 1978, Bhattacharya, 1985),

$$\frac{X(t) - ax_0 - t(\overline{b} + \overline{\beta})}{\sqrt{t}} \xrightarrow{\mathcal{L}} N(0, K) \tag{3.9}$$

as $t \to \infty$, where $\overline{b}, \overline{\beta}$ are as in (3.4), and

$$\begin{aligned} K &= \int_{[0,a)^k} [\text{Grad } \Psi(y) - I] D [(\text{Grad } \Psi(y) - I]' \frac{1}{a^k} dy \\ &= \int_{[0,a)^k} \text{Grad } \Psi(y) D (\text{Grad } \Psi(y))' \frac{1}{a^k} dy + D, \end{aligned} \tag{3.10}$$

$\Psi = (\Psi_1, \ldots, \Psi_k)$ being the unique mean-zero periodic (with period lattice $a\mathbb{Z}^k$) solution of

$$L_a \Psi_j(x) = b_j(x) + \beta_j(x/a) - \overline{b}_j - \overline{\beta}_j \quad (1 \le j \le k). \tag{3.11}$$

Here L_a is as given in (2.4).

The rest of this section is devoted to the task of determining the precise time scale for the validity of this Gaussian approximation (3.9), when 'a' is large. For this we need

(3.12)

(1) to analyze the asymptotic behavior of the dispersion matrix K as $a \to \infty$, and

(2) to prove that as $a \to \infty$ and t is sufficiently large relative to a, (3.9) takes hold.

To accomplish this it is convenient to rescale $X(t)$ as

$$Y(t) = \frac{X(a^2 t)}{a}\ ,\ \dot{Y}(t) = Y(t) \mod 1. \tag{3.13}$$

Then $Y(t)$ is governed by Itô's equation

$$Y(t) = Y(0) + \int_0^t a\{b(aY(s)) + \beta(Y(s))\}ds + \sigma\overline{B}(t), \tag{3.14}$$

where $\{\overline{B}(t) := B(a^2t)/a : t \geq 0\}$ is again a standard k-dimensional Brownian motion. Note that $\dot{Y}(\cdot)$ is a diffusion on the *unit torus* $\mathcal{T}_1 = \{x \mod 1 : x \in \mathbb{R}^k\}$, whose infinitesimal generator is

$$A_a = \frac{1}{2}\sum_{i,j=1}^{k} D_{ij}\frac{\partial^2}{\partial y_i \partial y_j} + a\{b(ay) + \beta(y)\} \cdot \nabla \tag{3.15}$$

The solution to (3.11) may be expressed as

$$\Psi_j(x) = a^2 g_j(x/a) \qquad (1 \leq j \leq k) \tag{3.16}$$

where g_j is the unique mean zero element of $L^2(\mathcal{T}_1)$ satisfying

$$A_a g_j(y) = b_j(ay) - \overline{b}_j + \beta_j(y) - \overline{\beta}_j. \tag{3.17}$$

Note that, if one writes

$$\begin{aligned} E_{jj'} &= \frac{1}{a^k}\int_{[0,1)^k} \nabla\Psi_j(x) \cdot D\nabla\Psi_{j'}(x)dx \\ &= a^2 \int_{[0,1)^k} \nabla g_j(y) \cdot D\nabla g_{j'}(y)dy, \end{aligned} \tag{3.18}$$

then, by (3.10), the elements of the dispersion matrix K may be expressed as

$$K_{jj'} = E_{jj'} + D_{jj'} \qquad (1 \leq j, j' \leq k). \tag{3.19}$$

It is convenient to introduce the complex Hilbert spaces

$$\begin{aligned} H^0 &= \{f \in L^2(\mathcal{T}_1) : \int_{[0,1)^k} f(x)dx = 0\} = 1^{\perp}, \\ H^1 &= \{f \in H^0 : \int_{[0,1)^k} |\Delta f(x)|^2 dx < \infty\}, \end{aligned} \tag{3.20}$$

with respective inner products

$$\langle f, g\rangle_0 = \int_{[0,1)^k} f(x)g^-(x)dx, \ \langle f, g\rangle_1 = \int_{[0,1)^k} \nabla f(x)\cdot D\nabla g^-(x)dx \quad (3.21)$$

