The Annals of Probability
2002, Vol. 30, No. 1, 101–134

LYAPUNOV EXPONENTS FOR SMALL RANDOM PERTURBATIONS OF HAMILTONIAN SYSTEMS

BY PETER H. BAXENDALE AND LEVON GOUKASIAN

University of Southern California

Consider the stochastic nonlinear oscillator equation

$$\ddot{x} = -x - x^3 + \varepsilon^2 \beta \dot{x} + \varepsilon \sigma x \dot{W}_t$$

with $\beta < 0$ and $\sigma \neq 0$. If $4\beta + \sigma^2 > 0$ then for small enough $\varepsilon > 0$ the system $(x, \dot{x})$ is positive recurrent in $\mathbf{R}^2 \setminus \{(0, 0\}$. Now let $\overline{\lambda}(\varepsilon)$ denote the top Lyapunov exponent for the linearization of this equation along trajectories. The main result asserts that

$$\overline{\lambda}(\varepsilon) = \varepsilon^{2/3} \overline{\lambda} + O(\varepsilon^{4/3}) \qquad \text{as } \varepsilon \to 0$$

with $\overline{\lambda} > 0$. This result depends crucially on the fact that the system above is a small perturbation of a Hamiltonian system. The method of proof can be applied to a more general class of small perturbations of two-dimensional Hamiltonian systems. The techniques used include (i) an extension of results of Pinsky and Wihstutz for perturbations of nilpotent linear systems, and (ii) a stochastic averaging argument involving motions on three different time scales.

1. Introduction. The stochastic version of the Duffing–van der Pol equation

$$\ddot{x} = \alpha x + \beta \dot{x} - a x^3 - b x^2 \dot{x} + \sigma x \dot{W}_t$$

(with $a \geq 0$, $b \geq 0$) has quickly become a standard test case for the theory of stochastic bifurcations. See Arnold [2], Arnold, Sri Namachchivaya and Schenk-Hoppé [4], Keller and Ochs [15], Liang and Sri Namachchivaya [18] and Schenk-Hoppé [22]. Here we consider only the case of a multiplicative perturbation by white noise. In phase space we get the equations

$$\begin{aligned} dx_1 &= x_2\, dt, \\ dx_2 &= (\alpha x_1 + \beta x_2 - a x_1^3 - b x_1^2 x_2)\, dt + \sigma x_1\, dW_t, \end{aligned} \tag{1}$$

where W_t denotes a standard one dimensional Brownian motion process. (The Wong–Zakai correction term is 0 in this system, so the Itô and Stratonovich versions have the same coefficients.) Linearizing this equation at $x = 0$ we get

$$du_t = \begin{bmatrix} 0 & 1 \\ \alpha & \beta \end{bmatrix} u_t\, dt + \begin{bmatrix} 0 & 0 \\ \sigma & 0 \end{bmatrix} u_t\, dW_t. \tag{2}$$

Received November 2000; revised April 2001.
AMS 2000 subject classifications. Primary, 37H10, 60H10; secondary, 37H15, 37H20, 60J60.
Key words and phrases. Stochastic oscillator, Lyapunov exponent, Hamiltonian, stochastic averaging, nilpotent stochastic differential equation.

The top Lyapunov exponent for the linearized equation is

$$\lambda(\alpha,\beta,\sigma) \equiv \lim_{t\to\infty} \frac{1}{t} \log \|u_t\|$$

where the limit exists almost surely and does not depend on $u_0 \neq 0$ if $\sigma \neq 0$. Clearly the sign of $\lambda(\alpha,\beta,\sigma)$ determines the stability or instability of the linearized process u_t. The sign of $\lambda(\alpha,\beta,\sigma)$ also determines the stability or instability for the original process x_t, in the following sense.

THEOREM 1. *Assume either $a > 0$ and $b > 0$; or $a = 0$, $b > 0$ and $\alpha < 0$; or $a > 0$, $b = 0$ and $\beta < 0$.*

(i) *If $\lambda(\alpha,\beta,\sigma) < 0$ then 0 is almost surely stable, in the sense that $P^x(x_t \to 0$ as $t \to \infty) = 1$ for all $x \in \mathbf{R}^2$.*

(ii) *If $\lambda(\alpha,\beta,\sigma) > 0$ then 0 is almost surely unstable, and the process $\{x_t : t \geq 0\}$ is positive recurrent in $\mathbf{R}^2 \setminus \{0\}$ with stationary probability μ, say. Moreover, μ has the property $\mu(B(0,r))/r^2 \to \infty$ as $r \to 0$ if $\beta < 0$ and $\mu(B(0,r))/r^2 \to 0$ as $r \to 0$ if $\beta > 0$.*

This is an application of results of Baxendale [7] and Baxendale and Stroock [10], see also [2], Section 9.5.1; more details are given in the remark following Lemma 6 in Subsection 3.1. The cases when $a = 0$ and $b = 0$ correspond to the stochastic versions of the van der Pol oscillator and the Duffing equation, respectively. The theorem implies that for the one point motion $\{x_t : t \geq 0\}$ we see a change in the dynamical behavior, or a D-bifurcation, when $\lambda(\alpha,\beta,\sigma)$ changes sign, and a qualitative change in the stationary probability measure μ, or P-bifurcation, when β changes sign.

