STOCHASTIC PROCESSES AND PERTURBATION PROBLEMS DEFINED BY PARABOLIC EQUATIONS WITH A SMALL PARAMETER

MICHAEL A. KOURITZIN

Institute for Mathematics and Its Applications, University of Minnesota, Minneapolis, MN 55455, U.S.A.

1. INTRODUCTION

The anticipation of an interesting limit theory for a system of singularly-perturbed parabolic equations:

$$\partial_\tau v^\varepsilon(\tau, x) = \varepsilon \sum_{|k| \leq q} A_k(\tau, x) \partial_x^k v^\varepsilon(\tau, x) + f^\varepsilon(\tau, x), \quad v^\varepsilon(0, x) = \varphi^\varepsilon(x), \tag{1.1}$$

over time intervals of at least $O(\varepsilon^{-1})$ is justified through the standard almost periodic, statistical mechanics, or probabilistic condition that the time dependence in the coefficients can be averaged out, meaning that

$$A_k^0(x) \doteq \lim_{T \to \infty} \frac{1}{T} \int_0^T A_k(\tau, x) \, d\tau \tag{1.2}$$

exists in some sense for each $0 \leq |k| \leq q$, $x \in \mathbb{R}^d$. It is then useful to perform a time scaling to transform (1.1) into an equivalent equation

$$\partial_t u^\varepsilon(t, x) = \sum_{|k| \leq q} A_k(t/\varepsilon, x) \partial_x^k u^\varepsilon(t, x) + g^\varepsilon(t, x), \quad u^\varepsilon(0, x) = \varphi^\varepsilon(x), \tag{1.3}$$

which is analyzed over regions like $(0, T] \times \mathbb{R}^d$. In this note, we study some limiting (as $\varepsilon \to 0$) properties of $\partial_x^m u^\varepsilon(t, x)$ when q is an even integer, "$\sum_{|k| \leq q}$" denotes the summation over all possible d-tuples k of non-negative integers such that $0 \leq |k| \doteq k_1 + k_2 + \cdots + k_d \leq q$, $\partial_x^k \doteq \partial_{x_1}^{k_1} \partial_{x_2}^{k_2} \cdots \partial_{x_d}^{k_d}$, and $\{A_k(s, x), s \geq 0\}$ is a $\mathbb{C}^{N \times N}$-valued, possibly random, process for each $x \in \mathbb{R}^d$. Suppose we define

$$\alpha_k^\varepsilon(t, x) \doteq \int_0^t \overline{A}_k(s/\varepsilon, x) \, ds \doteq \int_0^t [A_k(s/\varepsilon, x) - A_k^0(x)] \, ds \quad \forall t \in [0, T], \; x \in \mathbb{R}^d. \tag{1.4}$$

Then, $\alpha_k^\varepsilon(1, x) \to 0$ will be used in lieu of (1.2) for (pointwise in x) deterministic, sure, or almost sure convergence whereas $\alpha_k^\varepsilon(1, x) \Rightarrow 0$ will be used for pointwise convergence in distribution.

Notwithstanding the two original contributions contained herein, the primary purpose of our document is to encourage contemplation of the possible impact, interconnections, and applications of our three principle, essentially mathematical references Kouritzin [5],[6], and Dawson and Kouritzin [1]. Accordingly, we leave the historical account of averaging and the discussion of previous related works to the introductions of [5], [6], and [1]. Moreover, we will label the new results as corollaries and consider them as applications of Theorems 1 and 3. From a pragmatic point of view, the

interesting outcroppings of our work possibly include approximating Fokker Planck equations having random or almost periodic coefficients by simpler equations, approximating the higher order parabolic densities of certain composite processes (see e.g. Hochberg and Orsingher [3]) by higher order averaged parabolic equations, and the fact that the limiting object in our invariance principle satisfies a general linear stochastic partial differential equation of arbitrary order driven by a general noise source that need not be Gaussian, Markov, or a semimartingale. Moreover, suppose, motivated by Khas'minskii [4] and Kushner and Huang [8], we include the related problem of convergence for u^{ε} over expanding intervals $[0, T\varepsilon^{2h-2}]$ to a certain nonlinear stochastic partial differential equation. Then, we can add filtering and nonlinear stochastic partial differential equations to our list of potential application (see Kushner and Huang [8]). Still, it is our strong opinion that the major contribution in all this work is twofold: (i) Extending the underlying theory of arbitrary order parabolic equations to promote the study of the asymptotics of such perturbation problems; and (ii) the general *transfer type*, functional analytic methods we develop to prove our limit theorems. We only assume a limit theorem on the coefficients and then we infer one for the partial differential equation solutions and we believe that these transfer methods can be used for asymptotic problems on other media.

The Fundamental solution of a parabolic equation provides an elegant method of determining how this equation will propagate given initial data and/or forcing functions. Therefore, it is not surprising that much is exposed about our system (1.3) from a collection of like-constructed fundamental solutions $\{\Gamma^{\varepsilon}, 0 < \varepsilon \leq 1\}$. For instance, uniform (in ε) bounds on the solutions u^{ε} are establishable from uniform bounds of the correct type on $\{\Gamma^{\varepsilon}, 0 < \varepsilon \leq 1\}$ as well as bounds on g^{ε} and φ. However, we have already encountered our first quandary: To cede uniform (in ε) bounds for Γ^{ε} the classical theory would require that the principle coefficients $\{A_k(t/\varepsilon, x)\}_{|k|=q}$ are continuous in t uniformly with respect to $(x, t, \varepsilon) \in \mathbb{R}^d \times [0, \infty) \times (0, 1]$ which in turn would disallow dependence on ε or t and lead to a rather uninteresting averaging theory. Fortunately, this apparent dilemma has already been circumvented in [5] by showing that a modest strengthening of the classical uniformly parabolic condition eliminates any need for uniformity in the continuity of $t \to A_k(t/\varepsilon, x)$, $|k| = q$. Thus, for all results in the sequel we will assume:

(**C1**) The system (1.3) is uniformly parabolic in the sense that

$$- \sup_{t \geq 0} \sup_{x \in \mathbb{R}^d} \max_j \sup_{|\xi|=1} \lambda_j(\xi; x, t) > 0, \tag{1.5}$$

where, for all $\xi, x \in \mathbb{R}^d$, and $t \geq 0$, $\{\lambda_j(\xi; x, t)\}_{j=1}^{2N}$ are the (real) roots of

$$\det \left(\sum_{|k|=2p} \begin{bmatrix} \mathrm{Re}[A_k(x,t) + A_k^T(x,t)] & -\mathrm{Im}[A_k(x,t) - A_k^T(x,t)] \\ \mathrm{Im}[A_k(x,t) - A_k^T(x,t)] & \mathrm{Re}[A_k(x,t) + A_k^T(x,t)] \end{bmatrix} (i\xi)^k - \lambda I_{2N} \right). \tag{1.6}$$

When the coefficients are random, (1.5) is interpreted to hold for almost all ω.

