DESY 98-103
August 1998

ISSN 0418-9833

PARTICLE MOTION IN A RAPIDLY VARYING FIELD

Robert Cogburn and James A. Ellison

Department of Mathematics and Statistics
The University of New Mexico
Albuquerque, New Mexico 87131

Abstract

Averaging techniques are applied to particle motion in rapidly varying deterministic and random fields. In the deterministic case, we briefly review first and third order averaging and then present a new extension to longer time intervals on which the energy is approximately conserved. In the stochastic case, we review Khas'minskii's first and second order averaging and then discuss a new extension to longer time intervals on which the energy evolves approximately as a Markov diffusion process.

Particle motion in a rapidly varying field (RVF) appears to be of general and widespread interest. The general situation is governed by the initial value problem

$$\begin{aligned} \dot{x} &= \epsilon f(x,t,\omega) := \epsilon \bar{f}(x) + \epsilon p(x,t) + \epsilon q(x,t,\omega), \\ x(0,\epsilon,\omega) &= \zeta \end{aligned} \tag{1}$$

where $\epsilon > 0$ is a small parameter, p is deterministic with zero t-average and q is random with expected value zero ($\omega\epsilon\Omega$ where Ω is an underlying probability space and $Eq = 0$). Thus p and q can be considered perturbations in (1) and the goal is to find approximate solutions or approximate conservation laws based on knowledge of the unperturbed problem. RVF problems do not typically present themselves in the form of (1), but can often be put in this form by suitable scalings and transformations. In this letter, we focus on the special case where

$$\begin{aligned} x &= (x_1, x_2)^T, \ \bar{f}(x) = (x_2, -U'(x_1))^T, \\ p(x,t) &= (0, p_2(x_1,t))^T, \ q(x,t,\omega) = (0, q_2(x_1,t,\omega))^T. \end{aligned} \tag{2}$$

If we let $u = x_1$ and $t = \tau/\epsilon$ then (1-2) become $\frac{d^2u}{d\tau^2} + U'(u) = p_2(u,\tau/\epsilon) + q_2(u,\tau/\epsilon,\omega)$ which is a perturbed nonlinear oscillator. The term RVF is now clear because on $0(1)\tau$-intervals p_2 and q_2 are rapidly varying for ϵ small.

The importance of this problem was recognized early by Kapitza [1], and his derivation of an approximate solution to the deterministic, $q = 0$, version of (1-2) is presented in Landau and Lifschitz [2]. This derivation does not naturally lead to error estimates, although as we discuss the method of averaging [3] can be applied and gives such estimates. Khas'minskii [4] was the first to present rigorous results on the full stochastic case of (1). Our initial

motivation for this work came from the desire to understand the influence of lattice vibrations on the channeling motion of energetic charged particles in crystals [5, 6]. The equations of motion without lattice vibrations can be written (see [6], Eq. (2.11)), as (1) with $q = 0$ and $\bar{f}(x) + p(x,z) = (-\frac{1}{2}W_{x_3}(x_3,x_4,z), -\frac{1}{2}W_{x_4}(x_3,x_4,z), x_1, x_2)^T$ where $P_\perp := (x_1,x_2)$ and $Q_\perp := (x_3,x_4)$ are normalized momenta and coordinates perpendicular to the atomic strings lying along the channeling z-axis. Here, W is the crystal potential (usually taken as a sum, over the perfect crystal lattice sites, of screened coulomb potentials), which is periodic in z, and ϵ is a small parameter roughly equal to the ratio of transverse momentum to the longitudinal momentum. Thus, for ϵ small the channeled particle sees a RVF. If we allow the lattice atoms to be randomly displaced from their sites, we obtain (1) with $q \neq 0$ and if we assume planar motion ($x_2 = x_4 = 0$) we can obtain (1-2). A detailed analysis of this problem is in preparation [7]. A second example of recent interest involves the theory of rf noise in large synchrotron accelerators [8] such as the cancelled SSC in Texas, the Tevatron at Fermilab, the SPS at CERN and HERA at DESY. The equation of motion which approximately governs the "stationary bucket" longitudinal beam dynamics with rf phase noise is $\frac{d^2u}{d\tau^2} + \Omega^2 \sin(u + \psi(\tau/T_0)) = 0$, where τ is time, T_0 is beam revolution time, $\phi = u + \psi$ is the phase of the rf as seen by the particle, ψ is the noise and Ω is the small amplitude "synchrotron frequency." If we let $t = \tau/T_0, x_1 = u, \epsilon x_2 = dx_1/dt$ and $\epsilon = \Omega T_0$, then the equation is of the form (1-2). For SSC parameters $\epsilon \cong 0.008$ and thus on the time scale of a synchrotron oscillation the particle sees a RVF. Details are discussed in [9].