If the Lyapunov exponent $\lambda(\alpha,\beta,\sigma)$ is negative, it is clear that the trajectories $\{x_t : t \geq 0\}$ and $\{y_t : t \geq 0\}$ obtained by starting the SDE (1) at distinct points x and y will simultaneously converge to 0, and so in particular the distance apart $\|y_t - x_t\|$ will converge to 0. However if $\lambda(\alpha,\beta,\sigma) > 0$ it is not readily apparent what will happen to $\|y_t - x_t\|$ as $t \to \infty$. We will refer to this as the problem of stability along trajectories. In this paper we will address a closely related property, that of linearized stability along trajectories. (This property is also the main topic in the paper [8].) Linearizing the SDE (1) along the trajectory x_t we get the following equation

$$dv_t = \begin{bmatrix} 0 & 1 \\ \alpha - 3ax_1^2 - 2bx_1x_2 & \beta - bx_1^2 \end{bmatrix} v_t\,dt + \begin{bmatrix} 0 & 0 \\ \sigma & 0 \end{bmatrix} v_t\,dW_t.$$

Write

$$v_t = \|v_t\| \begin{bmatrix} \cos\theta_t \\ \sin\theta_t \end{bmatrix},$$

then

$$d\theta_t = \big[(\beta - bx_1^2)\sin\theta\cos\theta + (\alpha - 3ax_1^2 - 2bx_1x_2)\cos^2\theta - \sin^2\theta - \sigma^2\sin\theta\cos^3\theta\big]\,dt + \sigma\cos^2\theta\,dW_t$$

and, assuming $\lambda(\alpha,\beta,\sigma) > 0$ so that μ exists, we get the Furstenberg–Khas'-minskii–Carverhill formula for the Lyapunov exponent for the v_t process

$$\overline{\lambda}(\alpha,\beta,\sigma,a,b) \equiv \lim_{t\to\infty}\frac{1}{t}\log\|v_t\| = \int Q(x,\theta)\,dP(x,\theta) \tag{3}$$

where

$$Q(x,\theta) = (\beta - bx_1^2)\sin^2\theta + (1+\alpha-3ax_1^2-2bx_1x_2)\sin\theta\cos\theta + \frac{\sigma^2}{2}\cos 2\theta\cos^2\theta$$

and the measure P on $(\mathbf{R}^2\setminus\{0\})\times S^1$ is the stationary probability measure for the (x_t,θ_t) process. However the integral in (3) is hard to evaluate; we cannot determine the sign of $\overline{\lambda}(\alpha,\beta,\sigma,a,b)$ and so we cannot determine whether there is linearized stability or instability along the trajectories.

In this paper we will consider the special case of equation (1) where $\alpha = -\omega^2$, $\omega > 0$, and $a > 0$ are fixed, and the other parameters are small. Using the transformation $\overline{x}_t = (\sqrt{a}/\omega)x_{t/\omega}$ we see that

$$\lambda(-\omega^2,\beta,\sigma) = \omega\lambda(-1,\beta/\omega,\sigma/\omega^{3/2}),$$
$$\overline{\lambda}(-\omega^2,\beta,\sigma,a,b) = \omega\overline{\lambda}(-1,\beta/\omega,\sigma/\omega^{3/2},1,\omega b/a).$$

Therefore we can and will assume henceforth that $\omega = a = 1$. Replacing the dissipation terms $\beta\dot{x}$ and $-bx^2\dot{x}$ by $\varepsilon^2\beta\dot{x}$ and $-\varepsilon^2 bx^2\dot{x}$, respectively, and replacing the noise intensity σ^2x^2 by $\varepsilon^2\sigma^2x^2$, we get the system

$$\begin{aligned} dx_1 &= x_2\,dt,\\ dx_2 &= (-x_1 - x_1^3 + \varepsilon^2\beta x_2 - \varepsilon^2 bx_1^2x_2)\,dt + \varepsilon\sigma x_1\,dW_t. \end{aligned} \tag{4}$$

Here $b \ge 0$. If $b = 0$ we need to assume also that $\beta < 0$. The equation with $b = 0$ appears in the study of stochastically forced vibrations of a thin beam; see Wedig [24]. In this setting $-\beta$ represents friction, so the restriction $\beta < 0$ is physically meaningful. We note, in justification of our title, that the equations (4) with $\varepsilon = 0$ give the flow of the Hamiltonian system with $H(x_1,x_2) = x_1^2/2 + x_1^4/4 + x_2^2/2$.

Imkeller and Lederer [14] give an exact formula for $\lambda(-1,\beta,\sigma)$. Using this result, or the earlier asymptotic result of Auslender and Mil'shtein [5], we get

$$\lambda(-1,\varepsilon^2\beta,\varepsilon\sigma) = \left(\frac{\beta}{2} + \frac{\sigma^2}{8}\right)\varepsilon^2 + O(\varepsilon^4). \tag{5}$$

Thus, if $\beta/2 + \sigma^2/8 < 0$, then for all sufficiently small $\varepsilon > 0$ the trajectories x_t converge to 0 almost surely. However if $\beta/2 + \sigma^2/8 > 0$, then for all sufficiently small $\varepsilon > 0$ the process x_t is positive recurrent on $\mathbf{R}^2 \setminus \{0\}$ with stationary probability μ_ε, say, and we need to consider the Lyapunov exponent

$$\overline{\lambda}(\varepsilon) \equiv \overline{\lambda}(-1, \varepsilon^2\beta, \varepsilon\sigma, 1, \varepsilon^2 b)$$

for the linearization of (4) along trajectories. Our main result is the following.