In order to make the asymptotic properties of $\partial_x^m u^{\varepsilon}$ manifest, we introduce a limiting or averaged deterministic differential equation:

$$\partial_t u(t, x) = \sum_{|k| \leq q} A_k^0(x) \partial_x^k u(t, x) + g^0(t, x), \quad u(0, x) = \varphi(x). \tag{1.7}$$

Since there is no ε- or t-dependence on the coefficients of (1.7) we can resort back to the usual, slightly more general parabolic condition:

(**C2**) (1.7) is uniformly parabolic in the sense that

$$- \sup_{x \in \mathbb{R}^d} \max_j \sup_{|\xi|=1} \mathrm{Re} \left\{ \lambda_j^0(\xi; x) \right\} > 0, \tag{1.8}$$

where, for all $\xi, x \in \mathbb{R}^d$, $\left\{\lambda_j^0(\xi; x)\right\}_{j=1}^N$ are the roots of the polynomial

$$\det\left(\sum_{|k|=2p} A_k^0(x)(i\xi)^k - \lambda I_N\right). \tag{1.9}$$

Under these parabolic conditions plus general, problem-specific bounds and Hölder continuity conditions on the coefficients and their derivatives, a highly developed theory for the existence of fundamental solutions, uniqueness of particular solutions, bounds for derivatives of the fundamental solutions, and solutions to the adjoint problem now exists. This theory facilitates and was partially developed for the study of the asymptotics of equations like (1.1) or (1.3). In the following section, we consider problems of averaging for the derivatives up to order $q - 1$ of fundamental solutions to (1.3). This theory is then used to prove a result on "total averaging" that is finding general conditions under which $\partial_x^m u^\varepsilon$ converges uniformly to $\partial_x^m u$ when $g^\varepsilon(t, x) = g(t/\varepsilon, x)$, $g^0(t, x) = g^0(x)$ does not depend on t, and $\varphi^\varepsilon = \varphi = 0$. In Section 3, we let $g^\varepsilon(t, x) = g^0(t, x) = 0$ and consider averaging for arbitrary order derivatives of the Cauchy problem. This is then put into a stochastic setting and given a strong law of large number interpretation from which we show how to infer a weak law of large numbers under weaker probabilistic conditions. In Section 4, we turn to the question of rates of convergence or fluctuation results and show that depending on the type of dependence exhibited by the coefficients there is an $h \in (0, 1)$ such that $\varepsilon^{h-1}(u^\varepsilon - u) \overset{\varepsilon \to 0}{\Rightarrow} \widehat{y}$, the mild solution to a linear stochastic partial differential equation driven by some continuous, Hilbert space valued noise source like Brownian motion or fractional Brownian motion (which is neither Markov nor a semimartingale).

We now summarize the notation which is germane to all of the remaining sections: For notational ease we will always take $T = 1$ and let $I \doteq (0, 1]$. $|\cdot|$ will denote Euclidean distance in $\mathbb{R}^d$ and $\mathbb{C}^N$ and $||\cdot||$ will be used for the $|\cdot|$−induced norm for $\mathbb{C}^{N \times N}$ matrices. (For technical reasons, it was convenient to define our norms in a different manner in the referenced works but finite dimensional norms are, of course, compatible.) Finally, $e_i \doteq (0, ..., 0, 1, 0, ..., 0)^T \in \mathbb{R}^d$ with the 1 in the ith row, $a \vee b$ and $a \wedge b$ will be used to denote the maximum respectively minimum of two real numbers a, b, and $a_{m,n} \overset{n,m}{\ll} b_{m,n}$ will imply that there is a constant $c > 0$ such that $|a_{m,n}| \leq c\, |b_{m,n}|$ for all n, m.

2. FUNDAMENTAL SOLUTION AND TOTAL AVERAGING

Theorem 1 of the current section is based on the work in [5]. As mentioned in the introduction, one of the main obstacles in obtaining an asymptotic theory for our parabolic equations is establishing uniform bounds on certain derivatives of the fundamental solutions to (1.3). This is addressed in Theorem A and Lemma 5 of [5] as well as in [1] and [6]. However, there is no need to involve ourselves in these technical issues here so we will proceed to discuss our results.

In addition to (C1,C2) we require the following Conditions in this section:

(C3) For each $|k| \leq 2p$: **(i)** A_k is continuous in t over $[0, \infty)$, and **(ii)** A_k and A_k^0 are uniformly bounded in $\mathbb{R}^d \times [0, \infty)$ respectively $\mathbb{R}^d$.

(C4) For each $|k| \leq 2p$, $\partial_{x_i} A_k$ and $\partial_{x_i} A_k^0$ exist and are uniformly bounded in $\mathbb{R}^d \times [0, \infty)$ respectively $\mathbb{R}^d$ for $i = 1, ..., d$.

(C5) $\partial_{x_i} A_k$ and $\partial_{x_i} A_k^0$ are Hölder continuous in x with exponent $0 < \varsigma \leq 1$ uniformly in $\mathbb{R}^d \times [0, \infty)$ respectively $\mathbb{R}^d$ for $i = 1, ..., d$ when $|k| = 2p$.