As a subspace of H^0, H^1 is the closure of $C^\infty(\mathcal{T}_1)\cap 1^\perp$ with respect to the norm $\|f\|_1 = (\langle f, f\rangle_1)^{1/2}$. Note that

$$\langle f, g\rangle_1 = -\langle \mathcal{D}f, g\rangle_0, \quad \mathcal{D} := \frac{1}{2}\sum_{i,j=1}^{k} D_{ij}\frac{\partial^2}{\partial y_i \partial y_j} \quad (3.22)$$

for all twice continuously differentiable f in H^0 and all $g \in H^1$. It is simple to check (see Bhattacharya, 1999 and (3.56)) that

$$\|f\|_0^2 \le \|f\|_1^2/(2\pi^2\alpha_1) \qquad (f \in H^1) \quad (3.23)$$

where α_1 is the smallest eigenvalue of D. We may now rewrite (3.18), (3.19) as

$$\begin{aligned} E_{jj'} &= a^2\langle g_j, g_{j'}\rangle_1, \\ K_{jj'} &= E_{jj'} + D_{jj'}, \qquad (1 \le j, j' \le k). \end{aligned} \quad (3.24)$$

The proof of the following Lemma may be found in Bhattacharya (1999), Lemma 3.5.

LEMMA 3.2. *Under the assumptions (A1)-(A3) in Section 2, one has for all* $j, 1 \le j \le k$,

$$\sup \|g_j\|_1^2 < \infty, \qquad \|g_j\|_0^2 \le \frac{1}{2\pi^2\alpha_1}\|g_j\|_1^2$$

where α_1 *is the smallest eigenvalue of* D. *The supremum is over all positive integers 'a'. Also,*

$$|\langle f(a\cdot), g\rangle| \le \frac{\|f\|_0\|g\|_1}{\alpha_1 a} \quad \forall f \in H^0, \quad g \in H^1, \quad (3.25)$$

where $f(a\cdot)$ *is the function* $f(ax), x \in \mathcal{T}_1$.

To proceed further in the analysis of K, express g_j as (see (3.15), (3.17))

$$\begin{aligned} &(\mathcal{I} + aS_a)g_j = \mathcal{D}^{-1}(b_j(a\cdot) - \overline{b}_j + \beta_j(\cdot) - \overline{\beta}_j), \\ &\mathcal{I} = \text{identity operator on } H^1, \quad S_a = \mathcal{D}^{-1}(b(a\cdot) + \beta(\cdot))\cdot\nabla, \end{aligned} \quad (3.26)$$

where $\mathcal{D}$ as defined in (3.22) acts on the space H^2 of functions in H^0 with second order derivatives in the L^2-sense. That is, H^2 comprises those elements f of H^0 such that the $\sum_{r\in\mathbb{Z}^k}(|r|^2\hat{f}(r)|)^2 < \infty$, where $\hat{f}$ is the Fourier transform of f. Since $b(a\cdot)$ is rapidly oscillating on $\mathcal{T}_1$, in view of (3.25) one may think of approximating it by $\overline{b}$ and approximate S_a by

$$\overline{S} = \mathcal{D}^{-1}(\overline{b} + \beta(\cdot)) \cdot \nabla. \tag{3.27}$$

Then $\overline{S}$ (as well as S_a) is a skew symmetric and compact operator of H^1 (see Bhattacharya, 1999, Proposition 3.2), and as such has a sequence of eigenvalues $i\lambda_n$ (λ_n real) and a corresponding sequence of orthonormal eigenfunctions $\varphi_n (n = 1, 2, \ldots)$ such that $\lambda_n \to 0$ as $n \to \infty$, and the linear span of $\{\varphi_n : n \geq 1\}$ is the range $\mathcal{R}$ of $\overline{S}$. The null space N of $\overline{S}$ then satisfies $N^{\perp} = \overline{\mathcal{R}}$ (the closure of $\mathcal{R}$ in H^1). Denote by $f_N, f_{N^\perp}$ orthogonal projections (in H^1) of f onto N and onto $N^\perp$, respectively. We will make the following assumption for the next result.

(A4) $\{(\mathcal{D}^{-1}(\beta_j \cdot - \overline{\beta}_j))_N, 1 \leq j \leq p\}$ *is a linearly independent subset of* H^1.

PROPOSITION 3.3. *Assume (A1)-(A3), and A4 for some* $p (1 \leq p \leq k)$. *In addition, suppose that* $(\mathcal{D}^{-1}(\beta_j \cdot - \overline{\beta}_j))_N$ *is twice continuously differentiable for* $1 \leq j \leq p$. *Then the* lim inf *as* $a \to \infty$ *of the smallest eigenvalue of the matrix* $((\langle g_j, g_{j'}\rangle_1))_{1\leq j,j'\leq p}$ *is positive.*

PROOF. Define an approximation of A_a, namely,

$$\overline{A} = \mathcal{D} + a(\overline{b} + \beta(\cdot)) \cdot \nabla = \mathcal{D}(\mathcal{I} + a\overline{S}), \tag{3.28}$$

and corresponding approximation h_j of g_j, which is the unique element of H^2 satisfying

$$(\mathcal{I} + a\overline{S})h_j = \mathcal{D}^{-1}(\beta_j(\cdot) - \overline{\beta}_j) \quad (1 \leq j \leq k). \tag{3.29}$$

