

Asymptotic Analysis of Deterministic and Stochastic Equations with Rapidly Varying Components

George C. Papanicolaou*

Courant Institute, New York University, New York, USA

Werner Kohler

Department of Mathematics, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061, USA

Abstract. The asymptotic character of deterministic and stochastic equations whose solutions have a rapidly varying component is studied. Of particular interest is the class of problems for which the limiting behavior can be described in a contracted and simplified framework.

1. Introduction

The object of this paper is the asymptotic analysis of equations whose solutions have the following three features: (i) rapidly oscillating components, (ii) rapidly decaying components, (iii) rapidly varying stochastic components. The asymptotic limit that emerges can be described in a considerably smaller space and is much simpler than the original problem. This contraction of the description due to the three features above is what we seek to analyze.

Naturally, problems of this form have received considerable attention due to their frequent appearance in many different areas of physics and elsewhere. Many of our references contain in one form or another such problems. In [1] and [2] the passage from the linearized Boltzmann equation to the linearized Navier-Stokes equation is considered. This problem is reconsidered here in Section 2. Our analysis is similar to that of [3] and [4] and we employ it because it extends easily to stochastic problems. We refer also to the work on Generalized Master equations [5] and in particular to its mathematical development [6] which is very similar to our problem. The work of Kurtz [7] also aims in the same general direction as does the work on random evolutions [8, and references therein].

In Section 3 we consider the asymptotic behavior of problems in the form of stochastic Boltzmann equations. Such problems are of interest in the study of transport phenomena in random media [9] and elsewhere. In Section 4, we consider

* Part of this research was supported by the Air Force Office of Scientific Research Grant No. AFOSR-71-2013. Support by an Alfred P. Sloan Fellowship is also acknowledged. The work was completed while this author was a visiting member of IRIA, Rocquencourt, France.

a concrete application of the results of Section 3 to stochastic differential equations. We state the results in a more general form than direct application of the theorems of Section 3 implies. This more general form, and the proof, is a simple extension of the theorem given in [10] wherein additional references to work on the asymptotics of stochastic equations are given. We also give an example illustrating the theorem.

We note that the results of Section 4 may be considered as stochastic analogs of the results of [11, 12] on the analysis of stiff systems of ordinary differential equations.

2. Deterministic Problems of Boltzmann Form

Let $L_{k,m}$, $k, m = 0, 1, 2, \dots, N$ be Banach spaces with norms $\| \cdot \|_{k,m}$ such that for each m , $0 \leq m \leq N$,

$$L_{0,m} \supset L_{1,m} \supset \dots \tag{2.1}$$

and for each k , $0 \leq k \leq N$,

$$L_{k,0} \supset L_{k,1} \supset \dots \tag{2.2}$$

We assume that each space is dense in the preceding one in the above inclusions.

Let A and B be linear operators such that

$$A \in \mathcal{L}(L_{k,m}, L_{k-1,m-1}), \quad B \in \mathcal{L}(L_{k,m}, L_{k,m-1}). \tag{2.3}$$

Here $\mathcal{L}(\cdot, \cdot)$ denotes the collection of bounded linear operators from one space into another. We assume that B and $A + B/\varepsilon$, $\varepsilon > 0$, generate contraction semigroups on $L_{0,0}$ which we denote by e^{Bt} and by $e^{t(A+B/\varepsilon)}$ respectively. We assume furthermore that for all $t \in [0, \infty)$ these semigroups map $L_{k,m}$ into itself, $0 \leq k, m \leq N$, with bound independent of t and ε .

We are interested in the asymptotic behavior of $e^{t(A+B/\varepsilon)}$ as $\varepsilon \downarrow 0$, $0 \leq t \leq T$, and in related questions. For this purpose we introduce the following additional hypotheses concerning e^{tB} .

$$e^{tB} \rightarrow P, \quad t \uparrow \infty, \quad \text{in } L_{k,m}, \quad 0 \leq k, m \leq N. \tag{2.4}$$

Here P is the projection operator into the nullspace of B , that is, the closure of the set $\{f \in L_{0,1} | Bf = 0\}$ in $L_{0,0}$. Note that we have

$$e^{tB}P = Pe^{tB} = P, \quad t \geq 0, \tag{2.5}$$

and on $L_{k,m}$, $m \geq 1$,

$$BP = PB = 0. \tag{2.6}$$

We assume that

$$\|(e^{tB} - P)f\|_{k,m} \leq C_{k,m} e^{-\gamma_{k,m}t} \|f\|_{k,m}, \quad 0 \leq k, m \leq N. \tag{2.7}$$

Here $C_{k,m}$ and $\gamma_{k,m}$ are positive constants.

In the following theorems it suffices to take $N = 4$.

Theorem 1. Assume that PAP generates a semigroup on $PL_{0,0}$ that maps $PL_{k,m} \rightarrow PL_{k,m}$, $0 \leq k, m \leq N$ boundedly on compact time intervals. Then for $0 < t \leq T$

$$\lim_{\varepsilon \downarrow 0} e^{t(A+B/\varepsilon)} = e^{tPAP} P, \quad \text{in } L_{0,0}. \tag{2.8}$$

Theorem 2. Assume that PAP generates a group of isometries on $PL_{0,0}$ such that e^{tPAP} maps $PL_{k,m} \rightarrow PL_{k,m}$ boundedly for all t .

Let $V \in \mathcal{L}(PL_{k,m}, PL_{k-2,m-2})$ be defined by

$$V = PA \int_0^\infty (e^{tB} - P) dt AP. \tag{2.9}$$

Let $\bar{V} \in \mathcal{L}(PL_{k,m}, PL_{k-2,m-2})$ be defined by

$$\bar{V} = \lim_{T \uparrow \infty} T^{-1} \int_{t_0}^{t_0+T} e^{-sPAP} V e^{sPAP} ds,$$

the limit existing uniformly in $t_0 \geq 0$. Specifically, assume that

$$\|(T^{-1} \int_{t_0}^{t_0+T} e^{-sPAP} V e^{sPAP} ds - \bar{V})Pf\|_{0,0} \leq C \|Pf\|_{2,2}/(1+T). \tag{2.9'}$$

Assume further that \bar{V} generates a contraction semigroup $e^{t\bar{V}}$ on $PL_{0,0}$ and that $e^{t\bar{V}}$ maps $PL_{k,m} \rightarrow PL_{k,m}$, $0 \leq k, m \leq N$, $0 \leq t \leq T$. Then for $0 \leq t \leq T$,

$$\lim_{\varepsilon \downarrow 0} e^{-(t/\varepsilon)PAP} e^{(t/\varepsilon)(A+B/\varepsilon)} P = e^{t\bar{V}} P, \tag{2.10}$$

in $L_{0,0}$ and the error in the approximation is $O(\sqrt{\varepsilon})$.

Remark. If we do not assume (2.9') but define \bar{V} merely as the limit as $T \uparrow \infty$ then (2.10) is valid but we can no longer assert that the error is $O(\sqrt{\varepsilon})$. Note also that \bar{V} commutes with PAP .

Examples. 1. As we mentioned in the introduction, Theorems 1 and 2 are simple abstractions of the situation encountered in the asymptotic analysis of the linearized Boltzmann equation [1]. For this case the spaces $L_{k,m}$ are Hilbert spaces of real valued functions in $R^3 \times R^3$ with norm

$$\|f\|_{k,m}^2 = \sum_{j=0}^k \sum_{i=0}^m \iint (1+|v|^2)^i |\nabla^j f(x, v)|^2 dx dv,$$

where

$$|\nabla^j f|^2 = \sum_{\substack{i_1, \dots, i_n \\ i_1 + \dots + i_n = j}} |\partial^{i_1} f / \partial x_{i_1} \partial x_{i_2} \dots \partial x_{i_n}|^2.$$

The operator B corresponds to the operator L of Grad [1, Section 2] and $A = -v \cdot \nabla$. The properties (2.3) follow from (4.18) and (4.20) of Grad [1] for the class of collision operators that he considers. Grad's a priori estimates [1, Section 6] more than suffice for the analysis given here based on a priori regularity properties. Theorem 2 above includes the averaging due to the pulling-back by the flow e^{-tPAP} in (2.10) as was done by Ellis and Pinsky [2].

2. In order to obtain a better understanding of the asymptotic limit developed in Theorem 2 we consider the following simple finite-dimensional example.

Let A be an $(N - n) \times (N - n)$, $N > n$, dimensional matrix with positive eigenvalues and let

$$B = \begin{matrix} & n & N-n \\ n & \begin{bmatrix} 0 & 0 \\ \vdots & \vdots \end{bmatrix} \\ N-n & \begin{bmatrix} 0 & -A \end{bmatrix} \end{matrix}$$

i.e. B has an n -dimensional nullspace. Let A be an $N \times N$ matrix with block form

$$A = \begin{matrix} & n & N-n \\ n & \begin{bmatrix} A_{11} & A_{12} \\ \vdots & \vdots \end{bmatrix} \\ N-n & \begin{bmatrix} A_{21} & A_{22} \end{bmatrix} \end{matrix},$$

where A_{11} is skew-symmetric. Clearly the projection operator P is

$$P = \begin{matrix} & n & N-n \\ n & \begin{bmatrix} I & 0 \\ \vdots & \vdots \end{bmatrix} \\ N-n & \begin{bmatrix} 0 & 0 \end{bmatrix} \end{matrix}$$

and $PAP = A_{11}$.