Theorem 1. *Suppose the Regularity Conditions (C1-C5) hold and $\alpha^\varepsilon(1, y) \overset{\varepsilon \to 0}{\Rightarrow} 0$ pointwise for each $y \in \mathbb{R}^d$. Then, for any $0 < \chi, \nu < 1$ it follows that there exists a positive constant $\widetilde{c} = \widetilde{c}_{\chi,\nu}$ and a $\mathbb{R}$-valued function $\gamma(\cdot) = \gamma_{\chi,\nu}(\cdot)$ satisfying $\lim_{\varepsilon \to 0} \gamma(\varepsilon) = 0$ such that*

$$\left\|\partial_x^b[\Gamma^\varepsilon(x, t; \xi, \tau) - \Gamma(x, t - \tau, \xi)]\right\| \leq \gamma(\varepsilon)\left|1 + |\xi|^2\right|^\nu (t - \tau)^{-\frac{d+|b|+\chi}{q}} h_{\widetilde{c}}(x - \xi, t - \tau) \tag{2.1}$$

for all $0 \leq |b| < q, \varepsilon \in (0,1], 0 \leq \tau \leq t \leq 1$ and $x, \xi \in \mathbb{R}^d$, where $\Gamma^\varepsilon, \Gamma$ are the fundamental solutions to (1.3),(1.7) and $h_c(y,s) \doteq \exp\{-c\, s^{-1/(q-1)} |y|^{q/(q-1)}\}$.

Comparing the right hand side of (2.1) to classical bounds for a single fundamental solution in, for example, Theorem 9.4.2 of Friedman [2], one can see that they are the same with the exception that (2.1) has the extra multiplicative term $\gamma(\varepsilon) |1 + |\xi|^2|^\nu (t-\tau)^{-\chi/q}$. Whereas the first factor in this term establishes that $\partial_x^b \Gamma^\varepsilon$ will approach $\partial_x^b \Gamma$ as $\varepsilon \to 0$ the remaining two factors (or alternative factors which grow slowly to infinity as $\xi \to \infty$ and $t - \tau \to 0$) are required to allow a uniform result over all $0 \leq \tau \leq t \leq 1$ and $x, \xi \in \mathbb{R}^d$. Still, these last two factors can be taken small enough as not to encumber our main goal of averaging for systems like (1.3) as the following corollary on total averaging portrays.

Corollary 2. *Suppose the conditions of Theorem 1 hold and h and h^0 are continuous, bounded functions on $[0,\infty) \times \mathbb{R}^d$ and $\mathbb{R}^d$ respectively such that*

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T h(s,y)\,ds = h^0(y) \quad \forall y \in \mathbb{R}^d. \tag{2.2}$$

Moreover, suppose h and h^0 are Hölder continuous in y uniformly on compact sets (with an exponent which can depend on the set). Then, it follows that $\sup_{t \in I} |u^\varepsilon(t) - u(t)|_q \overset{\varepsilon \to 0}{\to} 0$, where

$$u^\varepsilon(t,x) \doteq -\int_0^t \int_{\mathbb{R}^d} \Gamma^\varepsilon(t,x;\tau,\xi)\, h(\tfrac{\tau}{\varepsilon},\xi)\,d\xi d\tau, \; u(t,x) \doteq -\int_{\mathbb{R}^d} \int_0^t \Gamma(x,\tau,\xi) d\tau\, h^0(\xi)\,d\xi, \tag{2.3}$$

$\Gamma^\varepsilon, \Gamma$ being the fundamental solutions to (1.3,1.7), and

$$|f|_q \doteq \sum_{|k| < q} \sup_{x \in \mathbb{R}^d} |\partial_x^k f(x)|/(1 + |x|^2)^\nu. \tag{2.4}$$

Remark 1. *We chose to label the quantities in (2.3) u^ε and u because they will correspond to solutions to (1.3) and (1.7) respectively subject to $\varphi^\varepsilon = \varphi \equiv 0$, $g^\varepsilon(t,x) = h(\tfrac{t}{\varepsilon},x)$, and $g^0(t,x) = h^0(x)$. Hence, this corollary illustrates total averaging as mentioned in the introduction. Under our general conditions, u^ε and u may not be the only bounded solutions to (1.3) and (1.7) so we had to define them explicitly in terms of the fundamental solutions constructed in the classical manner.*

Proof of Corollary 2. Fix a $0 \leq |k| < q$ and let $u_k \doteq \partial_x^k u$, $u_k^\varepsilon \doteq \partial_x^k u^\varepsilon$. Then, the ∂_x^k commutes with the integrals and, by Theorem 1, one finds that

$$|u_k^\varepsilon(t,x) - u_k(t,x)| \leq \gamma(\varepsilon) \int_0^t \int_{\mathbb{R}^d} |1 + |\xi|^2|^\nu (t-\tau)^{-\frac{d+|k|+\chi}{q}} h_{\underset{\sim}{c}}(x-\xi, t-\tau) d\xi\, d\tau \tag{2.5}$$

$$+ \left| \int_0^t \int_{\mathbb{R}^d} \partial_x^k \Gamma(x, t-\tau, \xi)[h(\tau/\varepsilon, \xi) - h^0(\xi)]\, d\xi d\tau \right| \quad \forall t \in (0,1].$$

However, using the facts $|1 + |\xi|^2|^\nu/|1 + |x|^2|^\nu \overset{x,\xi}{\ll} 1 + |\xi - x|^{2\nu}$ and $\int_0^t \int_{\mathbb{R}^d} s^{-(d+|k|+\chi)/q} h_c(y,s) dy ds < \infty$ for any $c > 0$, all $0 \leq \chi < 1$, and all $|k| \leq q - 1$, one easily finds that

$$|u_k^\varepsilon(t,x) - u_k(t,x)|/(1 + |x|^2)^\nu \overset{t,x,\varepsilon}{\ll} \gamma(\varepsilon) + |q^\varepsilon(t,x)|/(1 + |x|^2)^\nu, \tag{2.6}$$

where

$$q^\varepsilon(t,x) \doteq \int_0^t \int_{\mathbb{R}^d} \partial_x^k \Gamma(x, t-\tau, \xi)[h(\tau/\varepsilon, \xi) - h^0(\xi)]\, d\xi d\tau. \tag{2.7}$$