THEOREM 2. *Consider* (4) *with* $\sigma > 0$ *and* $b \geq 0$ *and* $\beta/2 + \sigma^2/8 > 0$. *If* $b = 0$ *suppose also that* $\beta < 0$. *Then*

$$\overline{\lambda}(\varepsilon) = \varepsilon^{2/3}\overline{\lambda} + O(\varepsilon^{4/3}) \qquad \text{as } \varepsilon \to 0 \tag{6}$$

where

$$\overline{\lambda} = \gamma_0 \int_{(0,\infty)} G(h)\, d\rho(h) > 0. \tag{7}$$

Here γ_0 *is a constant, given in* (33) *with a numerical value approximately* 0.29; *and* $G(h)$ *is a positive function, given in* (45); *and* ρ *is the invariant probability measure on* $(0, \infty)$ *for the generator given explicitly in* (24), (46), (47).

COROLLARY 1. *Fix* $\omega \neq 0$, $\beta \in \mathbf{R}$, $\sigma \neq 0$, $a > 0$ *and* $b \geq 0$. *If* $4\omega^2\beta + \sigma^2 > 0$ *and either* $b > 0$ *or* $\beta < 0$ *then there is* $\varepsilon_0 > 0$ *such that* $\lambda(-\omega^2, \varepsilon^2\beta, \varepsilon\sigma) > 0$ *and* $\overline{\lambda}(-\omega^2, \varepsilon^2\beta, \varepsilon\sigma, a, \varepsilon^2 b) > 0$ *whenever* $0 < \varepsilon \leq \varepsilon_0$.

COROLLARY 2. *For every* $\alpha < 0$ *and* $a > 0$ *there exist* $\beta \in \mathbf{R}$, $\sigma \neq 0$ *and* $b \geq 0$ *such that* $\lambda(\alpha, \beta, \sigma) > 0$ *and* $\overline{\lambda}(\alpha, \beta, \sigma, a, b) > 0$.

For a stochastic flow $\{\xi_t : t \geq 0\}$ on a compact manifold M, the fact that the top Lyapunov exponent is positive, together with suitable non-degeneracy conditions, implies that the associated forward Markov invariant measure (or statistical equilibrium measure) almost surely contains no atoms. (For a discussion of forward Markov invariant measures, see [2]; for the result see [6], Remark 4.12.) The major step in this argument is the fact that linearized instability implies instability for the two point motion $(\xi_t(x), \xi_t(y))$ on $M \times M$. If we assume that a similar result is valid for the strictly forward complete stochastic flow generated by (1) on the non-compact space $\mathbf{R}^2 \setminus \{0\}$, then Corollary 2 implies that for certain parameter values in the stochastic Duffing–van der Pol equation the forward Markov invariant measure contains no atoms.

Our technique of proof allows us also to determine the behavior of the stationary probability μ_ε as $\varepsilon \to 0$. Let $\tilde{\rho}(h)$ denote the density of ρ with respect to Lebesgue measure on $(0, \infty)$, and let $T(h)$ denote the period of the Hamiltonian flow (obtained by putting $\varepsilon = 0$ in (4)) around the orbit $H(x_1, x_2) = h$. Define μ_0 to be the measure on $\mathbf{R}^2 \setminus \{0\}$ with density $\tilde{\rho}(H(x_1, x_2))/T(H(x_1, x_2))$ with respect to two-dimensional Lebesgue measure.

THEOREM 3. *Let $g(x_1, x_2)$ be a smooth function with compact support in $\mathbf{R}^2 \setminus \{0\}$. Then*

$$\left| \int g \, d\mu_\varepsilon - \int g \, d\mu_0 \right| = O(\varepsilon^2) \qquad \text{as } \varepsilon \to 0.$$

In particular $\mu_\varepsilon \to \mu_0$ as $\varepsilon \to 0$ weakly as probability measures on $\mathbf{R}^2$.

The formula for the density of μ_0 and the weak convergence of μ_ε to μ_0 are well known as a general principle for stochastic averaging of Hamiltonian systems, see Khas'minskii [16]. Here we observe that our techniques provide a rate for the convergence. The proof of this result appears in Subsection 3.6. It is clear from the proof that the convergence of order ε^2 will remain true for a wider class of functions g with appropriate behavior near 0 and ∞, but we not pursue that matter here.

The appearance in Theorem 2 of the scaling factor $\varepsilon^{2/3}$, rather than ε^2 as in the formula (5) for the Lyapunov exponent at 0, is perhaps surprising. Its presence is essentially based on the following sequence of observations.

(i) Equation (4) is a small perturbation of a Hamiltonian system.

(ii) The linearization of a two-dimensional Hamiltonian system, when written with respect to a suitable moving frame, is a nilpotent linear system.

(iii) A remarkable paper of Pinsky and Wihstutz [20] shows how to handle small random perturbations of a nilpotent system. In its simplest form, the method of [20] shows that the linear SDE

$$du_t = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} u_t \, dt + \begin{bmatrix} 0 & 0 \\ \varepsilon & 0 \end{bmatrix} u_t \, dW_t$$

has top Lyapunov exponent $\varepsilon^{2/3}\gamma_0$ where γ_0 is the constant appearing in Theorem 2.

The paper by Pinsky and Wihstutz [20] deals with constant coefficient linear SDEs, whereas the coefficients in our system will depend on the current position of the underlying trajectory x_t. Thus in addition to the Pinsky–Wihstutz transformation and the consequent averaging over an angle $\theta_t \in S^1$, we will have to do some averaging for the process x_t. A simplified version of this technique appears in [9]. In the present situation, it turns out that we have to deal with motions on three distinct times scales. We elaborate on this in the discussion following Proposition 1 in Section 2.