Consider now the linear system of equations for the vector functions $x^\varepsilon(t) \in R^n$, $y^\varepsilon(t) \in R^{N-n}$:

$$\begin{aligned} dx^\varepsilon(t)/dt &= \varepsilon^{-1} A_{11}x^\varepsilon(t) + \varepsilon^{-1} A_{12}y^\varepsilon(t), & x^\varepsilon(0) &= x_0 \\ dy^\varepsilon(t)/dt &= -\varepsilon^{-2} A y^\varepsilon(t) + \varepsilon^{-1} A_{21}x^\varepsilon(t) + \varepsilon^{-1} A_{22}y^\varepsilon(t), & y^\varepsilon(0) &= y_0. \end{aligned}$$

According to Theorem 1,

$$x(t) \sim e^{tA_{11}/\varepsilon} x_0, \quad \text{for } 0 \leq t \leq C\varepsilon, \varepsilon \downarrow 0,$$

for any constant C . Theorem 2 gives information in $0 \leq t \leq C$ i.e., in a much larger interval of time. In fact, we have here that

$$\bar{V} = \lim_{T \uparrow \infty} T^{-1} \int_0^T e^{-A_{11}s} (A_{12}A^{-1}A_{21}) e^{A_{11}s} ds$$

and that

$$x^\varepsilon(t) \sim e^{tA_{11}/\varepsilon} e^{t\bar{V}} x_0, \quad 0 \leq t \leq C, \varepsilon \downarrow 0.$$

Thus, Theorem 2 characterizes the modulation $e^{t\bar{V}}$ of the rapidly oscillating part $e^{tA_{11}/\varepsilon}$ of $x^\varepsilon(t)$ (recall that A_{11} is skew-symmetric).

Proof of Theorem 1. From the hypotheses above follow the variation of constants formulas

$$e^{t(A+B/\varepsilon)} = e^{tB/\varepsilon} + \int_0^t e^{(t-s)B/\varepsilon} A e^{s(A+B/\varepsilon)} ds, \tag{2.11}$$

$$e^{t(A+B/\varepsilon)} = e^{tB/\varepsilon} + \int_0^t e^{(t-s)(A+B/\varepsilon)} A e^{sB/\varepsilon} ds. \tag{2.12}$$

These formulas hold in $L_{k,m}$, $1 \leq k, m \leq N$. They will be used frequently in the sequel.

First we show that

$$e^{t(A+B/\varepsilon)}(I-P) \rightarrow 0 \quad \text{in } L_{0,0}, \quad t > 0. \tag{2.13}$$

This follows from (2.12) and our hypotheses

$$e^{t(A+B/\varepsilon)}(I-P) = e^{tB/\varepsilon} - P + \int_0^t e^{(t-s)(A+B/\varepsilon)} A(e^{sB/\varepsilon} - P) ds. \tag{2.14}$$

Here we have used (2.5). For $f \in L_{1,1}$ we have

$$\|e^{t(A+B/\varepsilon)}(I-P)f\|_{0,0} \leq C_{0,0} e^{-\gamma_0 t/\varepsilon} \|f\|_{0,0} + \int_0^t C \cdot C_{1,0} e^{-\gamma_1 s/\varepsilon} ds \|f\|_{1,1}. \tag{2.15}$$

The right hand side of (2.15) tends to zero as $\varepsilon \rightarrow 0, t > 0$. Since $L_{1,1}$ is dense in $L_{0,0}$ the assertion (2.13) follows.

In (2.15) and in the sequel we use C to denote any constant whose particular value is unimportant. This simplifies the notation.

Now we work on the range of P only. Let $\Delta = \Delta(\varepsilon) \downarrow 0$ as $\varepsilon \downarrow 0$, the precise dependence to be chosen later, and decompose $[0, t]$ into m intervals of length Δ so that $m \uparrow \infty$ as $\varepsilon \downarrow 0$. We have

$$\begin{aligned} & \| (e^{t(A+B/\varepsilon)}P - e^{tPAP}P) f \|_{0,0} \\ &= \left\| \sum_{k=0}^{m-1} e^{k\Delta(A+B/\varepsilon)} [e^{\Delta(A+B/\varepsilon)} - e^{\Delta PAP}] e^{(m-k-1)\Delta PAP} P f \right\|_{0,0} \\ &\leq \sum_{k=0}^{m-1} \| (e^{\Delta(A+B/\varepsilon)} - e^{\Delta PAP}) f_k \|_{0,0}. \end{aligned} \tag{2.16}$$

Here we have set $f_k = e^{(m-k-1)\Delta PAP} P f$ which is in the range of P and we have used the hypothesis that $e^{t(A+B/\varepsilon)}$ is a contraction on $L_{0,0}$.

For $f \in L_{2,2}$ using the variation of constants formulas and (2.5) we have:

$$\begin{aligned} e^{\Delta(A+B/\varepsilon)} f_k &= f_k + \Delta(PAP) f_k + \int_0^\Delta (e^{sB/\varepsilon} - P) A f_k ds \\ &\quad + \int_0^\Delta \int_0^s e^{\sigma(A+B/\varepsilon)} A e^{(s-\sigma)B/\varepsilon} A f_k d\sigma ds. \end{aligned} \tag{2.17}$$

Hence,

$$\begin{aligned} & \| e^{\Delta(A+B/\varepsilon)} f_k - f_k - \Delta PAP f_k \|_{0,0} \\ &\leq \int_0^\Delta \| (e^{sB/\varepsilon} - P) A f_k \|_{0,0} ds + C\Delta^2 \| f_k \|_{2,2} \\ &\leq \int_0^\Delta C_{0,0} e^{-\gamma_0 s/\varepsilon} ds \| A f_k \|_{0,0} + C\Delta^2 \| f_k \|_{2,2} \\ &\leq C'\varepsilon \| f_k \|_{1,1} + C\Delta^2 \| f_k \|_{2,2}. \end{aligned} \tag{2.18}$$

We also have that

$$\| e^{\Delta PAP} f_k - f_k - \Delta PAP f_k \|_{0,0} \leq C\Delta^2 \| f_k \|_{2,2}. \tag{2.19}$$

On combining (2.18) and (2.19) it follows that

$$\begin{aligned} \| (e^{t(A+B/\varepsilon)} - e^{tPAP}) P f \|_{0,0} &\leq \sum_{k=0}^{m-1} [C'\varepsilon \| f_k \|_{1,1} + C\Delta^2 \| f_k \|_{2,2}] \\ &\leq \tilde{C} \| f \|_{2,2} [\varepsilon/\Delta + \Delta]. \end{aligned} \tag{2.20}$$

The quantity $(\varepsilon/\Delta + \Delta)$ will go to zero provided Δ is chosen appropriately, for example $\Delta(\varepsilon) = \sqrt{\varepsilon}$. In this case (2.20) shows that we also have an error estimate in the approximation, namely $O(\sqrt{\varepsilon})$.

Proof of Theorem 2. Let

$$U^\varepsilon(t, s) = e^{-PAPt/\varepsilon} e^{(A+B/\varepsilon)(t-s)/\varepsilon} e^{PAPs/\varepsilon}, \quad 0 \leq s \leq t \leq T, \tag{2.21}$$

where we define $e^{t(PAP)}(I - P) = I - P$. We note that $U^\varepsilon(t, s)$ is not a semigroup since it does not depend on $t - s$ only. One can easily verify, however, that with the help of (2.5) and our regularity hypotheses, the following variation of constants formulas hold in $L_{k,m}$, $1 \leq k, m \leq N$.

$$U^\varepsilon(t, s) = e^{(t-s)B/\varepsilon^2} + \varepsilon^{-1} \int_s^t e^{(t-\sigma)B/\varepsilon^2} A^\varepsilon(\sigma) U^\varepsilon(\sigma, s) d\sigma, \tag{2.22}$$

$$U^\varepsilon(t, s) = e^{(t-s)B/\varepsilon^2} + \varepsilon^{-1} \int_s^t U^\varepsilon(t, \sigma) A^\varepsilon(\sigma) e^{(\sigma-s)B/\varepsilon^2} ds, \tag{2.23}$$

$$A^\varepsilon(t) \equiv e^{-PAPt/\varepsilon} (A - PAP) e^{PAPt/\varepsilon}. \tag{2.24}$$

From the regularity hypotheses on e^{PAPt} and $e^{t(A+B/\varepsilon)}$ follows that $U^\varepsilon(t, s)$, $0 \leq s \leq t \leq T$, maps $L_{k,m}$ into itself with bound independent of $\varepsilon > 0$, $0 \leq k, m \leq N$.