Now, by classical bounds (Theorem 9.6.7 of Friedman with $r = 1$) there is a $c > 0$ such that

$$\|\partial_x^{k+a}\Gamma(x,s,\xi)\| \overset{x,s,\xi}{\ll} s^{-(d+|k|+|a|)/q} h_c(x-\xi,s) \quad \forall x,\xi \in \mathbb{R}^d,\ s \in I,\ 0 \le |a| \le 1 \tag{2.8}$$

so it follows easily by reusing the second fact above (2.6) that

$$|q^\varepsilon(t,x)| \overset{t,x,\varepsilon}{\ll} 1 \quad \forall x \in \mathbb{R}^d,\ s \in I,\ \varepsilon \in (0,1]\ . \tag{2.9}$$

Moreover; fixing $x \in \mathbb{R}^d$, $\Delta \in \mathbb{R}$, and $i \in \{1,...,d\}$; and defining $x' = x + \Delta e_i$; we note that if $|\Delta| < s^{1/q}$ then we have by the mean value theorem and (2.8)

$$\int_{\mathbb{R}^d} \|\partial_x^k[\Gamma(x,s,\xi) - \Gamma(x',s,\xi)]\|\, d\xi \overset{x,x',s}{\ll} |\Delta|^{\frac{1}{2}} \int_{\mathbb{R}^d} s^{\frac{1/2-d-q}{q}} h_c(y-\xi,s)\, d\xi \overset{x,x',s}{\ll} |\Delta|^{\frac{1}{2}} s^{\frac{1}{2q}-1} \tag{2.10}$$

for some y between x and x'. Conversely, if $|\Delta| \ge s^{1/q}$ then we have by (2.8) that

$$\int_{\mathbb{R}^d} \|\partial_x^k\Gamma(x',s,\xi)\|\, d\xi \overset{s,x'}{\ll} |\Delta|^{\frac{1}{2}} \int_{\mathbb{R}^d} s^{\frac{1/2-d-q}{q}} h_c(x'-\xi,s)\, d\xi \overset{s,x,x'}{\ll} |\Delta|^{\frac{1}{2}} s^{\frac{1}{2q}-1}. \tag{2.11}$$

Using (2.10) and (2.11) repeatedly with $s = t - \tau$, one easily finds that

$$|q^\varepsilon(t,x) - q^\varepsilon(t,x')| \overset{\varepsilon,x,x',t}{\ll} |x - x'|^{\frac{1}{2}} \quad \forall x,x' \in \mathbb{R}^d,\ t \in I,\ \varepsilon \in (0,1]. \tag{2.12}$$

Unfortunately, we will not be able to get away with the classical bounds for the fundamental solutions in order to obtain a similar type bound in t but rather we must turn to Proposition 3 (iii) of Kouritzin [6] to find under our conditions that

$$\|\partial_s\partial_x^k\Gamma(x,s,\xi)\| \overset{x,s,\xi}{\ll} s^{-(d+|k|+q)/q} h_c(x-\xi,s) \quad \forall x,\xi \in \mathbb{R}^d,\ s \in I = (0,1]. \tag{2.13}$$

Therefore, it follows easily from (2.7), (2.8), and (2.13) that for any $t > t'$

$$|q^\varepsilon(t,x) - q^\varepsilon(t',x)| \tag{2.14}$$
$$\le \int_{t'}^t \int_{\mathbb{R}^d} \|\partial_x^k\Gamma(x,t-\tau,\xi)\|\, d\xi d\tau + \int_{t'}^t \int_0^{t'} \int_{\mathbb{R}^d} \|\partial_\sigma\partial_x^k\Gamma(x,\sigma-\tau,\xi)\|\, d\xi d\tau d\sigma$$
$$\overset{\varepsilon,x,t',t}{\ll} \int_{t'}^t \left[(t-\sigma)^{\frac{1-q}{q}} + \int_0^{t'}(\sigma-\tau)^{\frac{1-2q}{q}}d\tau\right] d\sigma \overset{\varepsilon,x,t',t}{\ll} |t-t'|^{1/q}.$$

Equicontinuity for $(t,x) \to q^\varepsilon(t,x)$, $\varepsilon \in (0,1]$ follows from (2.12) and (2.14). It then follows by extending these functions uniquely to $[0,1] \times \mathbb{R}^d$, (2.9), this equicontinuity, and the arguments in (66-77) of Kouritzin [5] that $\{r^\varepsilon,\ 0 < \varepsilon \le 1\}$ is relatively compact in the continuous bounded functions $C_B(I \times \mathbb{R}^d)$, where

$$r^\varepsilon(t,x) \doteq q^\varepsilon(t,x)/(1+|x|^2)^\nu \quad \forall t \in I,\ x \in \mathbb{R}^d. \tag{2.15}$$

Now, we show that the only possible limit point for r^ε is 0. Indeed, suppose we fix $t \in I$ and $x \in \mathbb{R}^d$. Then, it follows from (2.8) that

$$\int_{t-\delta}^t \int_{\mathbb{R}^d} |\partial_x^k\Gamma(x,t-\tau,\xi)[h(\tau/\varepsilon,\xi) - h^0(\xi)]|\, d\xi d\tau \overset{\varepsilon,\delta}{\ll} \delta^{1/q} \quad \forall 0 \le \delta \le t. \tag{2.16}$$

and from integration by parts that

$$\int_{\mathbb{R}^d} \int_0^{t-\delta} \partial_x^k \Gamma(x, t-\tau, \xi)[h(\tau/\varepsilon, \xi) - h^0(\xi)]d\tau \, d\xi \tag{2.17}$$

$$= \int_{\mathbb{R}^d} \partial_x^k \Gamma(x, \delta, \xi) \int_0^{t-\delta} [h(s/\varepsilon, \xi) - h^0(\xi)]ds \, d\xi$$

$$- \int_{\mathbb{R}^d} \int_0^{t-\delta} \partial_\tau \partial_x^k \Gamma(x, t-\tau, \xi) \int_0^\tau [h(s/\varepsilon, \xi) - h^0(\xi)]ds \, d\tau \, d\xi \quad \forall 0 < \delta < t.$$