As mentioned above, our main result depends crucially upon the fact that we are dealing with a small random perturbation of a two-dimensional Hamiltonian system. In fact, much of what we do is valid in this general setting. In Section 2 we develop a formalism in the setting of small random perturbations of Hamiltonian systems. The main result of this section, Theorem 4, asserts that, under suitable conditions, the top Lyapunov exponent (corresponding to linearizing along trajectories) has the behavior described in (6) and (7) except that now the function $G(h)$ and the operator $\mathcal{N}$ are described in more general terms. Then in Section 3 we verify the conditions of Theorem 4, and obtain the more precise description of the function G and the coefficients for the operator $\mathcal{N}$, thus proving Theorem 2. The proof of Theorem 3 also appears in Section 3.

2. Stochastic Hamiltonian systems. Throughout this section we assume that $H\colon \mathbf{R}^2 \to \mathbf{R}$ is a smooth function with isolated critical points such that $H(x) \to \infty$ as $\|x\| \to \infty$. We write $H_i = \partial H/\partial x^i$ and $H_{ij} = \partial^2 H/\partial x^i \partial x^j$. Then $\nabla H(x) = [H_1(x), H_2(x)]^T$ and $\overline{\nabla} H(x) = [H_2(x), -H_1(x)]^T$. The Hamiltonian system associated with H is

$$\dot{x} = \overline{\nabla} H(x)$$

with flow denoted $x_t = \Phi_t(x)$. Then $v_t = D\Phi_t(x)v$ satisfies

$$\dot{v}_t = \begin{bmatrix} H_{12}(x_t) & H_{22}(x_t) \\ -H_{11}(x_t) & -H_{12}(x_t) \end{bmatrix} v_t.$$

In order to see the nilpotent structure of the linearized Hamiltonian flow we define vector fields

$$U_1(x) = \overline{\nabla} H(x) \quad \text{and} \quad U_2(x) = \frac{\nabla H(x)}{\|\nabla H(x)\|^2},$$

and then we write v_t in the moving frame given by $\Gamma(H(x))U_1(x)$ and $U_2(x)$. Here Γ is a smooth positive function we shall choose later. More precisely, we write

$$v_t = w_{1,t}\Gamma\big(H(x_t)\big)U_1(x_t) + w_{2,t}U_2(x_t) \tag{8}$$

and derive the equation for the vector

$$w_t = \begin{bmatrix} w_{1,t} & w_{2,t} \end{bmatrix}^T.$$

Throughout the paper we will use the notation $V.f$ to denote the action of a vector field V as a first order differential operator acting on a function f. Thus $V.f(x) = DF(x)(V(x)) = \langle \nabla f(x), V(x)\rangle$.

LEMMA 1. *If x is not a critical point of H then* $(U_1.H)(x) = 0$, $(U_2.H)(x) = 1$ *and*

$$DU_1(x)\big(U_2(x)\big) - DU_2(x)\big(U_1(x)\big) = J(x)U_1(x) \tag{9}$$

where

$$J(x) = \frac{[H_2(x)^2 - H_1(x)^2][H_{22}(x) - H_{11}(x)] + 4H_1(x)H_2(x)H_{12}(x)}{[H_1(x)^2 + H_2(x)^2]^2}.$$

PROOF. Direct calculation. □

It can now easily be shown that

$$\dot{w}_t = \begin{bmatrix} 0 & J(x_t)/\Gamma\big(H(x_t)\big) \\ 0 & 0 \end{bmatrix} w_t. \tag{10}$$

The next result contains a generalization which will be useful when we consider perturbations of Hamiltonian systems.

LEMMA 2. *For any vector field V on $\mathbf{R}^2$ write*

$$V(x) = \alpha^1(x)U_1(x) + \alpha^2(x)U_2(x) \tag{11}$$

away from the critical points of H. Consider the equation $\dot{x}_t = V(x_t)$ and its linearized version $\dot{v}_t = DV(x_t)v_t$. Write v_t in terms of w_t as in (8). *Then away from the critical points of H we have $\dot{w}_t = M(x_t)w_t$ where*

$$M(x) = \begin{bmatrix} (U_1.\alpha^1)(x) - \big[J(x) + (\log\Gamma)'\big(H(x)\big)\big]\alpha^2(x) & \dfrac{(U_2.\alpha^1)(x) + J(x)\alpha^1(x)}{\Gamma(H(x))} \\ \Gamma(H(x))(U_1.\alpha^2)(x) & (U_2.\alpha^2)(x) \end{bmatrix}.$$

PROOF. Substitute $V(x) = \alpha^1(x)U_1(x) + \alpha^2(x)U_2(x)$ in the equation $\dot{v}_t = DV(x_t)(v_t)$ to get an expression for $\dot{v}$ first in terms of α^i, U_i and v and then, using (8), in terms of α^i, U_i, Γ and w. Differentiate (8) with respect to t to give an expression for $\dot{v}$ in terms of U_i, Γ, w and $\dot{w}$. Equate these two expressions for $\dot{v}$ to get an equation involving α^i, U_i, Γ, w and $\dot{w}$. Take the components of this equation in the directions U_1 and U_2, and use (9) to obtain the desired result. □