Regular perturbation methods when applied to (1-2) give rise to secular terms and thus give approximate solutions valid only on $O(1)$ t-intervals. However, it is in a standard form for the method of averaging, and in this letter we summarize known averaging results which give approximate solutions on $O(1/\epsilon)$ t-intervals and discuss our new results for adiabatic invariants on $O(1/\epsilon^2)$ t-intervals. In the latter case, we give some indication of how the proofs work. In what follows, p is of zero mean and 1-periodic in t and q is a stochastic process such that $Eq = 0$ and q at t and $t + s$ become independent for large s (mixing condition). For convenience, we take U to be a symmetric bowl type potential with zero minimum such that all solutions of the unperturbed problem ($p = q = 0$) are periodic; more general potentials can be treated. Precise regularity conditions can be found in Ref. [10].

The method of averaging is a standard tool for the analysis of deterministic differential equations, [3,6,11], such as (1-2) with $q = 0$, which we now study. The averaged problem for (1-2) is

$$\dot{u} = \epsilon \bar{f}(u), \quad u(0) = \zeta, \tag{3}$$

and if we let $\varphi(\tau,\zeta)$ denote the general solution of (3) with $\epsilon = 1$ and fix $T > 0$ then the first order averaging theorem gives

$$x(t,\epsilon) = \varphi(\epsilon t, \zeta) + 0(\epsilon) \tag{4}$$

uniformly for $t \in [0, T/\epsilon]$. The averaged IVP at third order is

$$\begin{aligned} \dot{v}_1 &= \epsilon v_2, & v_1(0,\epsilon) &= \xi - \epsilon^2 \mathcal{I}^2 p_2(\xi,0) \\ \dot{v}_2 &= -\epsilon U'_{eff}(v_1,\epsilon), & v_2(0,\epsilon) &= \eta - \epsilon \mathcal{I} p_2(\xi,0) + \epsilon^2 \eta \mathcal{I}^2 D_1 p_2(\xi,0) \end{aligned} \tag{5}$$

where $U_{eff}(x_1,\epsilon) = U(x_1) + \frac{1}{2}\epsilon^2 \int_0^1 [(\mathcal{I}p_2)(x_1,t)]^2 dt$ and $\zeta = (\xi,\eta)^T$. Here, $\mathcal{I}g(t)$ denotes the zero-mean integral of a zero-mean periodic function and D_1 denotes the derivative with respect to the first argument, be it vector or scalar. The averaging theorem at third order gives

$$\begin{pmatrix} x_1(t,\epsilon) \\ x_2(t,\epsilon) \end{pmatrix} = \begin{pmatrix} v_1(t,\epsilon) + \epsilon^2(\mathcal{I}^2 p_2)(v_1(t,\epsilon),t) \\ v_2(t,\epsilon) + \epsilon(\mathcal{I}p_2)(v_1(t,\epsilon),t) - \epsilon^2 v_2(t,\epsilon)(\mathcal{I}^2 D_1 p_2)(v_1(t,\epsilon),t) \end{pmatrix} + O(\epsilon^3) \quad (6)$$

uniformly for $t\ \epsilon[0, T/\epsilon]$. More details can be found in Refs. [6] and [11] and [12] contains the quantum case.