We employ again, as in Theorem 1, the decomposition of the interval $[0, t]$ into m intervals of length $\Delta(\varepsilon) \downarrow 0$ as $\varepsilon \downarrow 0$. We have the following

$$\begin{aligned} & \| (U^\varepsilon(t, 0) - e^{t\bar{V}}) P f \|_{0,0} \\ &= \| \sum_{k=0}^{m-1} U^\varepsilon(t, (k+1)\Delta) [U^\varepsilon((k+1)\Delta, k\Delta) - e^{A\bar{V}}] e^{kA\bar{V}} P f \|_{0,0} \\ &\leq \sum_{k=0}^{m-1} \| U^\varepsilon(t, (k+1)\Delta) [U^\varepsilon((k+1)\Delta, k\Delta) - P U^\varepsilon((k+1)\Delta, k\Delta)] e^{kA\bar{V}} P f \|_{0,0} \\ &\quad + \sum_{k=0}^{m-1} \| (P U^\varepsilon((k+1)\Delta, k\Delta) - e^{A\bar{V}}) e^{kA\bar{V}} P f \|_{0,0} \equiv \sum_{k=0}^{m-1} (I_{1,k}^\varepsilon + I_{2,k}^\varepsilon). \end{aligned} \tag{2.25}$$

Here we have also used the hypotheses that e^{tPAP} and $e^{t(A+B/\varepsilon)}$ are contractions in $L_{0,0}$.

We estimate $I_{1,k}^\varepsilon$ first by using the variation of constants formulas (2.22) and (2.23). Set $f_k = e^{kA\bar{V}} P f$. Then

$$\begin{aligned} I_{1,k} &= \| \{ e^{B(t-(k+1)\Delta)/\varepsilon^2} + \varepsilon^{-1} \int_{(k+1)\Delta}^t U^\varepsilon(t, s) A^\varepsilon(s) e^{(s-(k+1)\Delta)B/\varepsilon^2} ds \} \\ &\quad \cdot \{ e^{A\bar{V}} + \varepsilon^{-1} \int_{k\Delta}^{(k+1)\Delta} e^{((k+1)\Delta-\sigma)B/\varepsilon^2} A^\varepsilon(\sigma) U^\varepsilon(\sigma, k\Delta) d\sigma \\ &\quad - P - \varepsilon^{-1} \int_{k\Delta}^{(k+1)\Delta} P A^\varepsilon(\sigma) U^\varepsilon(\sigma, k\Delta) d\sigma \} f_k \|_{0,0} \\ &\leq \| \varepsilon^{-1} \int_{k\Delta}^{(k+1)\Delta} (e^{(t-\sigma)B/\varepsilon^2} - P) A^\varepsilon(\sigma) U^\varepsilon(\sigma, k\Delta) d\sigma f_k \|_{0,0} \\ &\quad + \| \varepsilon^{-2} \int_{(k+1)\Delta}^t \int_{k\Delta}^{(k+1)\Delta} U^\varepsilon(t, s) A^\varepsilon(s) \\ &\quad \cdot (e^{(s-\sigma)B/\varepsilon^2} - P) A^\varepsilon(\sigma) U^\varepsilon(\sigma, k\Delta) d\sigma ds f_k \|_{0,0}. \end{aligned} \tag{2.26}$$

We estimate separately the last two terms on the right side of (2.26).

$$\begin{aligned} & \| \varepsilon^{-1} \int_{k\Delta}^{(k+1)\Delta} (e^{(t-s)B/\varepsilon^2} - P) A^\varepsilon(\sigma) U^\varepsilon(\sigma, k\Delta) d\sigma f_k \|_{0,0} \\ &\leq C\varepsilon^{-1} \int_{k\Delta}^{(k+1)\Delta} C_{0,0} e^{-\gamma_0,0(t-\sigma)/\varepsilon^2} d\sigma \| f_k \|_{1,1} \\ &\leq \begin{cases} C\Delta\varepsilon^{-1} e^{-\gamma_0,0A/\varepsilon^2} \| P f \|_{1,1}, & 0 \leq k \leq m-2 \\ C\varepsilon \| P f \|_{1,1}, & k = m-1. \end{cases} \end{aligned} \tag{2.27}$$

We have used here the hypotheses about the smoothness preserving properties of the semigroups, including $e^{t\bar{V}}$, and (2.3) and (2.7). We recall also the convention about constants that we are employing.

We proceed to the second term on the right side of (2.26) which vanishes when $k = m - 1$. We have

$$\begin{aligned} & \|\varepsilon^{-2} \int_{(k+1)\Delta}^t \int_{k\Delta}^{(k+1)\Delta} U^\varepsilon(t, s) A^\varepsilon(s) (e^{(s-\sigma)B/\varepsilon^2} - P) A^\varepsilon(\sigma) U^\varepsilon(\sigma, k\Delta) f_k d\sigma ds\|_{0,0} \\ & \leq C\varepsilon^{-2} \int_{(k+1)\Delta}^t \int_{k\Delta}^{(k+1)\Delta} e^{-\gamma_{1,1}(s-\sigma)/\varepsilon^2} d\sigma ds \|Pf\|_{2,2} \leq C\varepsilon^2 \|Pf\|_{2,2}. \end{aligned} \tag{2.28}$$

The estimates (2.27) and (2.28) combine to yield

$$\sum_{k=0}^{m-1} I_{1,k} \leq C\Delta^{-1} \varepsilon^2 \|Pf\|_{2,2}. \tag{2.29}$$

By choosing $\Delta(\varepsilon) \rightarrow 0$ so that $\Delta/\varepsilon^2 \rightarrow \infty$ as $\varepsilon \rightarrow 0$ we see from (2.29) that $\sum_{k=0}^{m-1} I_{1,k}$ goes to zero for Pf in $L_{2,2}$.

Next we estimate $I_{2,k}$ where,

$$I_{2,k} = \|(PU^\varepsilon((k+1)\Delta, k\Delta) - e^{A\bar{V}})f_k\|_{0,0}. \tag{2.30}$$

By iterating the variation of constants formulas we obtain the following identity valid in $L_{4,4}$.

$$\begin{aligned} U^\varepsilon((k+1)\Delta, k\Delta) &= e^{A\bar{B}/\varepsilon^2} + \varepsilon^{-1} \int_{k\Delta}^{(k+1)\Delta} e^{((k+1)\Delta-s_1)B/\varepsilon^2} \\ & A^\varepsilon(s_1) e^{(s_1-k\Delta)B/\varepsilon^2} ds_1 + \varepsilon^{-2} \int_{k\Delta}^{(k+1)\Delta} \int_{k\Delta}^{s_1} e^{((k+1)\Delta-s_1)B/\varepsilon^2} \\ & A^\varepsilon(s_1) e^{(s_1-s_2)B/\varepsilon^2} A^\varepsilon(s_2) e^{(s_2-k\Delta)B/\varepsilon^2} ds_2 ds_1 + \varepsilon^{-3} \int_{k\Delta}^{(k+1)\Delta} \int_{k\Delta}^{s_1} \int_{k\Delta}^{s_2} e^{((k+1)\Delta-s_1)B/\varepsilon^2} \\ & A^\varepsilon(s_1) e^{(s_1-s_2)B/\varepsilon^2} A^\varepsilon(s_2) e^{(s_2-s_3)B/\varepsilon^2} A^\varepsilon(s_3) e^{(s_3-k\Delta)B/\varepsilon^2} \\ & ds_3 ds_2 ds_1 + \varepsilon^{-4} \int_{k\Delta}^{(k+1)\Delta} \int_{k\Delta}^{s_1} \int_{k\Delta}^{s_2} \int_{k\Delta}^{s_3} e^{((k+1)\Delta-s_1)B/\varepsilon^2} A^\varepsilon(s_1) e^{(s_1-s_2)B/\varepsilon^2} A^\varepsilon(s_2) U^\varepsilon(s_2, s_4) \\ & A^\varepsilon(s_4) e^{(s_4-s_3)B/\varepsilon^2} A^\varepsilon(s_3) e^{(s_3-k\Delta)B/\varepsilon^2} ds_4 ds_3 ds_2 ds_1. \end{aligned} \tag{2.31}$$

We observe that

$$e^{tPAP}P = Pe^{tPAP}. \tag{2.32}$$

From this observation, (2.5) and the definition (2.24) of A^ε we may rewrite (2.31) as follows.

$$\begin{aligned} PU^\varepsilon((k+1)\Delta, k\Delta) f_k &= f_k + \varepsilon^{-2} \int_{k\Delta}^{(k+1)\Delta} \int_{k\Delta}^{s_1} PA^\varepsilon(s_1) \\ & (e^{(s_1-s_2)B/\varepsilon^2} - P) A^\varepsilon(s_2) f_k ds_2 ds_1 + \varepsilon^{-3} \int_{k\Delta}^{(k+1)\Delta} \int_{k\Delta}^{s_1} \int_{k\Delta}^{s_2} PA^\varepsilon(s_1) \\ & (e^{(s_1-s_2)B/\varepsilon^2} - P) A^\varepsilon(s_2) (e^{(s_2-s_3)B/\varepsilon^2} - P) A^\varepsilon(s_3) f_k ds_3 ds_2 ds_1 \\ & + \varepsilon^{-4} \int_{k\Delta}^{(k+1)\Delta} \int_{k\Delta}^{s_1} \int_{k\Delta}^{s_2} \int_{k\Delta}^{s_3} PA^\varepsilon(s_1) (e^{(s_1-s_2)B/\varepsilon^2} - P) A^\varepsilon(s_2) U^\varepsilon(s_2, s_4) \\ & A^\varepsilon(s_4) (e^{(s_4-s_3)B/\varepsilon^2} - P) A^\varepsilon(s_3) f_k ds_4 ds_3 ds_2 ds_1. \end{aligned} \tag{2.33}$$