Now we consider a small perturbation of the original Hamiltonian system

$$dx_t = \overline{\nabla} H(x_t)\,dt + \varepsilon^2 V_0(x_t)\,dt + \varepsilon \sum_{i=1}^{r} V_i(x_t) \circ dW_t^i \tag{12}$$

for given smooth vector fields $V_0, V_1, \ldots, V_r$. Notice that we have chosen to write (12) using Stratonovich rather than Itô stochastic differentials; this appears to offer some slight reduction in the lengths of some of the calculations, but has no serious effect on the theory. The linearization of (12) is given by

$$dv_t = D(\overline{\nabla} H)(x_t)v_t\,dt + \varepsilon^2 DV_0(x_t)v_t\,dt + \varepsilon \sum_{i=1}^{r} DV_i(x_t)v_t \circ dW_t^i. \tag{13}$$

Following the notation in Lemma 2 we write

$$V_i(x) = \alpha_i^1(x)U_1(x) + \alpha_i^2(x)U_2(x)$$

for $i \geq 0$. Then (12) becomes

$$dx_t = U_1(x_t)\,dt + \varepsilon^2(\alpha_0^1 U_1 + \alpha_0^2 U_2)(x_t)\,dt + \varepsilon \sum_{i=1}^{r} (\alpha_i^1 U_1 + \alpha_i^2 U_2)(x_t) \circ dW_t^i \tag{14}$$

and the process $H_t \equiv H(x_t)$ satisfies

$$\begin{aligned} dH_t &= \varepsilon^2 \alpha_0^2(x_t)\,dt + \varepsilon \sum_{i=1}^{r} \alpha_i^2(x_t) \circ dW_t^i \\ &= \varepsilon^2 \left(\alpha_0^2(x_t) + \frac{1}{2}\sum_{i=1}^{r} V_i.\alpha_i^2(x_t)\right) dt + \varepsilon \sum_{i=1}^{r} \alpha_i^2(x_t)\,dW_t^i. \end{aligned} \tag{15}$$

LEMMA 3. *Let w_t represent the linearized process v_t in the moving frame given by* (8). *Up to the first time that the process x_t given by* (12) *hits a critical point of H, or explodes to infinity, we have*

$$dw_t = \begin{bmatrix} 0 & \dfrac{J(x_t)}{\Gamma(H(x_t))} \\ 0 & 0 \end{bmatrix} w_t\,dt + \varepsilon^2 M_0(x_t) w_t\,dt + \varepsilon \sum_{i=1}^{r} M_i(x_t) w_t \circ dW_t^i$$

where for $i = 0, 1, \ldots, r$,

$$M_i(x) = \begin{bmatrix} (U_1.\alpha_i^1)(x) - [J(x) + (\log\Gamma)'(H(x))]\alpha_i^2(x) & \dfrac{(U_2.\alpha_i^1)(x) + J(x)\alpha_i^1(x)}{\Gamma(H(x))} \\ \Gamma(H(x))(U_1.\alpha_i^2)(x) & (U_2.\alpha_i^2)(x) \end{bmatrix}. \tag{16}$$

PROOF. The proof is essentially the same as that of Lemma 2. Instead of equating two expressions for $\dot{v}_t$, we equate two semimartingale expressions for dv_t. First we equate the martingale parts (which can be read off from the Stratonovich differentials) to obtain the matrices $M_i(x)$ for $i \geq 1$. Then we can subtract all the Stratonovich differential terms from both sides, and equate the remaining (bounded variation) terms to obtain $M_0(x)$. $\square$

At this point it is clear that we are dealing with a small perturbation of a nilpotent system. Accordingly, we follow the method of Pinsky and Wihstutz [20] and transform $\mathbf{R}^2$ using the transformation

$$T = \begin{bmatrix} \varepsilon^{2/3} & 0 \\ 0 & 1 \end{bmatrix}.$$

It is easy to see that for each fixed ε the process Tw_t will have the same Lyapunov exponent as w_t. For simplicity of notation we continue to write w_t; it is now given by the equation

$$
\begin{aligned}
dw_t = {}& \varepsilon^{2/3}\begin{bmatrix} 0 & \dfrac{J(x_t)}{\Gamma(H(x_t))} \\ 0 & 0 \end{bmatrix} w_t\,dt + \varepsilon^2 M_0^\varepsilon(x_t) w_t\,dt \\
& + \varepsilon \sum_{i=1}^r M_i^\varepsilon(x_t) w_t \circ dW_t^i
\end{aligned}
\tag{17}
$$

where for $i = 0, 1, \dots, r$,

$$
\begin{aligned}
M_i^\varepsilon(x) &= T M_i(x) T^{-1} \\
&= \begin{bmatrix} (U_1.\alpha_i^1)(x) - [J(x) + (\log\Gamma)'(H(x))]\alpha_i^2(x) & \dfrac{\varepsilon^{2/3}[(U_2.\alpha_i^1)(x) + J(x)\alpha_i^1(x)]}{\Gamma(H(x))} \\ \varepsilon^{-2/3}\Gamma(H(x))(U_1.\alpha_i^2)(x) & (U_2.\alpha_i^2)(x) \end{bmatrix}.
\end{aligned}
$$