The differential equation in (5) is equivalent to the effective problem in Ref. [2] although the approximation to x_2 is not discussed there. In fact, it is not clear that the approach of Ref. [2] can give the approximation to x_1 given in (6) because it is not clear that an accurate enough initial condition for v_2 can be obtained in that context. Furthermore, unless the initial condition can be determined to the order in (5), the order of the error in (6) will become $O(\epsilon^2)$ and it becomes questionable whether the use of U_{eff} in Eq. (5) is warranted.

Typically, it is difficult to extend averaging results to longer than $O(1/\epsilon)$ t-intervals, unless $\bar{f} = 0$. For example, in the general setting of (1-2), it is probably not true that (4) and (6) are valid on longer t-intervals or even that the approximations can be improved to make it so. Thus, it makes sense to look at invariants of the unperturbed problem, and here we examine the evolution of the energy

$$e(x) = \frac{1}{2}x_2^2 + U(x_1). \quad (7)$$

We will show that $y_\epsilon(t) := e(x(t,\epsilon))$ is approximately constant on $O(1/\epsilon^2)$ t-intervals. More precisely, we sketch the proof of the following theorem.

Theorem 1 *Fix $T > 0$ and $e_1 > e(\zeta)$ and let $\mathcal{D} = \{x \mid e(x) < e_1\}$. Then there exist positive constants ϵ^* and C such that for all $0 \le t \le T/\epsilon^2, 0 < \epsilon \le \epsilon^*$, the IVP (1-2) has a unique solution in $\mathcal{D}$ and $\mid y_\epsilon(t) - e(\zeta) \mid\le C\epsilon$.*

Remarks:

(1) In the proof we need the following quantities:

$$\hat{p}(x_1) := \int_0^1 t p_2(x_1,t)dt, \quad (8)$$

$$\tilde{\mu}_1(x) := x_2^2 D\hat{p}(x_1) - U'(x_1)\hat{p}(x_1), \quad (9)$$

$$\bar{\mu}_1(z) := \frac{1}{P(z)}\int_0^{P(z)} \tilde{\mu}_1(\varphi(\tau,\zeta))d\tau, \quad (10)$$

where $z = e(\zeta)$ and P is the period of φ, that is $\varphi(\tau + P(e(\zeta)),\zeta) = \varphi(\tau,\zeta)$. We will show that $\bar{\mu}_1(z) = 0$.

(2) In light of the theorem and because the flow is area preserving (the vector field $\bar{f}+p$ is divergence free), one might ask if there are in fact invariant curves for the period one Poincaré map thus keeping the energy confined for all t. We have not yet investigated this; however we do have an example [10] of the form of (1), which is not area preserving, in which our method shows the energy varies $O(1)$ on $O(1/\epsilon^2)$ t-intervals. Futhermore, the restrictions on ϵ here are much less restrictive than for a typical invariant curve theorem.

Proof: Let $[0,\beta(\epsilon))$ denote the maximal right interval of existence of $x(t,\epsilon)$ in $\mathcal{D}$. Note that all functions are bounded and globally x-Lipschitz on $\mathcal{D}$ and that $y_\epsilon(t) = e(x(t,\epsilon))$ gives

$$\dot{y}_\epsilon = \epsilon x_2 p_2(x_1,t) =: \epsilon g(x,t) \quad (11)$$

with g of zero t-mean. The proof consists of two lemmas.

Lemma 1: $y_\epsilon(t_2) - y_\epsilon(t_1) = \epsilon^2\int_{t_1}^{t_2}\bar{\mu}_1(x(s,\epsilon))ds + \epsilon O(\epsilon^2(t_2-t_1)+1)$ *uniformly for* $t_1,t_2 \in [0,\beta(\epsilon))$.