Pointwise convergence for r^ε follows by (2.16), our hypothesis (2.2), (2.17), (2.13), (2.8), and the dominated convergence theorem. The result is then proved by (2.6), (2.7), and the relative compactness. ∎

3. LIMITS FOR HIGH ORDER DERIVATIVES

As borne out by juxtaposing the statements and proofs of Theorems 9.4.2 and 9.6.7 in Friedman [2], the relatively wieldy methods and mild conditions suitable for proving the classical bounds on the derivatives up to order $q-1$ of a fundamental solution are inadequate for establishing similar bounds for higher order derivatives. Hence, there is little hope that the methods and conditions used in establishing Theorem 1 are sufficient for higher order derivative analogues. Yet, comparing the high order x-derivatives of u^ε to those of u is an important problem. Indeed, in the case where $g^\varepsilon = g \equiv 0$ i.e. the Cauchy initial data problems, this type of comparison appears crucial for obtaining fluctuation results characterizing the rate of convergence of u^ε into u. Hence, in this section, we will consider approximating arbitrary order derivatives of u^ε with those of u for the initial data problem under necessarily more stringent conditions on the coefficients and initial data than those of the previous section. The main result of this section, Theorem 3, is proved in Kouritzin [6].

Throughout this section, $(l_1, ..., l_d)$ and $(b_1, ..., b_d)$ are fixed vectors of non-negative integers such that $l_1 + l_2 + \cdots + l_d \leq q$. Moreover, we also introduce a partial ordering on this set of non-negative integer vectors by letting $k \leq m$ denote the relation $k_i \leq m_i$ for all $i = 1, 2, ..., d$ and $k < m$ mean $k \leq m$ and $k \neq m$. Making f denote any continuous function on $I \times \mathbb{R}^d$, we define for each $\gamma, \nu > 0$

$$|f|_{l,\gamma} \doteq \sup_{t,x} \frac{|t^{\frac{(|l| - \gamma/2) \vee 0}{2p}} f(t,x)|}{(1 + |x|^2)^\nu} + \sup_{\substack{t,t' \\ 0 < |x - x'| \leq 1}} \frac{|t^{\frac{(|l| - \gamma/2) \vee 0}{2p}} f(t,x) - t'^{\frac{(|l| - \gamma/2) \vee 0}{2p}} f(t',x')|}{(1 + |x|^2)^\nu \left[|t - t'|^{\frac{\gamma}{2p}} + |x - x'|^\gamma\right]}. \tag{3.1}$$

Our additional regularity required for this section is:

(C3) For each $|k| \leq 2p$ and $|m| \leq |b|$, $\partial_x^m A_k$ and $\partial_x^m A_k^0$ exist and are continuous and uniformly bounded on $\mathbb{R}^d \times [0, \infty)$ respectively $\mathbb{R}^d$.

(C4) When $|k| \leq 2p$ and $|m| = |b|$, $\partial_x^m A_k$ and $\partial_x^m A_k^0$ are Hölder continuous in x with exponent $0 < \varsigma \leq 1$ uniformly on $\mathbb{R}^d \times [0, \infty)$ respectively $\mathbb{R}^d$.

(C5) For each $m \leq b$, $\partial_x^m \varphi$ exists and is a continuous, bounded function on $\mathbb{R}^d$. Furthermore, $\partial_x^b \varphi$ is Hölder continuous in x with exponent $0 < \varsigma \leq 1$ uniformly on $\mathbb{R}^d$.

(C6) There is only one continuous, bounded solution of

$$\partial_t v(t,x) = \sum_{|k| \leq 2p} A_k^0(x) \partial_x^k v(t,x) + f(t,x) \quad \text{subject to } v(0,x) = \varphi(x) \tag{3.2}$$

for any continuous, bounded f on $I \times \mathbb{R}^d$ which is Hölder continuous in x uniformly on bounded subsets of $\mathbb{R}^d$.

In contrast to the previous section, the methods used for the results in this section do require uniqueness for (3.2). However, in the case where $|b| \geq q$, (C6) follows automatically from Conditions

(C2-C4) and Theorem 9.5.6 of Friedman [2] and need not be included in the assumptions. Uniqueness in (1.3) is not required and by Condition (C5) we can define

$$u^\varepsilon(t,x) \doteq \int_{I\!R^d} \Gamma^\varepsilon(x,t;\xi,0)\,\varphi(\xi)\,d\xi \quad \forall t \in I,\, x \in \mathbb{R}^d. \tag{3.3}$$

Moreover, under Condition (C6), we know from Friedman p. 257 that the unique solution to (1.7) when $g^0 \equiv 0$ is given by

$$u(t,x) \doteq \int_{I\!R^d} \Gamma(x,t;\xi,0)\varphi(\xi)d\xi. \tag{3.4}$$

Now, in preparation of stating the main result of this section (to follow immediately), we set $\gamma \doteq \varsigma/12$, let $\nu \in (0,1/4)$, and define $\alpha_{k,m}^\varepsilon(t,x) \doteq \partial_x^m \alpha_k^\varepsilon(t,x) = \int_0^t \partial_x^m[A_k(\tau/\varepsilon,x) - A_k^0(x)]\,d\tau$ for all $t \in I$, $x \in \mathbb{R}^d$, $|k| \le q$, $m \le b$.

Theorem 3. *Suppose Regularity Conditions (C1-C6) hold and for each (fixed) $x \in \mathbb{R}^d$, $|k| \le q$, and $m \le b$ one has that $\alpha_{k,m}^\varepsilon(1,x) \stackrel{\varepsilon \to 0}{\Rightarrow} 0$. Then, it follows that*

$$\left|\partial_x^{l+b}(u^\varepsilon - u)\right|_{l,\gamma} \stackrel{\varepsilon \to 0}{\Rightarrow} 0. \tag{3.5}$$

Recalling (3.1), one finds that the convergence in (3.5) is uniform with respect to $(t,x) \in I \times \mathbb{R}^d$ and based upon a weighted Hölder continuity norm. This may at first appear somewhat contrived or unnatural. However, this convergence problem can be well motivated by the proof of our fluctuation problem stated in the following section and represents a very strong form of convergence for the conditions that we have set. Furthermore, even for the low order derivatives, Theorem 3 represents somewhat of an improvement over Theorem 1 applied to initial data problems: While Theorem 1 could be used to compare $\partial_x^l u^\varepsilon$ to $\partial_x^l u$ for $1 \le |l| < q$ the comparison would blow up as t approaches 0. On the other hand, Theorem 3 allows one to control or eliminate this explosion for the same order derivative at the price of more stringent assumptions by choosing $b > 0$.