Write

$$
w_t = \|w_t\| \begin{bmatrix} \cos\theta_t \\ \sin\theta_t \end{bmatrix}.
$$

Then from (17) we can obtain Stratonovich SDEs for $\log\|w_t\|$ and θ_t. Converting these to Itô form, using (14) and (15) we get

$$
\begin{aligned}
d\big(\log\|w_t\|\big) = {}& \big[\varepsilon^{2/3} Q_0(x_t,\theta_t) + \varepsilon^{4/3} Q_1(x_t,\theta_t) + \varepsilon^2 Q_2(x_t,\theta_t) \\
& \quad + \varepsilon^{8/3} Q_3(x_t,\theta_t) + \varepsilon^{10/3} Q_4(x_t,\theta_t)\big]\,dt \\
& + \sum_{i=1}^r \big[\varepsilon^{1/3} Q_5^i(x_t,\theta_t) + \varepsilon Q_6^i(x_t,\theta_t) + \varepsilon^{5/3} Q_7^i(x_t,\theta_t)\big]\,dW_t^i
\end{aligned}
\tag{18}
$$

and

$$
\begin{aligned}
d\theta_t = {}& \big[\varepsilon^{2/3} t_0(x_t,\theta_t) + \varepsilon^{4/3} t_1(x_t,\theta_t) + \varepsilon^2 t_2(x_t,\theta_t) \\
& \quad + \varepsilon^{8/3} t_3(x_t,\theta_t) + \varepsilon^{10/3} t_4(x_t,\theta_t)\big]\,dt \\
& + \sum_{i=1}^r \big[\varepsilon^{1/3} t_5^i(x_t,\theta_t) + \varepsilon t_6^i(x_t,\theta_t) + \varepsilon^{5/3} t_7^i(x_t,\theta_t)\big]\,dW_t^i
\end{aligned}
\tag{19}
$$

where

$$
\begin{aligned}
Q_0(x,\theta) = {}& \frac{J(x)}{\Gamma(H(x))}\cos\theta\sin\theta \\
& + \Gamma^2\big(H(x)\big)\sum_{i=1}^r \big[U_1.\alpha_i^2(x)\big]^2 \bigg[\frac{1}{2}\cos^2\theta - \sin^2\theta\cos^2\theta\bigg]
\end{aligned}
$$

and

$$t_0(x,\theta) = -\frac{J(x)}{\Gamma(H(x))}\sin^2\theta - \Gamma^2\big(H(x)\big)\sum_{i=1}^{r}\big[U_1.\alpha_i^2(x)\big]^2\cos^3\theta\sin\theta$$

and

$$t_5^i(x,\theta) = \Gamma\big(H(x)\big)(U_1.\alpha_i^2)(x)\cos^2\theta.$$

We shall see that the remaining Q_k and t_k do not contribute to the leading term in the asymptotic expansion for $\overline{\lambda}(\varepsilon)$. However the following result will be important in obtaining integrability and growth estimates on these terms; its proof is by direct calculation, using Itô's formula and the formula for the Wong–Zakai correction term.

LEMMA 4. *Each of the functions* Q_1, Q_2, Q_3, Q_4, $(Q_5^i)^2$, $(Q_6^i)^2$, $(Q_7^i)^2$, t_1, t_2, t_3, t_4, $(t_5^i)^2$, $(t_6^i)^2$ *and* $(t_7^i)^2$ *is a sum of terms of the form* $f(\theta)g(x)$ *where each* f *is a trigonometric polynomial and each* g *is either* (i) *an entry in* $M_0(x)$, *or* (ii) *an entry in* $DM_i(x)(V_i(x))$ *for* $i \geq 1$, *or* (iii) *a product of two entries in* $M_i(x)$ *for* $i \geq 1$. *Here* $M_i(x)$ *is the matrix-valued function given in* (16).

At this point we need to make some assumptions about the Hamiltonian function H and the perturbing vector fields $V_0, V_1, \ldots, V_r$. Let M denote the space $\mathbf{R}^2$ with the critical points of H removed.

(A1) For each sufficiently small $\varepsilon > 0$ the process $\{(x_t, \theta_t) : t \geq 0\}$ given by (12) and (19) is a positive recurrent diffusion process on $M \times S^1$ with (unique) stationary probability P_ε, say. We write μ_ε for the M marginal of P_ε.
(A2) For each sufficiently small $\varepsilon > 0$, the functions $\log\|\nabla H\|$ and $\log\Gamma$ are integrable with respect to μ_ε.
(A3) For each sufficiently small $\varepsilon > 0$, the functions $Q_0, \ldots, Q_4$, $[Q_5^i]^2$, $[Q_6^i]^2$ and $[Q_7^i]^2$ are all integrable with respect to P_ε.
(A4) There is $K < \infty$ so that for each sufficiently small $\varepsilon > 0$ we have $|\int Q_k\,dP_\varepsilon| \leq K$ for $k = 1, 2, 3, 4$.

PROPOSITION 1. *Assume* (A1)–(A4). *The Lyapunov exponent* $\overline{\lambda}(\varepsilon)$ *for the linearization* v_t *given by* (13) *satisfies*

$$\overline{\lambda}(\varepsilon) = \varepsilon^{2/3}\int_{M\times S^1} Q_0(x,\theta)\,dP_\varepsilon(x,\theta) + O(\varepsilon^{4/3})$$

as $\varepsilon \to 0$.