Proof: Let $n_1 = [t_1]$ and $n_2 = [t_2]$, where [ ] denotes the greatest integer less than, then

$$y_\epsilon(t_2) - y_\epsilon(t_1) = \epsilon\sum_{k=n_1}^{n_2-1}\int_k^{k+1} g(x(s,\epsilon),s)ds + O(\epsilon).$$

Now for $k \le s \le k+1$ and $x(k,\epsilon) = x_k, g(x(s,\epsilon),s) = g(x_k,s) + D_1 g(x_k,s)(x(s,\epsilon)-x_k) + O(\|x(s,\epsilon)-x_k\|^2)$. It follows from (1-2) that $\| x(s,\epsilon) - x_k \| = O(\epsilon)$, thus $x(s,\epsilon) - x_k = \epsilon\int_k^s f(x_k,\tau)d\tau + O(\epsilon^2)$ where $f = \bar{f} + p$. Hence $\int_k^{k+1} g(x(s,\epsilon),s)ds = 0 + \epsilon\tilde{\mu}_1(x_k) + O(\epsilon^2) = \epsilon\int_k^{k+1}\tilde{\mu}_1(x(s,\epsilon))ds + O(\epsilon^2)$. The lemma is immediate. □

Lemma 2 (Ergodic): $y_\epsilon(t) - y_\epsilon(0) = \epsilon O(\epsilon^2 t + 1)$ *uniformly for* $t \in [0,\beta(\epsilon))$.

Proof: Let $t_{k+1} = t_k + \frac{1}{\epsilon}P_k, P_k = P(y_\epsilon(t_k)), t_0 = 0, x(0) = \zeta$. Define m by $t_m \le t \le t_{m+1}$. Clearly $\int_0^t \tilde{\mu}_1(x(s,\epsilon))ds = \sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}}\tilde{\mu}_1(x(s,\epsilon))ds + O(1/\epsilon)$, and from the first order averaging theorem $x(s,\epsilon) = \varphi(\epsilon(s-t_k), x(t_k,\epsilon)) + O(\epsilon)$ uniformly for $t_k \le s \le t_{k+1}$ and $0 \le k \le m$. Thus

$$\begin{aligned}\int_0^t \tilde{\mu}_1(x(s,\epsilon))ds &= \sum_{k=0}^{m-1} P_k\tfrac{1}{P_k}\int_{t_k}^{t_{k+1}}\tilde{\mu}_1(\varphi(\epsilon(s-t_k),x(t_k,\epsilon)))ds + O(\epsilon t) + O(1/\epsilon) \quad (12)\\ &= \sum_{k=0}^{m-1}\tfrac{1}{\epsilon}P_k\bar{\mu}_1(y_\epsilon(t_k)) + O(\epsilon t) + O(1/\epsilon).\end{aligned}$$

Let $u(\tau) = \varphi(\tau,\zeta)$, then $\dot{u}_1 = u_2$ and $\dot{u}_2 = -U'(u_1)$. Combining this with (9) gives $\tilde{\mu}_1(u(\tau)) = u_2\dot{u}_1 D\hat{p}(u_1) + \dot{u}_2\hat{p}(u_1) = \frac{d}{d\tau}u_2\hat{p}(u_1)$ and integrating over a period gives $\bar{\mu}_1 = 0$. Lemma 2 follows from Lemma 1. □

Lemma 2 gives $|y_\epsilon(t) - e(\zeta)| < C\epsilon$ for $t \in J(\epsilon) := [0,\beta(\epsilon)) \cap [0,T/\epsilon^2]$ where $C = C(T) = M(T+1)$ and M is the order constant in Lemma 2. The theorem is proven if we can show $\beta(\epsilon) > T/\epsilon^2$. Now $|y_\epsilon(t)| \le C\epsilon + e(\zeta)$. Choose ϵ^* such that $C\epsilon^* + e(\zeta) := e^* < e_1$ and assume $\epsilon < \epsilon^*$ and $\beta(\epsilon) \le T/\epsilon^2$. Then $x(t,\epsilon)$ stays away from the boundary of $\mathcal{D}$ as $t \uparrow \beta(\epsilon)$ which contradicts $[0,\beta(\epsilon))$ being the maximal forward interval of existence of $x(t,\epsilon)$ in $\mathcal{D}$. Therefore $\beta(\epsilon) > T/\epsilon^2$, and this completes the proof of the theorem. □