From (2.33) we obtain further

$$\begin{aligned} & \|PU^\varepsilon((k+1)\Delta, k\Delta) f_k - f_k - \Delta \bar{V} f_k\|_{0,0} \\ & \leq \Delta \|\varepsilon^{-2} \Delta^{-1} \int_{k\Delta}^{(k+1)\Delta} \int_{k\Delta}^{s_1} PA^\varepsilon(s_1) (e^{(s_1-s_2)B/\varepsilon^2} - P) A^\varepsilon(s_2) f_k ds_2 ds_1 - \bar{V} f_k\|_{0,0} \\ & + \varepsilon^{-3} \int_{k\Delta}^{(k+1)\Delta} \int_{k\Delta}^{s_1} \int_{k\Delta}^{s_2} C \cdot C_{1,1} e^{-\gamma_{1,1}(s_1-s_2)/\varepsilon^2} C_{2,2} e^{-\gamma_{2,2}(s_2-s_3)/\varepsilon^2} \\ & ds_3 ds_2 ds_1 \|f_k\|_{3,3} + \varepsilon^{-4} \int_{k\Delta}^{(k+1)\Delta} \int_{k\Delta}^{s_1} \int_{k\Delta}^{s_2} \int_{k\Delta}^{s_3} C \cdot C_{1,1} e^{-\gamma_{1,1}(s_1-s_2)/\varepsilon^2} \\ & C_{3,3} e^{-\gamma_{3,3}(s_4-s_3)/\varepsilon^2} ds_4 ds_3 ds_2 ds_1 \|f_k\|_{4,4}. \end{aligned} \tag{2.34}$$

In the last two terms on the right side of (2.34) we have used (2.7) and the regularity

hypotheses along with (2.3). Elementary calculations yield

$$\begin{aligned} & \|PU^\varepsilon((k+1)\Delta, k\Delta)f_k - f_k - \Delta\bar{V}f_k\|_{0,0} \\ & \leq \Delta\|\varepsilon^{-2}\Delta^{-1}\int_{k\Delta/\varepsilon^2}^{(k+1)\Delta/\varepsilon^2}\int_{k\Delta/\varepsilon^2}^{s_1}PA^\varepsilon(s_1)(e^{(s_1-s_2)B/\varepsilon^2} - P)A^\varepsilon(s_2)f_k ds_2 ds_1 - \bar{V}f_k\|_{0,0} \\ & \quad + C_1\Delta\varepsilon\|f_k\|_{3,3} + C_2\Delta^2\|f_k\|_{4,4}. \end{aligned} \tag{2.35}$$

We shall now analyze the first term on the right side of (2.35) which we denote, without the factor Δ , by J_k^ε . First we observe the following.

$$\begin{aligned} & PA^\varepsilon(s_1)(e^{(s_1-s_2)B/\varepsilon^2} - P)A^\varepsilon(s_2)P = Pe^{-s_1PAP/\varepsilon}(A - PAP)e^{s_1(PAP)/\varepsilon} \\ & (e^{(s_1-s_2)B/\varepsilon^2} - P)e^{-s_2PAP/\varepsilon}(A - PAP)e^{s_2PAP/\varepsilon}P = Pe^{-s_1PAP/\varepsilon}PA \\ & (e^{(s_1-s_2)B/\varepsilon^2} - P)APe^{s_2PAP/\varepsilon}P. \end{aligned} \tag{2.36}$$

Here we have used the definition $e^{tPAP}(I - P) = I - P$ and (2.5). In view of (2.36) we have

$$\begin{aligned} J_k^\varepsilon &= \|\varepsilon^2\Delta^{-1}\int_{k\Delta/\varepsilon^2}^{(k+1)\Delta/\varepsilon^2}\int_{k\Delta/\varepsilon^2}^{s_1}Pe^{-s_1\varepsilon PAP}PA(e^{(s_1-s_2)B} - P)AP \\ & e^{s_2\varepsilon PAP}f_k ds_2 ds_1 - \bar{V}f_k\|_{0,0} \leq \|\varepsilon^2\Delta^{-1}\int_{k\Delta/\varepsilon^2}^{(k+1)\Delta/\varepsilon^2}e^{-s_1\varepsilon PAP}\int_{k\Delta/\varepsilon^2}^{s_1} \\ & PA(e^{(s_1-s_2)B} - P)AP(e^{-s_1\varepsilon PAP} - I)ds_2 e^{s_1\varepsilon PAP}f_k ds_1\|_{0,0} \\ & + \|\varepsilon^2\Delta^{-1}\int_{k\Delta/\varepsilon^2}^{(k+1)\Delta/\varepsilon^2}e^{-s_1\varepsilon PAP}(PA\int_0^\infty(e^{\sigma B} - P)d\sigma AP)e^{s_1\varepsilon PAP}f_k ds_1 - \bar{V}f_k\|_{0,0} \\ & + \|\varepsilon^2\Delta^{-1}\int_{k\Delta/\varepsilon^2}^{(k+1)\Delta/\varepsilon^2}e^{-s_1\varepsilon PAP}(PA\int_{s_1-k\Delta/\varepsilon^2}^\infty(e^{\sigma B} - P)d\sigma AP)e^{s_1\varepsilon PAP}f_k ds_1\|_{0,0} \\ & = J_{k,1}^\varepsilon + J_{k,2}^\varepsilon + J_{k,3}^\varepsilon. \end{aligned} \tag{2.37}$$

For $J_{k,1}^\varepsilon$ we have the following estimate

$$\begin{aligned} J_{k,1}^\varepsilon &\leq \varepsilon^2\Delta^{-1}\int_{k\Delta/\varepsilon^2}^{(k+1)\Delta/\varepsilon^2}\int_{k\Delta/\varepsilon^2}^{s_1}C \cdot C_{1,1}e^{-\gamma_{1,1}(s_1-s_2)}\varepsilon(s_1-s_2) \\ & ds_2 ds_1\|f_k\|_{3,3} \leq C \cdot \varepsilon\|f_k\|_{3,3}. \end{aligned} \tag{2.38}$$

Here as usual, we have employed the regularity hypotheses and (2.7). For $J_{k,2}^\varepsilon$ we have, using (2.9),

$$\begin{aligned} J_{k,2}^\varepsilon &= \|\varepsilon\Delta^{-1}\int_{k\Delta/\varepsilon^2}^{(k+1)\Delta/\varepsilon^2}e^{-s_1PAP}V e^{s_1PAP}f_k ds_1 - \bar{V}f_k\|_{0,0} \\ & \leq C(1 + \Delta/\varepsilon)^{-1}\|f_k\|_{2,2}. \end{aligned} \tag{2.39}$$

The last inequality is hypothesis (2.9'). For $J_{k,3}^\varepsilon$ we have

$$J_{k,3}^\varepsilon \leq \varepsilon^2\Delta^{-1}\int_{k\Delta/\varepsilon^2}^{(k+1)\Delta/\varepsilon^2}C \cdot \int_{s_1-k\Delta/\varepsilon^2}^\infty C_{1,1}e^{-\gamma_{1,1}\sigma}d\sigma ds_1\|f_k\|_{2,2} \leq C\varepsilon^2\Delta^{-1}\|f_k\|_{2,2}. \tag{2.40}$$

We return now to (2.35). Using (2.38), (2.39), and (2.40) in (2.37) and then in (2.35) we find that

$$\begin{aligned} & \|PU^\varepsilon((k+1)\Delta, k\Delta)f_k - f_k - \Delta\bar{V}f_k\|_{0,0} \leq \Delta[C_1\varepsilon\|f_k\|_{3,3} + C_2(1 + \Delta/\varepsilon)^{-1}\|f_k\|_{2,2} \\ & \quad + C_3\varepsilon^2\Delta^{-1}\|f_k\|_{2,2} + C_4\varepsilon\|f_k\|_{3,3} + C_5\Delta\|f_k\|_{4,4}]. \end{aligned} \tag{2.41}$$

Our hypotheses about \bar{V} yield readily the estimate

$$\|e^{\Delta\bar{V}}f_k - f_k - \Delta\bar{V}f_k\|_{0,0} \leq C_6\Delta^2\|f_k\|_{4,4}. \tag{2.42}$$

Thus, from (2.41) and (2.42) it follows that

$$\begin{aligned} \sum_{k=0}^{m-1} I_{2,k} &= \sum_{k=0}^{m-1} \|(PU^\varepsilon((k+1)\Delta, k\Delta) - e^{A\bar{V}})f_k\|_{0,0} \leq C_1\varepsilon\|Pf\|_{3,3} \\ &\quad + C_2(1+\Delta/\varepsilon)^{-1}\|Pf\|_{2,2} + C_3\varepsilon^2\Delta^{-1}\|Pf\|_{2,2} + C_4\varepsilon\|Pf\|_{3,3} \\ &\quad + (C_5 + C_6)\Delta\|Pf\|_{4,4}. \end{aligned} \tag{2.43}$$

Now if we choose $\Delta(\varepsilon) \rightarrow 0$ as $\varepsilon \rightarrow 0$ so that $\Delta(\varepsilon)/\varepsilon \rightarrow \infty$ then, the right side of (2.43) goes to zero and so does the right side of (2.29). This proves the theorem.

For $\Delta(\varepsilon) = \sqrt{\varepsilon}$ we find that the overall error in the approximation is $O(\sqrt{\varepsilon})$ as asserted in the statement of the theorem.