PROOF. Formula (8) relating v_t and w_t yields the inequalities

$$\|v_t\| \leq \max\big(\Gamma(x_t)\|\nabla H(x_t)\|, \|\nabla H(x_t)\|^{-1}\big)\|w_t\|$$

and

$$\|w_t\| \le \max\bigl(\Gamma(x_t)^{-1}\|\nabla H(x_t)\|^{-1}, \|\nabla H(x_t)\|\bigr)\|v_t\|.$$

Assumptions (A1) and (A2) imply that $\|\nabla H(x_t)\|$ and $\Gamma(x_t)$ are tempered (see [2], Section 4.1), so that $\lim_{t\to\infty}(1/t)[\log\|v_t\| - \log\|w_t\|] = 0$, and therefore for each ε the processes v_t and w_t have the same Lyapunov exponent. It now follows from (18) and assumption (A3) that

$$\overline{\lambda}(\varepsilon) = \lim_{t\to\infty}\frac{1}{t}\log\|w_t\| = \int_{M\times S_1} Q_\varepsilon(x,\theta)\,dP_\varepsilon(x,\theta)$$

where

$$\begin{aligned} Q_\varepsilon(x,\theta) &= \varepsilon^{2/3}Q_0(x,\theta) + \varepsilon^{4/3}Q_1(x,\theta) + \varepsilon^2 Q_2(x,\theta) + \varepsilon^{8/3}Q_3(x,\theta) \\ &\quad + \varepsilon^{10/3}Q_4(x,\theta). \end{aligned}$$

The result now follows easily using assumption (A4). $\square$

We now proceed to estimate the integral $\int Q_0(x,\theta)\,dP_\varepsilon(x,\theta)$. Recall that P_ε is the invariant probability measure for the process (x_t,θ_t) on $M\times S^1$. From the equations (12) for x_t and (19) for θ_t we can write down the generator $\mathcal{L}_\varepsilon$, say, for the (x_t,θ_t) process. Define functions

$$\begin{aligned} a(x) &= J(x)/\Gamma\bigl(H(x)\bigr), \\ b(x) &= \bigl[\Gamma\bigl(H(x)\bigr)\bigr]^2\sum_{i=1}^r\bigl[U_1.\alpha_i^2(x)\bigr]^2, \\ c(x) &= \alpha_0^2(x) + \tfrac12\sum_{i=1}^r V_i.\alpha_i^2(x), \\ d(x) &= \tfrac12\sum_{i=1}^r\bigl[\alpha_i^2(x)\bigr]^2, \end{aligned}$$

and operators

$$\begin{aligned} \overline{L} &= V_0 + \tfrac12\sum_{i=1}^r V_i^2, \\ \widetilde{\mathcal{L}}_\varepsilon &= U_1 + \varepsilon^2\overline{L} \\ &\quad + \varepsilon^{2/3}\Bigl[-a(x)\sin^2\theta\frac{\partial}{\partial\theta} + b(x)\Bigl(-\cos^3\theta\sin\theta\frac{\partial}{\partial\theta} + \frac12\cos^4\theta\frac{\partial^2}{\partial\theta^2}\Bigr)\Bigr]. \end{aligned}$$

Then $\mathcal{L}_\varepsilon = \widetilde{\mathcal{L}}_\varepsilon$ plus terms of order $\varepsilon^{4/3}$ and higher involving at least one derivative with respect to θ. Also, for any function F of a real variable we have

$$\mathcal{L}_\varepsilon(F\circ H)(x) = \varepsilon^2\overline{L}(F\circ H)(x) = \varepsilon^2\bigl[c(x)F'\bigl(H(x)\bigr) + d(x)F''\bigl(H(x)\bigr)\bigr].$$

From the formula for $\mathcal{L}_\varepsilon$, or equivalently from the equations (12), (19) and (15), we see a separation of time scales. The x_t process moves around the orbits of the Hamiltonian system at rate 1; the angle θ moves around S^1 at rate $\varepsilon^{2/3}$; and the process $H_t \equiv H(x_t)$ moves at rate ε^2. The existence of two different time scales for motion around orbits and motion across orbits is the basis of many results on stochastic averaging for small perturbations of Hamiltonian systems, from early work by Khas'minskii [16] to more recent studies by Freidlin and Wentzell [13], Freidlin and Weber [11, 12] and Liang and Sri Namachchivaya [18]. Here we see that to estimate the Lyapunov exponent we need to deal also with the motion θ_t running at a third, intermediate, rate.

Our technique will involve averaging over each of the orbits of the unperturbed Hamiltonian system, then averaging over the angle θ, and finally averaging over the space of orbits. More precisely, we will use the adjoint method and build, for each $\varepsilon > 0$ a function $f_\varepsilon(x, \theta)$ so that

$$\mathcal{L}_\varepsilon f_\varepsilon(x,\theta) = Q_0(x,\theta) - \overline{\lambda} + \varepsilon^{2/3} \sum_{k=0}^{4} \varepsilon^{2k/3} \phi_k(x,\theta)$$

for suitable constant $\overline{\lambda}$ and functions ϕ_k. Then, subject to growth and integrability conditions we get

$$0 = \int Q_0 \, dP_\varepsilon - \overline{\lambda} + \varepsilon^{2/3} \sum_{k=0}^{4} \varepsilon^{2k/3} \int \phi_k \, dP_\varepsilon \tag{20}$$

so that

$$\int Q_0 \, dP_\varepsilon = \overline{\lambda} + O(\varepsilon^{2/3}). \tag{21}$$

In the adjoint method the stochastic averaging is carried out by averaging certain of the coefficients of the operator $\widetilde{\mathcal{L}}_\varepsilon$, and then constructing f_ε in terms of the averaged version of $\widetilde{\mathcal{L}}_\varepsilon$. In order to carry this out, we will make some further assumptions.

(A5) The function H has a single critical point.