By expressing h_j as $(h_j)_N + (h_j)_{N^\perp}$, and using the eigenfunction expansion of $(h_j)_{N^\perp}$ in terms of $\{\varphi_n : n \geq 1\}$, one can show (see Bhattacharya, 1999, relation (3.75))

$$h_j \to (\mathcal{D}^{-1}(\beta_j - \overline{\beta}_j))_N \text{ in } H^1 \text{ as } a \to \infty \quad (1 \leq j \leq k). \tag{3.30}$$

We now compare g_j and h_j. First note that

$$g_j - h_j = \mathcal{D}^{-1}(b_j(a\cdot) - \overline{b}_j) - a\mathcal{D}^{-1}\{b(a\cdot) - \overline{b}) \cdot \nabla g_j\}, \tag{3.31}$$

and

$$\langle g_j - h_j, \varphi_n\rangle_1 = \frac{1}{1+ia\lambda_n}\langle \mathcal{D}^{-1}(b_j(a\cdot) - \overline{b}_j), \varphi_n\rangle_1 - \frac{a}{1+ia\lambda_n}\langle \mathcal{D}^{-1}\{(b(a\cdot) - \overline{b})\cdot\nabla g_j\}, \varphi_n\rangle_1. \quad (3.32)$$

The first term on the right in (3.32) equals $-(1+ia\lambda_n)^{-1}\langle b_j(a\cdot) - \overline{b}_j, \varphi_n\rangle_0$ which goes to zero as $a\to\infty$. To estimate the second term, write

$$\begin{aligned}\langle \mathcal{D}^{-1}\{(b(a\cdot) - \overline{b})\cdot\nabla g_j\}, \varphi_n\rangle_1 &= -\langle (b(a\cdot) - \overline{b})\cdot\nabla g_j, \varphi_n\rangle_0\\ &= -\sum_{s=1}^{k}\left\langle \frac{\partial}{\partial x_s}\{(b_s(ax) - \overline{b}_s)g_j(x)\}, \varphi_n\right\rangle_0\\ &= -\sum_{s=1}^{k}\langle b_s(ax) - \overline{b}_s, g_j^- \frac{\partial\varphi_n(x)}{\partial x_s}\rangle_0\end{aligned} \quad (3.33)$$

Since φ_n solves $\overline{S}\varphi_n = \lambda_n\varphi_n$, or $(\overline{b}+\beta(\cdot))\cdot\nabla\varphi_n = \lambda_n\mathcal{D}\varphi_n$, by standard regularity theory (see, e.g., Reed and Simon, 1975, p.112) φ_n is infinitely differentiable, so that $\{g_j^- \frac{\partial\varphi_n(x)}{\partial x_s} : a = 1,2,\dots\}$ is a bounded subset of H^1 by Lemma 3.2 and the inequality (Bhattacharya, 1999, inequality (3.85))

$$\|uv\|_1^2 \le c(\|u\|_1^2\cdot\|v\|_\infty^2 + \|u\|_0^2\|\nabla v\|_\infty^2), \quad (3.34)$$

applied to $u = g_j^-$ and $v(x) = \partial\varphi_n(x)/\partial x_s$. Here c is a positive constant which only depends on D. Hence by (3.25), the term (3.33) is of the order $O(1/a)$. Thus $\langle g_j - hj, \varphi_n\rangle_1$ goes to zero as $a\to\infty$, uniformly for all n. It follows that $(g_j - h_j)_{N^\perp}$ converges to zero weakly in H^1. Now note that $\langle f, S_a f\rangle_1 = 0\ \forall f\in H^1$, since S_a is skew symmetric. Hence

$$\begin{aligned}\|g_j\|_1^2 &= \langle g_j, (\mathcal{I}+aS_a)g_j\rangle_1 = \langle g_j, \mathcal{D}^{-1}\{b_j(a\cdot) - \overline{b}_j + \beta_j(\cdot) - \overline{\beta}_j\}\rangle_1\\ &= -\langle g_j, b_j(a\cdot) - \overline{b}_j + \beta_j(\cdot) - \overline{\beta}_j\rangle_0 = -\langle g_j, \beta_j(\cdot) - \overline{\beta}_j\rangle_0 + o(1)\\ &= \langle g_j, \mathcal{D}^{-1}(\beta_j(\cdot) - \overline{\beta}_j)\rangle_1 + o(1),\ \text{as } a\to\infty.\end{aligned} \quad (3.35)$$