3. Stochastic Problems with Rapidly Varying Deterministic Part

We adopt the same framework as the one described at the beginning of Section 2. The operator B is defined as before including properties (2.4), (2.5), (2.6), and (2.7). The operator A is now different and it is defined as follows.

Let (Ω, \mathcal{F}, P) be a probability space. For each $t \geq 0$ and $\omega \in \Omega$ let $A(t, \omega)$ be a linear operator in $\mathcal{L}(L_{k,m}, L_{k-1,m-1})$ which is strongly measurable jointly in t and ω . To simplify the discussion of questions of measurability we assume that the spaces $L_{k,m}$, $k, m \geq 0$ are separable. Let $\mathcal{F}_s^t \subset \mathcal{F}$, $0 \leq s \leq t \leq \infty$ be an increasing family of σ -algebras, that is

$$\mathcal{F}_{s_1}^{t_1} \subset \mathcal{F}_{s_2}^{t_2}, \quad 0 \leq s_2 \leq s_1 \leq t_1 \leq t_2 \leq \infty.$$

We assume that the conditional probabilities $P(F|\mathcal{F}_0^s)$ have a regular version $P_s(F|\omega)$ as in [3] and that P is mixing relative to \mathcal{F}_s^t as follows.

$$\sup_{s \geq 0} \sup_{\substack{U \in \mathcal{F}_{s+t}^{\infty} \\ W \in \mathcal{F}_s^0}} |P(U|W) - P(U)| = \varrho(t) \downarrow 0, \quad t \uparrow \infty, \quad \int_0^\infty \varrho^{\frac{1}{2}}(s) ds < \infty. \tag{3.1}$$

We assume that $A(t) = A(t, \omega)$ is \mathcal{F}_t^t measurable for each $t \geq 0$.

We are interested in the asymptotic analysis of the stochastic equation

$$dy^\varepsilon(t)/dt = \varepsilon^{-1}By^\varepsilon(t) + A(t/\varepsilon)y^\varepsilon(t), \quad t > 0, \quad y^\varepsilon(0) = y_0, \quad \varepsilon \downarrow 0. \tag{3.2}$$

Naturally, this equation will be analyzed under different conditions corresponding to Theorems 1 and 2 of Section 2. First we make precise the sense in which (3.2) is to be taken.

It is convenient to introduce the evolution operator associated with (3.2)

$$y^\varepsilon(t) = U^\varepsilon(t, s)y^\varepsilon(s), \quad 0 \leq s \leq t. \tag{3.3}$$

We state the relevant hypotheses in terms of $U^\varepsilon(t, s)$. We assume the following:

- (i) $U^\varepsilon(t, s)$, $t \geq s$ are contraction operators on $L_{0,0} \rightarrow L_{0,0}$ and strongly $\mathcal{F}_{s/\varepsilon}^{t/\varepsilon}$ measurable. They are also bounded on $L_{k,m} \rightarrow L_{k,m}$ for $1 \leq k, m \leq 4$.
- (ii) The finite propagator property holds.

$$U^\varepsilon(t, \sigma)U^\varepsilon(\sigma, s) = U^\varepsilon(t, s), \quad 0 \leq s \leq \sigma \leq t, \quad U^\varepsilon(t, t) = I. \tag{3.4}$$

(iii) The variation of constants formulas hold in $L_{k,m}$, $k, m \geq 1$.

$$U^\varepsilon(t, s) = e^{(t-s)B/\varepsilon} + \int_s^t U^\varepsilon(t, \sigma)A(\sigma/\varepsilon)e^{(\sigma-s)B/\varepsilon}d\sigma, \tag{3.5}$$

$$U^\varepsilon(t, s) = e^{(t-s)B/\varepsilon} + \int_s^t e^{(t-\sigma)B/\varepsilon}A(\sigma/\varepsilon)U^\varepsilon(\sigma, s)d\sigma. \tag{3.6}$$

With these hypotheses we have the following theorems.

Theorem 1. Assume that the limit

$$\bar{A} = \lim_{T \uparrow \infty} T^{-1} \int_{t_0}^{t_0+T} E\{A(s)\}ds \tag{3.7}$$

exists uniformly in $t_0 \geq 0$ and $P\bar{A}P$ defines an operator in $\mathcal{L}(PL_{k,m}, PL_{k-1,m-1})$. Assume that $P\bar{A}P$ generates a semigroup of contractions $e^{tP\bar{A}P}$ in $PL_{0,0}$ such that $e^{tP\bar{A}P}$ maps $PL_{k,m}$ into itself for $0 \leq t \leq T$. Then for $0 \leq t \leq T$

$$\lim_{\varepsilon \downarrow 0} E\{U^\varepsilon(t, 0)Pf\} = e^{tP\bar{A}P}Pf, \text{ in } L_{0,0}. \quad \square \tag{3.8}$$

Theorem 2. Assume that the operator $P\bar{A}P$ in Theorem 1 is identically zero.

$$P\bar{A}P \equiv 0. \tag{3.9}$$

Let $\bar{V} \in \mathcal{L}(PL_{k,m}, PL_{k-2,m-2})$ be defined by

$$\bar{V} = \lim_{T \uparrow \infty} T^{-1} \int_{t_0}^{t_0+T} \int_{t_0}^s E\{PA(s)(e^{B(s-\sigma)} - P)A(\sigma)P\}d\sigma ds \tag{3.10}$$

which we assume exists uniformly in $t_0 \geq 0$ and is independent of t_0 . We assume that \bar{V} generates a semigroup $e^{t\bar{V}}$ on $PL_{0,0}$ such that $e^{t\bar{V}}$ maps $PL_{k,m} \rightarrow PL_{k,m}$ for $0 \leq t \leq T$, $k, m \geq 1$. Suppose $U^\varepsilon(t, s)$ satisfies, for all $\omega \in \Omega$,

$$\|U^\varepsilon(t/\varepsilon, s/\varepsilon)f\|_{1,1} \leq C_{1,1}\|f\|_{1,1}, \quad 0 \leq t-s \leq \varepsilon, \tag{3.11}$$

$$\|U^\varepsilon(t/\varepsilon, s/\varepsilon)f\|_{2,2} \leq C_{2,2}\|f\|_{2,2}, \quad 0 \leq t-s \leq \varepsilon. \tag{3.12}$$

Under these hypotheses

$$\lim_{\varepsilon \downarrow 0} E\{U^\varepsilon(t/\varepsilon, 0)Pf\} = e^{t\bar{V}}Pf, f \in L_{0,0}, \quad 0 \leq t \leq T. \quad \square \tag{3.13}$$

Theorem 3. Assume that $A(t)$, $t \geq 0$, is a stationary process, that $P\bar{A}P \neq 0$ and that $P\bar{A}P$ generates a group of contractions $e^{tP\bar{A}P}$ in $PL_{0,0}$ which maps $PL_{k,m} \rightarrow PL_{k,m}$, $k, m \geq 1$, $-\infty < t < \infty$. Assume that (3.11) and (3.12) hold for $0 \leq t-s \leq T$. Assume further that

$$V = \int_0^\infty PE\{A(t)(e^{Bt} - P)A(0)\}Pdt \tag{3.14}$$

defines an operator in $\mathcal{L}(PL_{k,m}, PL_{k-2,m-2})$ and that the limit

$$\bar{V} = \lim_{T \uparrow \infty} T^{-1} \int_{t_0}^{t_0+T} e^{-sP\bar{A}P}V e^{sP\bar{A}P}ds \tag{3.15}$$

exists uniformly in $t_0 \geq 0$, is independent of it and defines an operator in $\mathcal{L}(PL_{k,m}, PL_{k-2,m-2})$. Let \bar{V} satisfy the same hypotheses as in Theorem 2.

Then for $0 \leq t \leq T$,

$$\lim_{\varepsilon \downarrow 0} e^{-(t/\varepsilon)P\bar{A}P}E\{U(t/\varepsilon, 0)Pf\} = e^{t\bar{V}}Pf, f \in L_{0,0}. \quad \square \tag{3.16}$$

Remarks. 1. Note that Theorems 2 and 3 complement each other. The one holds when $P\dot{A}P \equiv 0$ but $A(t)$ need not be stationary although (3.10) holds uniformly in t_0 . The other holds when $P\dot{A}P \neq 0$, generates a group and $A(t)$ is stationary. It does not seem possible to have $P\dot{A}P \neq 0$ and $A(t)$ non-stationary at the same time and expect a limit like (3.13) or (3.16) to exist.

2. As we mentioned in the introduction, one application of these theorems is in transport theoretic problems in random media [9]. An example is provided by the Boltzmann equation of gas dynamics [1] which was discussed in Example 1 of Section 2 when A is not $-v \cdot \nabla$ but, say, $-v(t, \omega) \cdot \nabla$ with $v(t, \omega)$ a random function of time with values in R^3 . More complicated, space dependent, examples require a considerable amount of effort to verify that all hypotheses of the theorems hold. In particular, (3.11) and (3.12) for Theorem 2 are not obvious and their strengthened version in Theorem 3 is probably not true unless very special assumptions are made.

Proof of Theorems 1, 2, and 3. The proof of all three theorems is very similar to the proof of Theorems 1 and 2 of Section 2. Therefore, we shall not repeat the details here but we shall present those portions of the argument which are different. In particular we shall discuss Theorem 2, which is typical.