Remark: We call Lemma 2 an ergodic result since it averages $\tilde{\mu}_1$ over orbits in phase space. To extend these results to systems with more degrees of freedom [10], we must introduce an ergodic hypothesis that $\bar{\mu}_1$ and related functions, when averaged along orbits of the unperturbed system, yield functions depending only on the invariant y. However, for a certain class of deterministic systems we can remove the ergodicity requirement [10].

References [4] and [13] discuss extensions of averaging to stochastic systems and, when these ideas are applied to the IVP (1-2) with q satisfying a mixing condition, we obtain the following results on $O(1/\epsilon)$ t-intervals:

I. $x(\tau/\epsilon,\epsilon,\omega) \to \varphi(\tau,\zeta)$ in probability as $\epsilon \to 0$ uniformly for $\tau \in [0,\tau_0]$.

II. If we define $Y(t,\epsilon,\omega)$ by $x(t,\epsilon,\omega) = \varphi(\epsilon t,\zeta) + \sqrt{\epsilon}Y(t,\epsilon,\omega)$ then $Y(\tau/\epsilon,\epsilon,\omega)$ converges weakly to a Gauss-Markov process $Y_0(\tau,\omega)$ for $0 \le \tau \le \tau_0$, where $Y_0(\tau,\omega)$ is defined by the Itô stochastic differential equation (SDE)

$$dY_0 = \begin{pmatrix} 0 \\ k(\varphi_1(\tau,\zeta))^{\frac{1}{2}} \end{pmatrix} dW(\tau,\omega) + \begin{pmatrix} 0 & 1 \\ -U''(\varphi_1(\tau,\zeta)) & 0 \end{pmatrix} Y_0 d\tau, Y_0(0) = 0. \quad (13)$$

Here $\lim_{\ell\to\infty} \frac{1}{\ell}\int_t^{t+\ell}\int_t^{t+\ell} E(q_2(x_1,s)q_2(x_1,r))dsdr =: k(x_1)$ is assumed to exist uniformly for $x, t > 0$ and $W(\tau,\omega)$ is standard Brownian motion. This equation is easy to solve in terms of $\Phi(\tau) := D_2\varphi(\tau,\zeta)$, i.e.

$$Y_0(\tau) = \Phi(\tau)\int_0^\tau \Phi^{-1}(s) \begin{pmatrix} 0 \\ k(\varphi_1(s,\zeta))^{\frac{1}{2}} \end{pmatrix} dW. \quad (14)$$

Remarks:

(1) Results I and II require a global Lipschitz condition on $\mathbb{R}^2$ which is not satisfied by (1-2), however this condition can be removed as discussed in Ref. [10].

(2) Weak convergence is a strong result. By definition, a process $X_\epsilon(\tau)$ converges weakly to the process $X_0(\tau)$ as $\epsilon \to 0$ on $[0,\tau_0]$, if the finite-dimensional distributions of the process $X_\epsilon(\tau)$ converge to the finite-dimensional distributions of the process $X_0(\tau)$ and the distribution of $f(X_\epsilon(\tau))$ converges to the distribution of $f(X_0(\tau))$ for any continuous functional f on $C_{[0,\tau_0]}$ (See [14]).