Now, since $(g_j - h_j)_{N^\perp}\to 0$ weakly in H^1 and $\|h_{N^\perp}\|_1\to 0$ as $a\to\infty$, it follows that $(g_j)_{N^\perp}\to 0$ weakly. Hence

$$\begin{aligned}\|g_j\|_1^2 &= \langle g_j, \mathcal{D}^{-1}(\beta_j(\cdot) - \overline{\beta}_j)\rangle_1 + o(1) = \langle g_j, (\mathcal{D}^{-1}(\beta_j(\cdot) - \overline{\beta}_j))_N\rangle_1 + o(1)\\ &= \langle h_j, (\mathcal{D}^{-1}(\beta_j(\cdot) - \overline{\beta}_j))_N\rangle_1 + \langle g_j - h_j, (\mathcal{D}^{-1}(\beta_j(\cdot) - \overline{\beta}_j))_N\rangle_1 + o(1)\\ &= \|(\mathcal{D}^{-1}(\beta_j(\cdot) - \overline{\beta}_j))_N\|_1^2 + \langle g_j - h_j, (\mathcal{D}^{-1}(\beta_j(\cdot) - \overline{\beta}_j))_N\rangle_1 + o(1).\end{aligned} \quad (3.36)$$

By (3.31)

$$
\begin{aligned}
&\langle g_j - h_j, (\mathcal{D}^{-1}(\beta_j(\cdot) - \overline{\beta}_j))_N\rangle_1 \\
&\quad = -\langle b_j(a\cdot) - \overline{b}_j, (\mathcal{D}^{-1}(\beta_j(\cdot) - \overline{\beta}_j))_N\rangle_0 \\
&\qquad + a\langle (b(a\cdot) - \overline{b}) \cdot \nabla g_j, (\mathcal{D}^{-1}(\beta_j(\cdot) - \overline{\beta}_j))_N\rangle_0 \\
&\quad = a\sum_{s=1}^{k} \left\langle \frac{\partial}{\partial x_s}\{(b_s(a\cdot) - \overline{b}_s)g_j\}, (\mathcal{D}^{-1}(\beta_j(\cdot) - \overline{\beta}_j))_N \right\rangle_0 + o(1) \\
&\quad = -a\sum_{s=1}^{k} \left\langle (b_s(a\cdot) - \overline{b}_s), g_j \frac{\partial}{\partial x_s}(\mathcal{D}^{-1}(\beta_j(\cdot) - \overline{\beta}_j))_N \right\rangle_0 + o(1).
\end{aligned} \tag{3.37}
$$

By hypothesis, $v_s := \partial/\partial x_s(\mathcal{D}^{-1}(\beta_j(\cdot) - \overline{\beta}_j))_N$ is continuously differentiable, so that (3.23), (3.25) and (3.34) yield

$$
|\langle g_j - h_j, (\mathcal{D}^{-1}(\beta_j(\cdot) - \overline{\beta}_j))_N\rangle_1| \le c_1\|b\|_0 \ \|g_j\|_1 \le c_2\|g_j\|_1, \tag{3.38}
$$

for some constants c_1, c_2 independent of a. Using (3.38) in (3.36) one obtains

$$
\|g_j\|_1^2 \ge \|(\mathcal{D}^{-1}(\beta_j(\cdot) - \overline{\beta}_j))_N\|_1^2 - c_2\|g_j\|_1 + o(1), \tag{3.39}
$$

where $o(1) \to 0$ as $a \to \infty$. If $1 \le j \le p$, then the first term on the right side of (3.39) is positive, and (3.39) implies

$$
\liminf_{a\to\infty} \|g_j\|_1^2 > 0. \tag{3.40}
$$

One may apply the same argument to an arbitrary linear combination $\sum_{j=1}^{p} d_j g_j$ to arrive at the desired conclusion. □

We are now ready to prove one of the two main results of this article. Recall the assumptions (A1)-(A3) in Section 2 and the assumption (A4) stated before the statement of Proposition 3.3. The theorem below refers to the solution $X(t) = (X_1(t), \dots X_k(t))$ of (3.1). To state it, write

$$
\begin{aligned}
Z_j(t) &= j - th \text{ coordinate of } X(t) - X(0) - t(\overline{b} + \overline{\beta}), \\
Z(t) &= (Z_1(t), \dots, Z_p(t)), \quad K_1^p = ((K_{jj'}))_{1 \le j,j' \le p}, \\
\Gamma &= \frac{1}{a^2}\ K_1^p.
\end{aligned} \tag{3.41}
$$

THEOREM 3.4. *Assume (A1)-(A4). In addition, suppose* $(\mathcal{D}^{-1}(\beta_j(\cdot) - \overline{\beta}_j))_N$ *is twice continuously differentiable,* $1 \le j \le p$. *Then the following hold.*

(a) The smallest eigenvalue of Γ *is bounded away from zero and the largest eigenvalue of* Γ *is bounded away infinity, as* $a \to \infty$. *(b) For the time scale*

$$\frac{t}{a^2} \to \infty \quad \text{as } a \to \infty, \tag{3.42}$$

one has

$$\frac{1}{a\sqrt{t}}\, \Gamma^{-1/2} Z(t) \xrightarrow{\mathcal{L}} N(0, I_p), \tag{3.43}$$

where $\Gamma^{-1/2}$ *is the positive definite symmetric square root of* Γ^{-1} *and* I_p *is the* $p \times p$ *identity matrix.*

Proof. Part (a) follows from Lemma 3.2 and Proposition 3.3.