We begin with (2.25) including expectation values. We also change notation so that $U^\varepsilon(t, s)$ stands for $U^\varepsilon(t/\varepsilon, s/\varepsilon)$. We have

$$\begin{aligned} & \| (E\{U^\varepsilon(t, 0)\} - e^{t\bar{V}})Pf \|_{0,0} = \left\| \sum_{k=0}^{m-1} E\{U^\varepsilon(t, (k+1)\Delta)\} \right. \\ & [U^\varepsilon((k+1)\Delta, k\Delta) - e^{k\Delta\bar{V}}]e^{k\Delta\bar{V}}Pf \left. \right\|_{0,0} \leq \sum_{k=0}^{m-1} \| (E\{U^\varepsilon(t, (k+1)\Delta)\} \\ & U^\varepsilon((k+1)\Delta, k\Delta)\} - E\{U^\varepsilon(t, (k+1)\Delta)\}PE\{U^\varepsilon((k+1)\Delta, k\Delta)\})e^{k\Delta\bar{V}}Pf \|_{0,0} \\ & + \sum_{k=0}^{m-1} \| (PE\{U^\varepsilon((k+1)\Delta, k\Delta)\} - e^{k\Delta\bar{V}})e^{k\Delta\bar{V}}Pf \|_{0,0} = \sum_{k=0}^{m-1} (I_{1,k}^\varepsilon + I_{2,k}^\varepsilon). \end{aligned} \tag{3.17}$$

The treatment of $I_{2,k}^\varepsilon$ for both Theorems 2 and 3, is similar to the one of Section 2. The treatment of $I_{1,k}^\varepsilon$ is where differences arise and the mixing hypothesis (3.1) enters. We shall examine this in more detail.

We rewrite $I_{1,k}^\varepsilon$ using (3.5) and (3.6) and recalling that $U^\varepsilon(t, s)$ stands for $U^\varepsilon(t/\varepsilon, s/\varepsilon)$ here.

$$\begin{aligned} I_{1,k}^\varepsilon &= \| E\{ (e^{B(t-(k+1)\Delta)/\varepsilon^2} + \varepsilon^{-1} \int_{(k+1)\Delta}^t U^\varepsilon(t, s)A(s/\varepsilon^2) \\ & e^{(s-(k+1)\Delta)B/\varepsilon^2} ds) \cdot (e^{AB/\varepsilon^2} + \varepsilon^{-1} \int_{k\Delta}^{(k+1)\Delta} e^{((k+1)\Delta-\sigma)B/\varepsilon^2} A(\sigma/\varepsilon^2) \\ & U^\varepsilon(\sigma, k\Delta)d\sigma) f_k \} - E\{ (e^{B(t-(k+1)\Delta)/\varepsilon^2} + \varepsilon^{-1} \int_{(k+1)\Delta}^t U(t, s)A(s/\varepsilon^2) \\ & e^{(s-(k+1)\Delta)B/\varepsilon^2} ds) \} \cdot PE\{ (e^{AB/\varepsilon^2} + \varepsilon^{-1} \int_{k\Delta}^{(k+1)\Delta} e^{((k+1)\Delta-\sigma)B/\varepsilon^2} \\ & A(\sigma/\varepsilon^2)U^\varepsilon(\sigma, k\Delta)d\sigma) f_k \} \|_{0,0} \leq \| \varepsilon^{-1} \int_{k\Delta}^{(k+1)\Delta} (e^{(t-\sigma)B/\varepsilon^2} - P)E\{A(\sigma/\varepsilon^2) \\ & U^\varepsilon(\sigma, k\Delta) f_k\} d\sigma \|_{0,0} + \| \varepsilon^{-2} \int_{(k+1)\Delta}^t \int_{k\Delta}^{(k+1)\Delta} E\{U^\varepsilon(t, s)A(s/\varepsilon^2) \\ & (e^{(s-\sigma)B/\varepsilon^2} - P)A(\sigma/\varepsilon^2)U^\varepsilon(\sigma, k\Delta) f_k\} d\sigma ds \|_{0,0} \\ & + \| \varepsilon^{-2} \int_{(k+1)\Delta}^t \int_{k\Delta}^{(k+1)\Delta} [E\{U^\varepsilon(t, s)A(s/\varepsilon^2)PA(\sigma/\varepsilon^2)U^\varepsilon(\sigma, k\Delta)\} \\ & - E\{U^\varepsilon(t, s)A(s/\varepsilon^2)\}PE\{A(\sigma/\varepsilon^2)U^\varepsilon(\sigma, k\Delta)\}] f_k d\sigma ds \|_{0,0}. \end{aligned} \tag{3.18}$$

The first term on the right side of (3.18) is estimated in exactly the same way as (2.27). The second term, similarly, exactly as (2.28). The third term in turn is estimated in exactly the way $I_2^{(k)}$ in (2.24) of [3] was estimated. In particular

Lemma 1 of [3] is used here again and, as for other calculations, hypotheses (3.11) and (3.12). The remainder of the proof proceeds without any substantial differences from the analysis of Section 2 and [3].

4. Application to Stochastic Differential Equations

Let (Ω, F, P) be a probability space and consider for $(\tilde{X}^\varepsilon(t), Y^\varepsilon(t)) = (\tilde{X}^\varepsilon(t, \omega), Y^\varepsilon(t, \omega))$, $\omega \in \Omega$, the following system of stochastic differential equations.

$$\begin{aligned} d\tilde{X}^\varepsilon(t)/dt &= \varepsilon^{-2} A^{(1)} \tilde{X}^\varepsilon(t) + \varepsilon^{-1} \tilde{F}^{(1)}(\tilde{X}^\varepsilon(t), Y^\varepsilon(t), t/\varepsilon^2) \\ &\quad + \tilde{G}^{(1)}(\tilde{X}^\varepsilon(t), Y^\varepsilon(t), t/\varepsilon^2), \\ dY^\varepsilon(t)/dt &= -\varepsilon^{-2} A^{(2)} Y^\varepsilon(t) + \varepsilon^{-1} \tilde{F}^{(2)}(\tilde{X}^\varepsilon(t), Y^\varepsilon(t), t/\varepsilon^2) \\ &\quad + \tilde{G}^{(2)}(\tilde{X}^\varepsilon(t), Y^\varepsilon(t), t/\varepsilon^2), \quad t > 0. \\ \tilde{X}^\varepsilon(0) &= x, \quad Y^\varepsilon(0) = y. \end{aligned} \tag{4.1}$$

Here $\tilde{X}^\varepsilon(t)$ and $Y^\varepsilon(t)$ take values in R^n and R^m respectively, $A^{(1)}$ is an $n \times n$ skew-symmetric matrix, $A^{(2)}$ is an $m \times m$ positive definite matrix, $\tilde{F}^{(1)}(x, y, t, \omega)$, $\tilde{G}^{(1)}(x, y, t, \omega)$ and $\tilde{F}^{(2)}(x, y, t, \omega)$, $\tilde{G}^{(2)}(x, y, t, \omega)$ are random vector functions in R^n and R^m respectively.

To analyze the asymptotic behavior of $(\tilde{X}^\varepsilon, Y^\varepsilon)$ as $\varepsilon \downarrow 0, 0 \leq t \leq T$, it is necessary first to remove the rapidly oscillating term $\varepsilon^{-2} A^{(1)} \tilde{X}^\varepsilon$ by defining new dependent variables as follows

$$\tilde{X}^\varepsilon(t) = e^{tA^{(1)}/\varepsilon^2} X^\varepsilon(t). \tag{4.2}$$

The equations for $(X^\varepsilon(t), Y^\varepsilon(t))$ are

$$\begin{aligned} dX^\varepsilon/dt &= \varepsilon^{-1} F^{(1)}(X^\varepsilon, Y^\varepsilon, t/\varepsilon^2) + G^{(1)}(X^\varepsilon, Y^\varepsilon, t/\varepsilon^2) \\ dY^\varepsilon/dt &= -\varepsilon^{-2} A^{(2)} Y^\varepsilon + \varepsilon^{-1} F^{(2)}(X^\varepsilon, Y^\varepsilon, t/\varepsilon^2) + G^{(2)}(X^\varepsilon, Y^\varepsilon, t/\varepsilon^2), \quad t > 0. \\ X^\varepsilon(0) &= x, \quad Y^\varepsilon(0) = y. \end{aligned} \tag{4.3}$$

The vector functions $F^{(1)}, G^{(1)}$, and $F^{(2)}, G^{(2)}$ are given by

$$\begin{aligned} F^{(1)}(x, y, t, \omega) &= e^{-A^{(1)t}} \tilde{F}^{(1)}(e^{A^{(1)t}} x, y, t, \omega) \\ G^{(1)}(x, y, t, \omega) &= e^{-A^{(1)t}} \tilde{G}^{(1)}(e^{A^{(1)t}} x, y, t, \omega), \\ F^{(2)}(x, y, t, \omega) &= \tilde{F}^{(2)}(e^{A^{(1)t}} x, y, t, \omega), \\ G^{(2)}(x, y, t, \omega) &= \tilde{G}^{(2)}(e^{A^{(1)t}} x, y, t, \omega). \end{aligned}$$

Without loss of generality we take $A^{(2)} = \text{diagonal } (a_1^{(2)}, a_2^{(2)}, \dots, a_m^{(2)})$, $a_k^{(2)} > 0$, $k = 1, \dots, m$.