Again, extensions to $O(1/\epsilon^2)$ t-intervals are difficult, although some extensions are possible under suitable restrictions (e.g. $\bar{f}(x) = 0$), see [15,16]. These ideas do not apply to (1-2). However, a modification of the ideas in [15] and an extension of the deterministic theorem just proven do apply to the energy process $y_\epsilon(t) = e(x(t,\epsilon,\omega))$. If as in [15] we assume a mixing condition, that is if we assume that $q_2(x_1,t)$ and $q_2(x_1,t+s)$ become independent sufficiently fast as s gets large, then we have proven that $y_\epsilon(\tau/\epsilon^2)$ converges weakly to a Markov diffusion process $z(\tau)$ as $\epsilon \to 0$ for $0 \le \tau \le T$. Here $z(\tau)$ is defined by the Itô SDE

$$dz = (\bar{\mu}_1(z) + \bar{\mu}_2(z))d\tau + (\bar{\sigma}(z))^{1/2}dW,\ z(0) = e(\zeta), \quad (15)$$

or equivalently by its transition probability density $\rho(z,\tau|z')$ and its initial density $\delta(z-e(\xi))$. The transition probability density is the unique solution of the Fokker-Planck equation

$$\frac{\partial\rho}{\partial\tau}(z,\tau|z_0) + \frac{\partial}{\partial z}(\bar{\mu}_1(z) + \bar{\mu}_2(z))\rho(z,\tau|z_0) = \frac{1}{2}\frac{\partial^2}{\partial z^2}\bar{\sigma}(z)\rho(z,\tau|z_0), \quad (16)$$
$$\rho(z,0|z_0) = \delta(z-z_0),$$

where $\bar{\mu}_1(z)$ is defined in the deterministic case and $\bar{\mu}_2(x) = \frac{1}{2}k(x_2)$, $\bar{\sigma}(x) = x_2^2k(x_1)$ and $\bar{\mu}_2(z)$ and $\bar{\sigma}(z)$ are defined as in Eq. (10). Here $\bar{\mu}_1$ is zero, however, this is not necessarily true in the general setting of Eq. (1).

In the stochastic case the proof is based on dividing the process into increments of length $L = L(\epsilon)$, where $L(\epsilon)\epsilon^{\frac{1}{2}} \to 0$ slower than any power of ϵ, then using a Taylor series expansion and standard estimates to obtain, for $0 \le t_1 < t_2 \le \beta(\epsilon)$

$$y_\epsilon(t_2) - y_\epsilon(t_1) = \sum_{k=[t_1/L]}^{[t_2/L]-1}\left(\epsilon\zeta_k + \epsilon^2\gamma_k\right) + \left(\epsilon^2(t_2-t_1) + \epsilon^{\frac{1}{2}}\right)o(1)$$

where, letting $G(x,t,\omega) = x_2q_2(x_1,t,\omega)$,

$$\xi_k = \int_{(k+1)L}^{(k+2)L} G(x(kL,\epsilon),t,\omega)dt$$

$$\gamma_k = \int_{(k+1)L}^{(k+2)L} D_1G(x(kL,\epsilon),t,\omega)\int_{kL}^t f(x(kL,\epsilon)s,\omega)dsdt.$$

Under the mixing condition, the sequence $\{(\zeta_k,\gamma_k), k = 0,1,\dots\}$ behaves almost like a sequence of independent random variables. The $\sum\epsilon\xi_k$ term generates the diffusion $\bar{\sigma}(x)$, while the $\sum\epsilon^2\gamma_k$ term generates the drift $\bar{\mu}_2(x)$, on $O(1/\epsilon^2)$ time scales. A number of estimates and stochastic results on weak convergence are required. Most of this is a transparent adaptation of the procedures used in [15], the primary innovation being the introduction of the ergodic averaging, required to obtain the diffusion and drift coefficients in (15) that depend only on $z = z(\tau)$, the energy of the process at time τ. We have also obtained a more precise representation of the effects of the deterministic perturbing term p by applying results from the theory of almost periodic functions.