The proof of part (b) is along the lines of the proof of Theorem 5.2 in Bhattacharya (1999). Recall the process $Y(t) = X(a^2 t)/a$ and $\dot{Y}(t) = Y(t)$ mod 1. Fix a $j, 1 \le j \le p$. Then the convergence

$$\frac{1}{a\sqrt{t}\sqrt{\Gamma_{jj}}} Z_j(t) \xrightarrow{\mathcal{L}} N(0,1) \quad \text{as } \frac{t}{a^2} \to \infty, \tag{3.44}$$

is equivalent to the convergence

$$\frac{1}{a\sqrt{t}\sqrt{\Gamma_{jj}}}\{Y_j(t) - Y_j(0) - a\, t(\overline{b}_j + \overline{\beta}_j)\} \xrightarrow{\mathcal{L}} N(0,1) \tag{3.45}$$

as

$$a \to \infty, \quad t \to \infty. \tag{3.46}$$

To prove (3.45), use (3.14) to drop the term involving $\overline{B}(t)$ to show that (3.45) is equivalent to

$$\frac{1}{\sqrt{t}\sqrt{\Gamma_{jj}}} \int_0^t \{b_j(a\dot{Y}(s)) - \overline{b}_j + \beta_j(\dot{Y}(s)) - \overline{\beta}_j\} ds \xrightarrow{\mathcal{L}} N(0,1) \tag{3.47}$$

under (3.46).

First assume $\dot{Y}(0)$ has the uniform (equilibrium) distribution on $\mathcal{T}_1$. Let $\varphi(a)$ $(a = 1, 2, \dots)$ be an arbitrary sequence of positive integers increasing to infinity, and write the left side of (3.47) for $t = \varphi(a)$ as

$$\sum_{r=1}^{\varphi(a)} V_r, \quad V_r := \frac{1}{\sqrt{\varphi(a)\Gamma_{jj}}} \int_{r-1}^{r} \{b_j(a\dot{Y}(s)) - \overline{b}_j + \beta_j(\dot{Y}(s)) - \overline{\beta}_j)\} ds. \tag{3.48}$$

By Itô's Lemma,

$$V_r = \frac{1}{\sqrt{\varphi(a)\Gamma_{jj}}}[g_j(\dot{Y}(r)) - g_j(\dot{Y}(r-1)) - \int_{r-1}^{r} \text{grad } g_j(\dot{Y}(s))\sigma d\overline{B}(s)]. \tag{3.49}$$

It follows that

$$EV_r = 0,$$

and, since $E[g_j(\dot{Y}(r) - g_j(\dot{Y}(r-1))]^2 \leq c\|g_j\|_0^2 \leq c$ where c is a finite constant independent of 'a',

$$\begin{aligned} E\Big(\sum_1^{\varphi(a)} V_r\Big)^2 &= E\Big[\sum_{r=1}^{\varphi(a)} (\varphi(a)\Gamma_{jj})^{-1/2} \int_{r-1}^{r} \text{grad } g_j(\dot{Y}(s)) \cdot \sigma d\overline{B}(s)\Big]^2 + o(1) \\ &= \frac{1}{\varphi(a)} \sum_{r=1}^{\varphi(a)} \frac{1}{\Gamma_{jj}} \|g_j\|_1^2 = \|g_j\|_1^2/\Gamma_{jj} \to 1 \text{ as } a \to \infty, \end{aligned} \tag{3.50}$$

by (3.24) and (3.41).

The proof of the asymptotic normally of $\sum_{r=1}^{\varphi(a)} V_r$ now may be given by the classical method of splitting into sums over big blocks, separated by sums over relatively small blocks, which may be ignored. For this define

$$\begin{aligned} &\delta = [\varphi(a)^{1/4}], \eta = [\delta^{1/8}\varphi^{1/16}(a)], \Psi = [\delta^{1/8}\varphi^{1/4}(a)], \\ &m = \Big[\frac{\varphi(a)}{\eta + \Psi}\Big], \end{aligned} \tag{3.51}$$

where $[y]$ denotes the integer part of y. Define the 'big' block sums

$$U_n = \sum_{r=1}^{\Psi(a)} V_{r+(n-1)(\Psi+\eta)} \qquad (n = 1, \dots, m), \tag{3.52}$$

and the 'little' block sums

$$\xi_n = \sum_{r=1}^{\eta} V_{r+n(\Psi+\eta)-\eta} \qquad (n = 1, \dots, m) \tag{3.53}$$