Let $\mathcal{F}_s^t \subset \mathcal{F}$, $0 \leq s \leq t \leq \infty$, be the family of σ -algebras introduced in Section 3 satisfying condition (3.1) with the strengthened assumption $\int_0^\infty \varrho^{1/5}(s) ds < \infty$. Denote by $F(x, y, t, \omega)$ and $G(x, y, t, \omega)$ the R^{n+m} vector functions $(F^{(1)}, F^{(2)})$ and $(G^{(1)}, G^{(2)})$. We assume the following

(i) F and G are mappings from $R^{n+m} \times [0, \infty] \times \Omega$ into R^{n+m} which are jointly measurable with respect to their arguments and which for fixed x, y, t are \mathcal{F}_t^1 measurable functions of ω .

$$(ii) E\{F(x, y, t, \cdot)\} = 0. \tag{4.4}$$

$$(iii) |F(x, y, t, \omega)| + |G(x, y, t, \omega)| \leq C(1 + |x| + |y|)$$

$$|D^1 F(x, y, t, \omega)| + |D^1 G(x, y, t, \omega)| \leq C$$

$$|D^i F(x, y, t, \omega)| + |D^i G(x, y, t, \omega)| \leq C(1 + |x|^p + |y|^p), \quad i = 2, 3.$$

Here D^i stands for any partial derivative of order $i = 1, 2, 3$ with respect to x or y , $|\cdot|$ stands for norm of vectors or absolute value of scalars, C is a constant and p is a nonnegative integer.

We state shortly a theorem characterizing the asymptotic behavior of $X^\varepsilon(t)$ as $\varepsilon \rightarrow 0$. The limit process is a diffusion Markov process and we proceed now to introduce its infinitesimal generator.

We assume that the following limits exist independently of $t_0 \geq 0$.

$$a^{ij}(x) = \lim_{T \uparrow \infty} T^{-1} \int_{t_0}^{t_0+T} \int_{t_0}^s E\{F_i^{(1)}(x, 0, \sigma) F_j^{(1)}(x, 0, s)\} d\sigma ds \quad i, j = 1, \dots, n, \tag{4.5}$$

$$\begin{aligned} b^i(x) = & \lim_{T \uparrow \infty} T^{-1} \int_{t_0}^{t_0+T} \int_{t_0}^s [\sum_{k=1}^n E\{F_k^{(1)}(x, 0, \sigma) \partial F_i^{(1)}(x, 0, s) / \partial x_k\} \\ & + \sum_{k=1}^n E\{F_k^{(2)}(x, 0, \sigma) e^{-a_k^{(2)}(\sigma-s)} \partial F_i^{(1)}(x, 0, s) / \partial y_k\}] d\sigma ds \\ & + \lim_{T \uparrow \infty} T^{-1} \int_{t_0}^{t_0+T} E\{G_i^{(1)}(x, 0, s)\} ds, \quad i = 1, 2, \dots, n. \end{aligned} \tag{4.6}$$

We assume moreover that the magnitude of the difference between $a^{ij}(x)$ and the integral in (4.5) is less than $C(1 + |x|^2)/(1 + T)$, $T \uparrow \infty$, while the magnitude of the difference between $b^i(x)$ and the sum of the integrals in (4.6) is less than $C(1 + |x|)/(1 + T)$. In the space of twice continuously differentiable functions on R^n we define the differential operator \mathcal{L} as follows.

$$\mathcal{L} f(x) = \sum_{i,j=1}^n a^{ij}(x) \partial^2 f(x) / \partial x_i \partial x_j + \sum_{i=1}^n b^i(x) \partial f(x) / \partial x_i. \tag{4.7}$$

Let $C^{k,p}(R^n)$ denote the class of functions on R^n that are k -times continuously differentiable and for which all partial derivatives up to order k inclusive are bounded by a constant times $(1 + |x|^p)$, p a nonnegative integer. Consider the parabolic initial value problem

$$\partial u(\tau, x) / \partial \tau = \mathcal{L} u(\tau, x), \quad \tau > 0, x \in R^n, \quad u(0, x) = f(x) \in C^{4,p}(R^n). \tag{4.8}$$

To insure that, for some $\bar{p} \geq 0$, (4.8) has a unique solution in $C^{4,\bar{p}}(R^n)$ for $0 \leq \tau \leq t_0$ it is sufficient to assume that the matrix $a^{ij}(x)$ has a square root $c^{ij}(x)$ which is sufficiently regular:

$$\begin{aligned} |b^i(x)| + |c^{ij}(x)| & \leq C(1 + |x|), \quad |D^1 b^i(x)| + |D^1 c^{ij}(x)| \leq C \\ b^i(x), c^{ij}(x) & \in C^{4,p}(R^n), \quad i, j = 1, \dots, n. \end{aligned} \tag{4.9}$$

Ito's calculus can be used to insure existence and uniqueness of a solution in $C^{4,\bar{p}}(R^n)$ [13]. The hypotheses (4.9) can be frequently verified from those on F and G and (4.5), (4.6). When this is not possible, one may have to change (4.9) and apply the theory of Oleinik [14].

Theorem. *Let the above hypotheses hold and let $f(x, y) \in C^{4,p}(R^{n+m})$. Let $u(\tau, x)$ denote the solution of (4.8) with data $f(x, 0)$. Then there exists an integer \bar{p} such that for all ε sufficiently small.*

$$|E\{f(X^\varepsilon(\tau, \sigma, x, y), Y^\varepsilon(\tau, \sigma, x, y)) | \mathcal{F}_0^{\sigma/\varepsilon^2}\} - [f(x, e^{-A^{(2)}(\tau-\sigma)/\varepsilon^2}y) - f(x, 0)] - u(\tau - \sigma, x)| \leq \varepsilon^{\frac{1}{2}}C(1 + |x|^{\bar{p}} + |y|^{\bar{p}}), \quad 0 \leq \sigma \leq \tau \leq \tau_0 < \infty. \tag{4.10}$$

Let $X^0(\tau, x)$ denote the n -dimensional diffusion Markov process with infinitesimal generator \mathcal{L} and with $X^0(0, x) = x$. Then as $\varepsilon \rightarrow 0$, $X^\varepsilon(\tau, 0, x, y)$ converges weakly to $X^0(\tau, x)$ as a process in $C([0, \tau_0], R^n)$, with x, y in a compact subset of R^{n+m} .

Remarks. 1. Limit theorems for stochastic differential equations are considered in [12] and several references to the work of Stratonovich and Khasminskii are given in [4, 10]. The theorem as stated here is an extension of the result in [10] and its proof requires no essential changes at all.

2. As we mentioned in the introduction the theorem of this section can be viewed as a consequence of Theorem 2 of Section 3 and this is the reason for presenting it here. This is similar to the approach followed in Sections 3 and 4 in [3]. More specifically, the operator $A(t)$ of Section 2 is given on smooth functions by

$$A(t)f(x, y) = \sum_{i=1}^n (F_i^{(1)}(x, y, t) \partial f(x, y) / \partial x_i + \varepsilon G_i^{(1)}(x, y, t) \partial f(x, y) / \partial x_i) + \sum_{i=1}^m (F_i^{(2)}(x, y, t) \partial f(x, y) / \partial y_i + \varepsilon G_i^{(2)}(x, y, t) \partial f(x, y) / \partial y_i). \tag{4.11}$$

The fact that $A(t)$ has also terms of order ε is, of course, easily accommodated. The operator B is given by

$$Bf(x, y) = - \sum_{i,j=1}^m A_{ij}^{(2)} y_j \partial f(x, y) / \partial y_i, \tag{4.12}$$

so that

$$e^{Bt} f(x, y) = f(x, e^{-A^{(2)}t}y), \quad Pf(x, y) = f(x, 0). \tag{4.13}$$

With this formal identification of objects and the type of analysis of Sections 3 and 4 in [3] one can obtain a result similar to the theorem of this section. However, the result as stated here is considerably more general as far as the various hypotheses are concerned and also regarding the conclusion. Therefore, for the special problem (4.3) the direct approach of [10] is preferable. We note that, with $A(t)$ and B given by (4.11)–(4.13), the infinitesimal generator \mathcal{L} of (4.7) coincides with the operator \bar{V} of (3.10) as, of course, it should be.

We shall conclude this section with an example illustrating the content of the theorem.

Consider the following system of coupled oscillators

$$d^2 \xi_i^\varepsilon(t) / dt^2 + 2\alpha_i d \xi_i^\varepsilon(t) / dt + \tilde{\omega}_i^2 \xi_i^\varepsilon(t) = \varepsilon \sum_{j=1}^M \mu_{ij}(t) \xi_j^\varepsilon(t) + \varepsilon^2 \tilde{G}_i(\xi_1^\varepsilon, d \xi_1^\varepsilon / dt, \dots, \xi_M^\varepsilon, d \xi_M^\varepsilon / dt, t), \quad t > 0. \tag{4.14}$$

$\xi_i^\varepsilon(0)$ and $d \xi_i^\varepsilon(0) / dt$ given.