Allowing the coefficients in (1.3) but not (1.7) to be random on some probability space $(\Omega, \mathcal{F}, P)$, assuming that $\alpha_{k,m}^\varepsilon(1,x,\omega) \stackrel{\varepsilon \to 0}{\Rightarrow} 0$ for almost all $\omega \in \Omega$, and applying Theorem 3 pathwise, we discover that we have a strong law of large numbers or rather a *transfer principle* between pointwise strong laws of large numbers for the coefficients and uniform strong laws of large numbers for arbitrary derivatives of singularly-perturbed Cauchy problem solutions. Natural questions to pose at this point are: (i) Can one establish such a transfer principle for weak laws of large numbers? and (ii) Is there a reasonable fluctuation theory characterizing the rate of convergence of u^ε to u? Whereas the first of these questions is answered by the following corollary, the second question is addressed in the following section in the form of a very general invariance principle.

Corollary 4. *Suppose Regularity Conditions (C1-C6) hold and for each (fixed) $x \in \mathbb{R}^d$, $|k| \le q$, and $m \le b$ one has that $\alpha_{k,m}^\varepsilon(1,x) \stackrel{\varepsilon \to 0}{\Rrightarrow} 0$. Then, it follows that*

$$\left|\partial_x^{l+b}(u^\varepsilon - u)\right|_{l,\gamma} \stackrel{\varepsilon \to 0}{\Rrightarrow} 0. \tag{3.6}$$

Proof. Let $\{x_j\}_{j=1}^\infty$ be a dense subset of $\mathbb{R}^d$ and $\{\varepsilon_r\}_{r=1}^\infty$ be a sequence of the ε. Then, it follows from our hypothesis and a simple triangle argument that for almost all $\omega \in \Omega$ there is a subsequence $\{\varepsilon_i\}_{i=1}^\infty$ of $\{\varepsilon_r\}_{r=1}^\infty$ such that

$$\alpha_{k,m}^{\varepsilon_i}(1,x_j) \stackrel{i \to \infty}{\Rightarrow} 0 \quad \forall j = 1,2,...,|k| \le q, m \le b. \tag{3.7}$$

Fix an ω such that (3.7) holds. Then, equicontinuity of $x \to \alpha_{k,m}^{\varepsilon_i}(1,x)$, $i = 1, 2, ..., |k| \leq q$, $m \leq b$ is easily established by Conditions (C3,C4) and it follows easily by (3.7) that

$$\alpha_{k,m}^{\varepsilon_i}(1,x) \stackrel{i \to \infty}{\Rightarrow} 0 \quad \forall x \in \mathbb{R}^d, \ |k| \leq q, \ m \leq b. \tag{3.8}$$

Now, it follows trivially from the proof in [6] that Theorem 3 also holds for subsequences. Hence, we can conclude from (3.8) that

$$\rho(\mathcal{L}\left(\left.\left|\partial_x^{l+b}(u^{\varepsilon_i} - u)\right|\right|_{l,\gamma}\right), \delta_0) \stackrel{i \to \infty}{\Rightarrow} 0, \tag{3.9}$$

where ρ denotes the Prokhorov metric and $\mathcal{L}$ means distribution or law on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. The corollary follows by the arbitrariness of $\{\varepsilon_r\}_{r=1}^{\infty}$. $\blacksquare$

4. INVARIANCE PRINCIPLES

As opposed to the previous two sections where the coefficients could be deterministic, we assume forthright in this section that $(\Omega, \mathcal{F}, P)$ is a probability space and $\{A_k(s,x), s \geq 0\}$ is a $\mathbb{C}^{N \times N}$-valued stochastic process on $(\Omega, \mathcal{F}, P)$ for each $0 \leq |k| \leq q, x \in \mathbb{R}^d$. Moreover, we will now be interested in rates of convergence for the Cauchy problem in this stochastic setting instead of just convergence itself. Hence, we will still assume $g^{\varepsilon} = g \equiv 0$ and $\varphi^{\varepsilon} = \varphi$ but now we will be considering a weak invariance principle for $\varepsilon^{h-1}(u^{\varepsilon} - u)$ under general conditions, $h \in (0,1)$ being a fixed constant which is determined by the "amount" of dependence in t the coefficients exhibit. This result is based upon the work in Dawson and Kouritzin [1] and brings forth our rate of convergence.

When interpreted in a stochastic setting, the results of the previous section had the nice feature of really being *transfer methods*. Instead of assuming apriori specific stationarity, moment, and dependency conditions on the coefficients to prove our weak and strong laws of large numbers we just assumed that the coefficients themselves satisfied either a (pointwise) weak or strong law of large numbers. This had the advantage of yielding very general results handling several cases at once. We follow a similar approach here by just assuming the coefficients themselves satisfy a weak invariance principle:

$$\mathcal{A}^{\varepsilon} \Rightarrow \widehat{\Theta} \quad \text{in} \ \ C(I; \mathcal{H}_1) \ \ \text{as} \ \ \varepsilon \to 0, \tag{4.1}$$

where $I \doteq [0,1]$ for this section, $\left\{\widehat{\Theta}_t, \ t \in I\right\}$ can be any $C(I; \mathcal{H}_1)$-valued random object,

$$\mathcal{A}_m^{\varepsilon}(t,x) \doteq \varepsilon^h \int_0^{t \varepsilon^{-1}} \overline{A}_m(\tau, x) \, d\tau \quad \forall x \in \mathbb{R}^d, \ t \in I, \ \varepsilon \in (0,1], \ |m| \leq q \tag{4.2}$$