Then denoting by $y \simeq z$ the relating $y - z \to 0$ in probability as $a \to \infty$,

$$\sum_{r=1}^{\varphi(a)} V_r \simeq \sum_{n=1}^{m} U_n + \sum_{n=1}^{m} \xi_n, \tag{3.54}$$

which follows from the fact that the right side is missing no more than the last $\Psi+\eta$ terms V_r from the left side. By applying the convergence in (3.50), but with $\Psi+\eta$ in place of φ, the expected value of the square of the sum of the missing terms is of the order $\mathrm{O}((\Psi+\eta/\varphi)\to 0$ as $a\to\infty$. By a similar argument,

$$E\xi_n^2 \le c_1\eta/\varphi(a), \quad \sum_{n=1}^{m} E\xi_n^2 \le c_2\ m\eta/\varphi(a) \to 0. \tag{3.55}$$

Also,

$$\begin{aligned} E\xi_n\xi_{n+n'} &= \sum_{i=1}^{\eta}\sum_{i'1}^{\eta} E(V_{i+\Psi}V_{i'} + V_{i'+(n'+1)(\Psi+\eta)-\eta}) \\ &= \frac{1}{\varphi(a)}\sum_{i,i'=1}^{\eta}\int_0^1 \langle f_1, \tilde{T}_{i'-i-1+n'(\Psi+\eta)+s} f_2\rangle_0 ds, \end{aligned} \tag{3.56}$$

where $f_1(y) = b_j(ay) - \overline{b}_j + \beta_j(y) - \overline{\beta}_j$ and $f_2(y) = \int_0^1 \tilde{T}_s f_1(y)dy$, with $\tilde{T}_s$ as the transition operator for $\dot{Y}(\cdot)$ (i.e., $\tilde{T}_s f(y) = E[f(\dot{Y}(s))|\dot{Y}(0){=}y]$). By Proposition 2.1, the integral on the right in (3.56) is bounded by $\|f_1\|_0\ \|f_2\|_0 \exp\{-c_3 n'\Psi\}$, so that

$$\begin{aligned} |E\xi_n\xi_{n+n'}| &\le c_4\frac{\eta^2}{\varphi(a)}\exp\{-c_3n'\Psi\}, \\ \sum_{n=1}^{m}\sum_{n'=1}^{m}|E\xi_n\xi_{n+n'}| &\le c_5\frac{m\eta^2}{\varphi(a)}\exp\{-c_3\Psi\} \to 0. \end{aligned} \tag{3.57}$$

From (3.55), (3.57), one sees that $E(\sum_{n=1}^m \xi_n)^2 \to 0$, so that (3.54) becomes

$$\sum_{r=1}^{\varphi} V_r \simeq \sum_{n=1}^{m} U_n. \tag{3.58}$$

We now show that the characteristic function of the right side converges to that of the sum of m i.i.d. random variables each having the same distribution as U_1. For this write $f(y) = E[\exp\{i\theta U_1\}|\dot{Y}(0) = y]$. By Proposition 2.2 we then obtain the following relations:

$$\begin{aligned} &E\left|\exp\left\{i\theta\sum_{n=1}^{m}U_n\right\} - E\left(\exp\left\{i\theta\sum_{n=1}^{m-1}U_n\right\}\right)E\exp\{i\theta U_m\}\right| \\ &= \left|E[\exp\left\{i\theta\sum_{n=1}^{m-1}U_n\right\}(\tilde{T}_\eta f(\dot{Y}(r')) - \overline{f})]\right| \quad (r' := (m-1)(\Psi+\eta)-\eta) \\ &\le c_6\exp\{-c_3\eta\}. \end{aligned}$$

Continuing now with the sum $\sum_{n=1}^{m-1}$, etc., one finally gets

$$\left| E\exp\left\{ i\theta \sum_{n=1}^{m-1} U_n \right\} - \prod_{n=1}^{m} E\exp\{i\theta U_n\} \right| \leq c_7\ m\exp\{-c_3\eta\} \to 0. \tag{3.59}$$

It remains to apply Lindeberg's condition to the sum of m i.i.d. random variables ($m = m(a), a = 1, 2, \dots$). This is simple since $|U_n| \leq c_8\Psi/\sqrt{\varphi} \to 0$ as $a \to \infty$. Hence, for every $\varepsilon > 0$,

$$\sum_{n=1}^{m} E\left(U_n^2 1_{\{|U_n|>\varepsilon\}}\right) = 0 \text{ for all sufficiently large '}a\text{'}. \tag{3.60}$$

We have thus proved (3.45), when $\dot{Y}(0)$ has the equilibrium distribution.