Here $\xi_i^\varepsilon(t)$ represents the amplitude of the i^{th} oscillator. We assume that the following hypotheses hold.

$$(i) \alpha_1 = 0, \quad 0 < \alpha_2 \leq \alpha_3 \leq \dots \leq \alpha_M, \tag{4.15}$$

$$\tilde{\omega}_i > \alpha_i, \quad i = 1, \dots, M.$$

(ii) $\mu_{ij}(t) = \mu_{ij}(t, \omega)$, $i, j = 1, \dots, M$, are real, symmetric ($\mu_{ij} = \mu_{ji}$), \mathcal{F}_t^t measurable, stationary, zero mean processes a.s. bounded for all $t \geq 0$. The $\omega \in \Omega$, (Ω, F, P) a probability space, should not be confused with the circular frequencies $\tilde{\omega}_i$ (and ω_i below), $i = 1, \dots, M$.

(iii) $\tilde{G}_i(\xi_1, \eta_1, \dots, \xi_M, \eta_M, t) = \tilde{G}_i(\xi_1, \eta_1, \dots, \xi_M, \eta_M, t, \omega)$, $i = 1, \dots, M$ are \mathcal{F}_t^t measurable stationary random functions satisfying (i) and (iii) of (4.4).

We transform next (4.14) so that we may apply the theorem.

Define the following

$$\omega_i = \sqrt{\tilde{\omega}_i^2 - \alpha_i^2} \tag{4.16}$$

$$A_i(t) = \begin{bmatrix} \cos \omega_i t & \sin \omega_i t \\ -\omega_i \sin \omega_i t & \omega_i \cos \omega_i t \end{bmatrix}, \quad i = 1, \dots, M,$$

$$\tilde{X}^\varepsilon(t) = A^{-1}(t) \begin{bmatrix} \xi_1^\varepsilon(t) \\ d\xi_1^\varepsilon(t)/dt \end{bmatrix} \tag{4.17}$$

$$\tilde{Y}^\varepsilon(t) = \begin{bmatrix} A_2^{-1}(t) \begin{bmatrix} \xi_2^\varepsilon(t) \\ d\xi_2^\varepsilon(t)/dt \end{bmatrix} \\ \vdots \\ A_M^{-1}(t) \begin{bmatrix} \xi_M^\varepsilon(t) \\ d\xi_M^\varepsilon(t)/dt \end{bmatrix} \end{bmatrix}.$$

Let $\tau = \varepsilon^2 t$. Then $X^\varepsilon(\tau) \equiv \tilde{X}^\varepsilon(\tau/\varepsilon^2)$ and $Y^\varepsilon(\tau) \equiv \tilde{Y}^\varepsilon(\tau/\varepsilon^2)$ satisfy a system of equations having the form of (4.3), with τ replacing t . The functions $F^{(1)}$, $F^{(2)}$, $G^{(1)}$, and $G^{(2)}$ are identified in the obvious manner and we shall not write them explicitly. We proceed with the application of the theorem. Let

$$Q_{ij,kl}(t) = E\{\mu_{ij}(t+s)\mu_{kl}(s)\} \tag{4.18}$$

and set

$$\begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = (2\omega_1)^{-2} \int_0^\infty Q_{11,11}(t) \begin{bmatrix} \frac{1}{2} \cos 2\omega_1 t \\ \frac{1}{2} \sin 2\omega_1 t \\ 1 \end{bmatrix} dt, \tag{4.19}$$

$$\begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix} = \sum_{i=2}^M (2\omega_1 \omega_i)^{-1} \int_0^\infty e^{-\alpha_i t} Q_{1i,1i}(t) \sin \omega_i t \begin{bmatrix} \cos \omega_1 t \\ \sin \omega_1 t \end{bmatrix} dt, \tag{4.20}$$

$$\begin{bmatrix} \varkappa_1(r) \\ \varkappa_2(r) \end{bmatrix} = (2\pi\omega_1 r)^{-1} \int_0^{2\pi} E\{\tilde{G}(r \cos \varphi, \omega_1 r \sin \varphi, 0, \dots, 0)\} \begin{bmatrix} \cos \varphi \\ \sin \varphi \end{bmatrix} d\varphi. \tag{4.21}$$

The theorem now asserts that $X^\varepsilon(\tau)$, which takes values in R^2 , converges weakly to a diffusion Markov process $X^0(\tau)$ in R^2 , $0 \leq \tau \leq \tau_0$. Moreover, because of (4.10), moments of $X^\varepsilon(\tau)$ converge to the corresponding moments of $X^0(\tau)$. To describe the limiting Markov process $X^0(\tau)$ it is convenient to represent it in polar coordinates

$$X_1^0(\tau) = r^0(\tau) \cos \theta^0(\tau), \quad X_2^0(\tau) = r^0(\tau) \sin \theta^0(\tau). \tag{4.22}$$

The infinitesimal generator of $(r^0(\tau), \theta^0(\tau))$ is given by

$$\begin{aligned} &\alpha r^2 \partial^2 / \partial r^2 + (3\alpha - \delta_2 - \varkappa_2(r)) r \partial / \partial r + (\gamma + \alpha) \partial^2 / \partial \theta^2 \\ &\quad + (2\beta + \delta_1 - \varkappa_1(r)) \partial / \partial \theta. \end{aligned} \tag{4.23}$$

We omit calculations which are straightforward.

Let us also express $X^\epsilon(\tau)$ in polar coordinates

$$X_1^\epsilon(\tau) = r^\epsilon(\tau) \cos\theta^\epsilon(\tau), \quad X_2^\epsilon(\tau) = r^\epsilon(\tau) \sin\theta^\epsilon(\tau). \tag{4.24}$$

From (4.17) it follows that

$$(r^\epsilon(\tau))^2 = (\xi_1^\epsilon(\tau/\epsilon^2))^2 + \omega_1^{-2} (d\xi_1(\tau/\epsilon^2)/dt)^2$$

i.e., $(r^\epsilon(\tau))^2$ is proportional to the energy of the undamped oscillator. Since the theorem asserts that moments go to moments, we have that

$$\lim_{\epsilon \downarrow 0} E\{(r^\epsilon(\tau))^2\} = E\{(r^0(\tau))^2\} = r_0^2 e^{(8\alpha - 2\delta_2)\tau}, \quad r_0 = r(0) \tag{4.25}$$

$$\lim_{\epsilon \downarrow 0} E\{(r^\epsilon(\tau))^4\} = E\{(r^0(\tau))^4\} = r_0^4 e^{(24\alpha - 4\delta_2)\tau}. \tag{4.26}$$

In (4.25) and (4.26) we have also assumed that $\kappa_1(r) = \kappa_2(r) = 0$ which amounts to $\tilde{G}_1 \equiv 0$. Without this assumption explicit expressions as on the right sides of (4.25), (4.26) cannot be obtained.

An interesting feature revealed by (4.25) and (4.26) is that the damped oscillators can have a pumping as well as a damping effect on the undamped oscillator. The parameters δ_2 in (4.20) can be rewritten as the difference of two positive terms

$$\delta_2 = \sum_{i=2}^M (4\omega_1\omega_i)^{-1} \int_0^\infty e^{-\alpha t} q_{1i,i1}(t) \cos(\omega_1 - \omega_i)t dt - \sum_{i=2}^M (4\omega_1\omega_i)^{-1} \int_0^\infty e^{-\alpha t} q_{1i,i1}(t) \cos(\omega_1 + \omega_i)t dt. \tag{4.27}$$

Thus, it can be either positive or negative depending on the power spectra involved. One naturally expects a damping effect and it is somewhat surprising that pumping is also possible.

References

1. Grad, H.: Proceedings of Symposia in Appl. Math., Vol. 17, pp. 154—183. Providence, R. I.: Amer. Math. Soc. 1965
2. Ellis, R., Pinsky, M.: The first and second fluid approximations to the linearized Boltzmann equation. *J. Math. Pures et Appl.* **54**, 125—156 (1975)
3. Papanicolaou, G. C., Varadhan, S. R. S.: *Comm. Pure Appl. Math.* **26**, 497—524 (1973)
4. Papanicolaou, G. C.: *Bull. A.M.S.* **81**, 330—392 (1975)
5. Haake, F.: Statistical treatment of open systems by generalized master equations. *Springer Tracts in Modern Physics*, Vol. 66. Berlin-Heidelberg-New York: Springer 1973
6. Davies, E. B.: Markovian master equations I *Commun. Math. Phys.* **39**, 91—110 (1974)
7. Kurtz, T. G.: *J. Funct. Anal.* **12**, 55—67 (1973)
8. Hersh, R.: *Rocky Mountain J. Math.* **4**, 443—496 (1974)
9. Frisch, U.: Private communication
10. Papanicolaou, G. C., Kohler, W.: *Comm. Pure Appl. Math.* **27**, 641—668 (1974)
11. Hoppensteadt, F., Miranker, W.: Numerical methods for stiff systems of differential equations related with transistors, tunnel diodes, etc. In: Glowinski, R., Lions, J. L. (Eds.): *Lecture Notes in Computer Science*, Vol. 10, 413—428. Berlin-Heidelberg-New York: Springer 1973
12. Hoppensteadt, F., Miranker, W.: Differential equations having rapidly changing solutions; analytic methods for weakly nonlinear systems. *J. Diff. Eqs.* (to appear)
13. Gihman, I. I., Skorohod, A. V.: *Stochastic differential equations*. Berlin-Heidelberg-New York: Springer 1972
14. Oleinik, O. A., Radkevich, E. V.: *Second order equations with non-negative characteristic form*. New York: Plenum Press 1973

Communicated by J. Moser

(Received March 1, 1975)