In summary, we have reviewed deterministic and stochastic averaging results for (1-2) on $O(1/\epsilon)$ t-intervals. We have shown how these results can be extended to $O(1/\epsilon^2)$ intervals by looking at an invariant of the unperturbed problem rather than the process itself. In Ref. [10] we present a much more general theory of "stochastic adiabatic invariance" and a "4/3 law" for equations of the form of (l) with $x \in \mathbb{R}^n$.

This paper was finished in 1991, but never published. Theorem 1 has not been presented elsewhere and is presently being used and in addition we feel the general discussion of the RVF problem is still of interest. More recent references can be found in [17].

We thank J. Tapia for useful comments, Helmut Mais for his encouragement and Linda Cicarella and Susan Pinter for their preparation of this manuscript. JAE was supported by NSF Grant DMR-8704348.

References

[1] P. L. Kapitza, Zh. Eksp. Teor. Fiz. **21**(1951), p. 588.

[2] L. D. Landau and E. M. Lifshitz, "Mechanics," (Pergamon Press Ltd., London, England, second edition, 1969), p. 93.

[3] J. A. Murdock, "Perturbations," John Wiley & Sons, New York, 1991.

[4] R. Z. Khas'minskii, Theory Prob. Appl. **11**(1966), p. 211.

[5] D. S. Gemell, Rev. Mod. Phys. **46**(1974), p. 129.

[6] H. S. Dumas, J. A. Ellison and A. W. Sáenz, *Axial Channeling in Perfect Crystals, the Continuum Model and the Method of Averaging*, Annals of Phys. **209**, 97-123(1991).

[7] J. Tapia, Ph.D. dissertation, University of New Mexico (1991), and R. Cogburn, J. A. Ellison and J. Tapia, (to be submitted).

[8] G. Dôme, Proc. CERN Accelerator School, Antiprotons for Colliding Beam Facilities, CERN 84-15, 20 December 1984, p. 215.

[9] H.-J. Shih, J. A. Ellison, B. S. Newberger and R. Cogburn, *Longitudinal Beam Dynamics with RF Noise*, Part. Acc., **43**, 159-194(1994).

[10] R. Cogburn and J. A. Ellison, *A Stochastic Theory of Adiabatic Invariance*, Commun. Math. Phys. **149** 97-126 (1992). Commun. Math. Phys. **149** 97-126 (1992) and *A Four-thirds Law for Phase Randomization of Stochastically Perturbed Oscillators and Related Phenomena*, Commun. Math. Phys. **166**, 317-336 (1994).

[11] J. A. Ellison, A. W. Sáenz and H. S. Dumas, *Improved Nth Order Averaging Theory for Periodic Systems*, J. Diff. Eq., **84**, 383-403(1990).

[12] A. Ben Lemlih, "An Extension of the Method of Averaging to Partial Differential Equations," Ph.D. Thesis, University of New Mexico, 1986. See also A. Ben Lemlih and J. A. Ellison, Phys. Rev. Lett. **55**(1985), p. 1950.

[13] M. I. Freidlin and A. D. Wentzell, "Random Perturbations of Dynamical Systems," (Springer-Verlag, New York, 1984).

[14] P. Billingsley, "Convergence of Probability Measures," (Wiley, New York, 1968), and H. J. Kushner, "Approximation and Weak Convergence Methods for Random Processes," (MIT Press, Cambridge, 1984).

[15] R. Z. Khas'minskii, Theory Prob. Appl. **11**(1966), p. 390, and A. N. Borodin, Theory Prob. Appl., **22**(1977), p. 482.

[16] R. Cogburn and R. Hersh, Indiana Univ. Math. J., **22**(1973), p. 211 and G. C. Papanicolaou and S. R. S. Varadhan, Comm. Pure Appl. Math., **26**(1973), p. 497.

[17] J. A. Ellison, *Accelerators and Probability: The Special Effect of Noise in Beam Dynamics*, to be published in the proceedings of the workshop *Nonlinear and Stochastic Beam Dynamics in Accelerators - A Challenge to Theoretical and Computational Physics*, Lüneburg, October 1997.