Let now $\dot{Y}$ have an arbitrary initial distribution. Let $t = \varphi(a), s = \Psi(a)$ (as in (3.51)). Write

$$\begin{aligned} \frac{Y_j(t) - Y_j(0) - at(\overline{b}_j + \overline{\beta}_j)}{a\sqrt{t}\sqrt{\Gamma_{jj}}} &= \frac{Y_j(s) - Y_j(0) - as(\overline{b}_j + \overline{\beta}_j)}{a\sqrt{t}\sqrt{\Gamma_{jj}}} \\ &\quad + \frac{Y_j(t) - Y_j(s) - a(t-s)(\overline{b}_j + \overline{\beta}_j)}{a\sqrt{t}\sqrt{\Gamma_{jj}}}. \end{aligned} \tag{3.61}$$

Use (3.14) to see that the first term on the right goes to zero as $a \to \infty$. For the second term, note that the conditional distribution of $Y(t) - Y(s)$, given $\{Y(u) :\ 0 \leq u \leq s\}$, is the same as the distribution of $Y(t-s) - z$ with an initial state $z = \dot{Y}(s)$. By Proposition 2.2, the total variation distance between the distribution of $Y(t) - Y(s)$ under an arbitrary distribution of $\dot{Y}(0)$ is $O(\exp\{-\Psi\}) \to 0$ as $a \to \infty$. On the other hand, as proved above, under the equilibrium initial distribution, the second term on the right converges in distribution to $N(0,1)$ (Note that one may replace $\sqrt{t}$ by $\sqrt{t-s}$ in the denominator, as $\sqrt{\frac{t-s}{t}} \to 1$). This completes the proof of (3.45), and hence of (3.44). One may apply the above arguments to an arbitrary linear combination $\sum_{j=1}^{d} c_j X_j(t)$. □

REMARK 3.5. The technical condition that $(\mathcal{D}^{-1}(\beta_j(\cdot) - \overline{\beta}_j))_N$ is twice continuously differentiable ($1 \leq j \leq p$) is probably redundant, but we are unable to dispense with it. One simple, but important, sufficient condition for this is that $\mathcal{D}^{-1}(\beta_j(\cdot) - \overline{\beta}_j) \in N$ ($1 \leq j \leq p$).

Consider now the case complementary to (A4): $\mathcal{D}^{-1}(\beta_j(\cdot) - \overline{\beta}_j) \in N^{\perp} = \overline{\mathcal{R}}$. We will make a somewhat more restricted assumption:

(A5) *$\mathcal{D}^{-1}(\beta_j(\cdot) - \overline{\beta}_j), 1 \le j \le p$, are linearly independent elements of the range $\mathcal{R}$ of $\overline{S} = \mathcal{D}^{-1}(\overline{b} + \beta(\cdot)) \cdot \nabla$.*

The assumption (A5) implies that there exist $f_j \in H^1$ such that

$$(\overline{b} + \beta(\cdot)) \cdot \nabla f_j = \beta_j(\cdot) - \overline{\beta}_j \quad (1 \le j \le p). \tag{3.62}$$

The growth in dispersion under the assumption (A5) is dramatically smaller than under (A4), as seen from the following result in Bhattacharya (1999, Theorem 3.8.)

PROPOSITION 3.6. *Assume (A1)-(A3), (A5). Also suppose that the f_j in* (3.62) *are twice continuously differentiable. Then*

$$\limsup_{a\to\infty} a^2\|g_j\|_1^2 < \infty \qquad (1 \le j \le p), \tag{3.63}$$

so that the smallest eigenvalue $\lambda(a)$ and the largest eigenvalue $\Lambda(a)$ of the dispersion matrix $K_1^p = ((K_{jj'}))_{i\le j,j'\le p}$, satisfy

$$\lambda_1 \le \liminf \lambda(a) \le \limsup \Lambda(a) < \infty, \tag{3.64}$$

where λ_1 is the smallest eigenvalue of $D_1^p = ((D_{jj'}))_{1\le j,j'\le p}$.

This proposition and Propositions 2.1, 2.2, are crucial for the proof of our final result.