and $\mathcal{A}^{\varepsilon}$ denotes the $\mathcal{H}_1$-valued (see Condition (C3) below) process $\{\mathcal{A}_m^{\varepsilon}\}_{|m| \leq q}$. $\mathcal{H}_1 \doteq \bigotimes_{|k| \leq q} H_1$ is the Hilbert space of all possible $\{f_k\}_{|k| \leq q}$ such that each $f_k \in H_1$ and $\left\|\{f_k\}_{|k| \leq q}\right\|_1^2 \doteq \sum_{|k| \leq q} \|f_k\|_1^2$. Here, $(H_1, <,>_1)$ (respectively H_1^v) denote the separable Hilbert spaces of weighted $\mathbb{C}^{N \times N}$-valued (respectively $\mathbb{C}^N$-valued) $L^2(\mathbb{R}^d)$-functions f which are also Hölder continuous on average in the sense that

$$\|f\|_1 \doteq \sqrt{\int_{\mathbb{R}^d} |1 + |x|^2|^{-d} \left\{ \|f(x)\|^2 + \int_{B_x} |x - \xi|^{-d - \gamma} \|f(\xi) - f(x)\|^2 \, d\xi \right\} dx} < \infty, \tag{4.3}$$

B_x being the open unit ball of $\mathbb{R}^d$ centered at x and γ being a constant such that $0 < \gamma < \frac{1}{4}$. Finally, the constant h is used for universality: $h = 1/2$ provides the weighting for a Hilbert space version

of the classical functional or Donsker-type central limit theorem and $h < 1/2$ allows for long range t dependence in the coefficients.

The remaining regularity required for this section is:

(C3) $\partial_x^k A_m$ and $\partial_x^k A_m^0$ exist and are continuous and uniformly bounded on $[0, \infty) \times \mathbb{R}^d$ respectively $\mathbb{R}^d$ for all $0 \le |m| \le q, 0 \le |k| \le q + 2$.

(C4) $\partial_x^k \varphi$ exists and is a bounded, continuous function on $\mathbb{R}^d$ for all $0 \le |k| \le q + 2$.

In preparation for stating our main result, we define $(H_2, <, >_2)$ to be the Hilbert space of weighted $\mathbb{C}^N$-valued $L^2(\mathbb{R}^d)$-functions f such that $|f|_2 \doteq \sqrt{\int_{\mathbb{R}^d} (1 + |x|^2)^{-2d} |f(x)|^2 \, dx} < \infty$, let $T_s, s \ge 0$ denote the continuous operators from $H_1^v \to H_2$ defined by

$$T_s f(x) \doteq \int_{\mathbb{R}^d} \Gamma(x, s, \xi) f(\xi) \, d\xi \qquad \forall \ x \in \mathbb{R}^d, \ s \in (0, 1], \tag{4.4}$$

and recall our definition of $\alpha_{k,m}^\varepsilon(t, x)$ following (3.4).

Theorem 5. *Suppose Conditions (C1-C4) as well as (4.1) hold, $\widehat{\Theta}$ is a $C(I; \mathcal{H}_1)$-valued stochastic process and $\alpha_{k,m}^\varepsilon(1, x) \Rightarrow 0$ as $\varepsilon \to 0$ for each $|k| \le q, |m| \le q + 1, x \in \mathbb{R}^d$. Then, the stochastic integral*

$$\{\widehat{y}(t), 0 \le t \le 1\} \doteq \left\{ -\sum_{|k| \le q} \int_0^t T_{t-\tau} \, d\widehat{\Theta}_k(\tau) \, u_k(\tau), 0 \le t \le 1 \right\} \qquad a.s., \tag{4.5}$$

interpreted in the sense of Definition 7 (to follow), exists and

$$\begin{bmatrix} \varepsilon^{h-1} (u^\varepsilon - u) \\ \mathcal{A}^\varepsilon \end{bmatrix} \Rightarrow \begin{bmatrix} \widehat{y} \\ \widehat{\Theta} \end{bmatrix} \qquad in \ \ C(I; H_2 \times \mathcal{H}_1) \quad as \ \ \varepsilon \to 0. \tag{4.6}$$

Under a variety of probabilistic conditions, one finds that $\alpha_{k,m}^\varepsilon(1, x)$ satisfies a weak law of large numbers (for each fixed (m, k, x)) and $\mathcal{A}^\varepsilon$ satisfies a weak invariance principle in $\mathcal{H}_1$. Our main thesis is that these two weak convergence assumptions are sufficient to establish that $\varepsilon^{h-1} (u^\varepsilon - u)$ converges in distribution to $\widehat{y}$. The only condition imposed on our limit object $\left\{ \widehat{\Theta}_k \right\}_{|k| \le q}$ is that it belongs to $C(I; \mathcal{H}_1)$. Hence, we must define what we mean by stochastic integration with respect to $\widehat{\Theta}_k$. However, an important advantage of our general conditions and this *invariance principle transfer* approach is that the analysis includes non-semimartingale, non-Markov limit objects $\left\{ \widehat{\Theta}_k \right\}_{|k| \le q}$ like fractional Brownian motions which are typical for long range (in t) dependent coefficients. In the special case where semimartingale conditions prevail the limit can be represented as a classical stochastic integral (see Kouritzin [7]).

Suppose C_u^1 ($C_u^{1,v}$) denotes the space of continuous $\mathbb{C}^{N \times N}$-valued ($\mathbb{C}^N$-valued) functions f such that $\partial_{x_i} f(t, x)$ exists and is continuous on $I \times \mathbb{R}^d$ for $i = 1, ..., d$ and there exists a $c_f > 0$ such that

$$\|f(t, x)\| \le c_f \ \text{and} \ \|\partial_{x_i} f(t, x)\| \le c_f \ \forall \ x \in \mathbb{R}^d, t \in I, i = 1, ..., d. \tag{4.7}$$

Then, it is shown in Dawson and Kouritzin [1] that $T_{t-\tau} f(\tau) \partial_\tau u_k(\tau)$ and $\partial_\tau T_{t-\tau} f(\tau) u_k(\tau)$ are Riemann integrable and

$$\int_0^t T_{t-\tau} f(\tau) \partial_\tau u_k(\tau) d\tau, \int_0^t (\partial_\tau T_{t-\tau}) f(\tau) u_k(\tau) d\tau \in C(I; H_2). \tag{4.8}$$