THEOREM 3.7. *In addition to the hypothesis of Proposition 3.6, assume that $\overline{\lim}_{a\to\infty}\|a\nabla g_j\|_\infty < \infty$. Then, for times*

$$\frac{t}{a^2} \to \infty \text{ as } a \to \infty, \tag{3.65}$$

the random vector $t^{-1/2}Z(t)$ defined by (3.41) *is asymptotically Gaussian with a dispersion matrix K_1^p satisfying $D_1^p \le K_1^p \le cI_p$ for some constant c independent of a :*

$$\frac{1}{\sqrt{t}}(K_1^p)^{-1/2}Z(t) \to N(0, I_p). \tag{3.66}$$

The proof of this follows along the lines of that of Theorem 3.4 (See also the proofs of Theorems 5.3, 5.4, and Remark 5.3.1 in Bhattacharya, 1999). The technical condition "$a\|\nabla g_j\|_\infty$ bounded away from infinity, $1 \le j \le p$", is awkward and, as shown in Bhattacharya (1999), pp. 1000-1001, it may be dispensed with over larger time scales $t \gg a^4$. Once again, the logarithmic factor $(\log a)^2$ appearing with the time scale in Bhattacharya (1999) is removed with the help of Proposition 2.2 due to Franke (2001).

One may note that the assumption that $b(\cdot)$ and $\beta(\cdot)$ are divergence-free are crucial for the method used to derive the growth in dispersivity with 'a' (Propositions 3.3, 3.6) as well as the time scale for the final Gaussian phase to take hold (Theorems 3.4, 3.7). Although this assumption is natural for the physical problem of solute transport in porous media discussed in Section 2, from a purely mathematical point of view one may inquire as to what happens if it is dropped. This is difficult to answer in general. But by explicit computation in one dimension ($k = 1$) one observes that dispersivity does not grow, and sometimes decreases to zero. In the latter case the time scale for the final Gaussian phase may be exponentially large in 'a' (Bhattacharya et al., 1999, Bhattacharya, 1999).

References

BENSOUSSAN A., LIONS, J.L. and PAPANICOLAOU, G.C. (1978). *Asymptotic Analysis for Periodic Structures.* North-Holland, Amsterdam.

BHATTACHARYA, R.N. (1982). On the functional central limit theorem and the law of iterated logarithm for Markov processes, *Z. Wahrsh. Verw. Gebiete*, **60**, 185-201.

BHATTACHARYA, R.N. (1985). A central limit theorem for diffusions with periodic coefficients, *Ann. Probab.*, **13**, 385-396.

BHATTACHARYA, R.N. (1999). Multiscale diffusion processes with periodic coefficients and an application to solute transport in porous media, *Ann. Appl. Probab.*, **9**, 951-1020.

BHATTACHARYA, R.N., DENKER, M. and GOSWAMI, A. (1999). Speed of convergence to equilibrium and to normality for diffusions with multiple periodic scales, *Stochastic Process. Appl.*, **80**, 55-86.

BHATTACHARYA, R.N. and GÖTZE, F. (1995). Time-scales for Gaussian approximation and its break down under a hierarchy of periodic spatial heterogeneities, *Bernoulli*, **1**, 81-123. Correction, ibid (1996), pp. 107-108.

BHATTACHARYA, R.N. and GUPTA, V.K. (1983). A theoretical explanation of solute dispersion in saturated porous media at the Darcy scale, *Water Resour. Res.*, **19**, 938-944.

BHATTACHARYA, R.N., GUPTA, V.K. and WALKER, H.F. (1989). Asymptotics of solute dispersion in periodic porous media, *SIAM J. Appl. Math*, **49**, 86-98.

CUSHMAN, J. (ed) (1990). *Dynamics of Fluid in Hierarchical Porous Media.* Academic Press, New York.

FRANKE, B. (2001). *Heat Content Inequalities for Diffusion on Manifolds.* Dissertation, Georg August Universität, Göttingen.

FRIED, J.J. and COMBARNOUS, M.A. (1971). Dispersion in porous media, *Adv. Hydosci.*, **7**, 169-282.

GLIMM, J. and SHARP, D.H. (1997). Multiscale science: a challenge for the 21st century, *SIAM News*, **30**, No.8.

GELHAR, L.W. and AXNESS, C.L. (1983). Three-dimensional stochastic analysis of macrodispersion in aquifers, *Water Resour. Res.*, **19**, 161-180.

IKEDA, N. and WATANABE, S. (1981). *Stochastic Differential Equations and Diffusion Processes.* North-Holland, New York.

PAPANICOLAOU, G.C. and VARADHAN, S.R.S. (1979). Boundary value problems with rapidly oscillating coefficients, *Colloq. Math. Soc. J'anos Bolyai*, **27**, 835-875.

REED, M. and SIMON, B. (1975). *Methods of Mathematical Physics, Vol. 1.* Academic Press, New York.

WINTER, C.L., NEWMAN, C.M. and NEUMAN, S.P. (1984). A perturbation expansion for diffusion in a random velocity field, *SIAM J. Appl. Math.*, **44**, 425-442.

RABI BHATTACHARYA
MATHEMATICS DEPARTMENT
INDIANA UNIVERSITY
BLOOMINGTON, IN 47405, USA
E-mail: bhattach@indiana.edu