Hence, for $f \in C_u^1$ with $f(0) \equiv 0$ and $|k| \leq q$, we can define the integral of (T, u_k) with respect to f to be the $C(I; H_2)$ object defined for all $t \in I$ by the integration by parts formula

$$\int_0^t T_{t-\tau} f(d\tau) \, u_k(\tau) \doteq f(t) u_k(t) - \int_0^t \partial_\tau T_{t-\tau} f(\tau) u_k(\tau) d\tau - \int_0^t T_{t-\tau} \, f(\tau) \partial_\tau u_k(\tau) d\tau. \qquad (4.9)$$

However, it will often be the case that $\widehat{\Theta}_k \in C(I; H_1) \cap (C_u^1)^{\sim}$ (complement) and due to singularity

$$\int_0^t (\partial_\tau T_{t-\tau}) \, f(\tau) \, u_k(\tau) d\tau \, (x) = \int_0^t \int_{I\!R^d} (\partial_\tau \Gamma(x, t - \tau, \xi)) \, f(\tau, \xi) \, u_k(\tau, \xi) \, d\xi \, d\tau \qquad (4.10)$$

only exists as an iterated integral when $f \in C_u^1$. Consequently, we must extend our integral to $C(I; H_1)$ by continuity and completeness. One finds again from [1] that C_u^1 is dense in $C(I; H_1)$ and $f \to \int_0^t T_{t-\tau} \, f(\tau) \, \partial_\tau u_k(\tau) \, d\tau$ and $f \to \int_0^t \partial_\tau T_{t-\tau} \, f(\tau) \, u_k(\tau) \, d\tau$ are continuous mappings from $(C_u^1, |\cdot|_{C(I;H_1)})$ to $C(I; H_2)$ for each $|k| \leq q$. Hence, the following definitions make sense.

Definition 6. *Suppose now that $0 \leq |k| \leq q$, $f \in C(I; H_1)$ and $\{f_n\}_{n=1}^\infty \subset C_u^1$ is such that $f_n \to f$ in $C(I; H_1)$ and $f(0) = f_n(0) \equiv 0$. Then, we define the integral of (T, u_k) with respect to f to be the $C(I; H_2)$ object defined by*

$$\left\{ \int_0^t T_{t-\tau} f(d\tau) u_k(\tau), t \in I \right\} \doteq \lim_{n \to \infty} \left\{ \int_0^t T_{t-\tau} f_n(d\tau) u_k(\tau), t \in I \right\}. \qquad (4.11)$$

Finally, noting that our integral is a continuous mapping, we are in a position to define our stochastic integrals.

Definition 7. *When $|k| \leq q$ and $\left\{ \widehat{\theta}(t), t \in I \right\}$ is a $C(I; H_1)-$valued random variable on some probability space $(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{P})$, we define the stochastic integral of (T, u_k) with respect to $\widehat{\theta}$ to be the $C(I; H_2)$-valued random variable defined for each $\widehat{\omega} \in \widehat{\Omega}$ as in the above definition and denote this integral by*

$$\left\{ \int_0^t T_{t-\tau} \, d\widehat{\theta}(\tau) \, u_k(\tau), t \in I \right\}. \qquad (4.12)$$

We are now at the appropriate place to discuss our choice of H_1. First, since the coefficients A_k, A_k^0 and the solutions u^ε, u themselves exist as functions rather than just distributions it is desirable to treat them as such and avoid resorting to necessarily large spaces of generalized functions just to handle the limit objects $\widehat{\Theta}$ and $\widehat{y}$. Being able to show $\varepsilon^{h-1}(u^\varepsilon - u) \overset{\varepsilon \to 0}{\Rightarrow} \widehat{y}$ in a small space is better than in a large space. Secondly, no additional assumptions are required on the coefficients to ensure that they live in H_1. Thirdly, $|\cdot|_{C(I;H_1)}$ is exactly the norm required to make the continuity and density arguments work in the above definition of our stochastic integral and H_1 is the corresponding separable Hilbert space. Finally, H_1 fits beautifully into the methods used in Dawson and Kouritzin [1].

It is established in Kouritzin [7] that our stochastic integral from Definition 7 is equivalent to standard definitions when $\left\{ \widehat{\Theta}_t, t \in I \right\}$ is a semimartingale. Moreover, due to the way that everything has been constructed it would be natural to define $\widehat{y}$ as the mild solution to the linear stochastic partial differential equation

$$\partial_t \widehat{y}(t, x) = \sum_{|k| \leq 2p} A_k^0(x) \, \partial_x^k \widehat{y}(t, x) dt + \sum_{|m| \leq 2p} \widehat{\Theta}_m(dt, x) \partial_x^m u(t, x). \qquad (4.13)$$

REFERENCES

1. DAWSON D. A. & KOURITZIN M. A., Invariance principles for parabolic equations with random coefficients, *Institute for Mathematics and Its Applications Preprint Series* **# 1433** September (1996).
2. FRIEDMAN A., *Partial Differential Equations of Parabolic Type*, Prentice Hall, Englewood Cliffs, N.J. (1964).
3. HOCHBERG K. J. & ORSINGHER E., Composition of stochastic processes governed by higher-order parabolic and hyperbolic equations, *J. Theoret. Probab.* **9**, 511 − 532 (1996).
4. KHAS'MINSKII R. Z., A limit theorem for the solutions of differential equations with random right-hand sides, *Theory Probab. Appl.* **11**, 390−406 (1966).
5. KOURITZIN M. A., Averaging for fundamental solutions of parabolic equations, Revised for probable publication in *J. Differential Equations*.
6. KOURITZIN M. A., Approximations for singularly perturbed parabolic equations of arbitrary order, *Submitted*.
7. KOURITZIN M. A., On arbitrary order parabolic stochastic partial differential equations with general Hilbert-space-valued noise, *Under preparation*.
8. KUSHNER H. J. & HUANG H., Limits for parabolic partial differential equations with wide band stochastic coefficients and an application to filtering theory, *Stochastics* **23**, 115−148 (1985).