Under assumption (A5) we can assume without loss of generality that the critical point is at 0 and that $H(0) = 0$. Then for each $h > 0$ the unperturbed Hamiltonian flow has a simple closed orbit $H^{-1}(h)$. Let $T(h)$ denote the period of this orbit, and let m_h denote the probability measure given by time averaging around this orbit. Define functions

$$\overline{c}(h) = \int c(x) \, dm_h(x) = \int \left[\alpha_0^2(x) + \tfrac{1}{2} \sum_{i=1}^{r} (V_i . \alpha_i^2)(x) \right] dm_h(x) \tag{22}$$

and

$$\overline{d}(h) = \int d(x)\,dm_h(x) = \frac{1}{2}\int \sum_{i=1}^{r} [\alpha_i^2(x)]^2\,dm_h(x) \tag{23}$$

and the operator

$$\mathcal{N} = \overline{c}(h)\frac{\partial}{\partial h} + \overline{d}(h)\frac{\partial^2}{\partial h^2}. \tag{24}$$

(A6) $T'(h) \neq 0$ for all $h > 0$.
(A7) $\int \sum_{i=1}^{r} [(U_1.\alpha_i^2)(x)]^2\,dm_h(x) > 0$ for all $h > 0$.
(A8) $\mathcal{N}$ is the generator of a positive recurrent diffusion on $(0, \infty)$ with invariant probability ρ.

The following result explains why we will need (A6).

LEMMA 5. *Assume* (A5). *Then for* $h > 0$,

$$\int J(x)\,dm_h(x) = -\frac{T'(h)}{T(h)}.$$

PROOF. Let $v_t = D\Phi_t(x)(v)$ denote the linearized version of the unperturbed Hamiltonian flow, and write $v_t = w_1(t)U_1(x_t) + w_2(t)U_2(x_t)$. Taking $\Gamma \equiv 1$ in (10) we get

$$\dot{w}(t) = \begin{bmatrix} 0 & J(x_t) \\ 0 & 0 \end{bmatrix} w(t)$$

so that for $H(x) = h$

$$w(T(h)) = \begin{bmatrix} 1 & \int_0^{T(h)} J(x_t)\,dt \\ 0 & 1 \end{bmatrix} w(0). \tag{25}$$

But we also have the identity

$$\Phi_{T(H(x))}(x) = x$$

for all x. Differentiating this expression we get

$$U_1(x)T'\big(H(x)\big)DH(x)(v) + v_{T(H(x))} = v. \tag{26}$$

Taking $v = U_2(x)$ in (26) and comparing with (25) we obtain

$$\int J(x)\,dm_h(x) = \frac{1}{T(h)}\int_0^{T(h)} J(x_t)\,dt = -\frac{T'(h)}{T(h)}$$

as desired. □

Define, for $h > 0$,

$$G_1(h) = \int J(x)\,dm_h(x), \tag{27}$$

$$G_2(h) = \int \sum_{i=1}^{r} [(U_1.\alpha_i^2)(x)]^2\,dm_h(x) \tag{28}$$

and

$$G(h) = \left|G_1(h)\right|^{2/3}\left[G_2(h)\right]^{1/3}. \tag{29}$$

Up to this point the positive function Γ has been arbitrary, subject to the requirements of (A2)–(A4). Now we choose to define Γ by

$$\left[\Gamma(h)\right]^3 = \frac{G_1(h)}{G_2(h)} \tag{30}$$

for $h > 0$. Assumptions (A6)–(A7) together with Lemma 5 imply that $G(h)$ and $\Gamma(h)$ are positive and finite for all $h > 0$. The choice of Γ will ensure that the terms of order $\varepsilon^{2/3}$ in $\widetilde{\mathcal{L}}_\varepsilon$, when averaged with respect to m_h, are of the form $\varepsilon^{2/3}G(h)\overline{M}$ where the operator $\overline{M}$, defined in (32), contains no dependence on h.

We will construct f_ε in a sequence of four steps.

Step 1. The choice of Γ implies that

$$\int a(x)\,dm_h(x) = \int b(x)\,dm_h(x) = G(h).$$

Since $a(x) - G(H(x))$ and $b(x) - G(H(x))$ have zero mean with respect to m_h for all h, there exist functions $A(x)$ and $B(x)$ such that

$$(U_1.A)(x) = a(x) - G\big(H(x)\big),$$
$$(U_1.B)(x) = b(x) - G\big(H(x)\big).$$

It follows that

$$\begin{aligned} &U_1.\big(A(x)\cos\theta\sin\theta + B(x)\big[\tfrac{1}{2}\cos^2\theta - \cos^2\theta\sin^2\theta\big]\big) \\ &\qquad = Q_0(x,\theta) - G(H(x))\big[\cos\theta\sin\theta + \tfrac{1}{2}\cos^2\theta - \cos^2\theta\sin^2\theta\big]. \end{aligned} \tag{31}$$

Step 2. Define operators

$$\overline{M} = (-\sin^2\theta - \cos^3\theta\sin\theta)\frac{\partial}{\partial\theta} + \frac{1}{2}\cos^4\theta\frac{\partial^2}{\partial\theta^2} \tag{32}$$

and

$$\begin{aligned} \widetilde{M} = &-\big[a(x) - G\big(H(x)\big)\big]\sin^2\theta\frac{\partial}{\partial\theta} \\ &+ \big[b(x) - G\big(H(x)\big)\big]\left[-\cos^3\theta\sin\theta\frac{\partial}{\partial\theta} + \frac{1}{2}\cos^4\theta\frac{\partial^2}{\partial\theta^2}\right] \end{aligned}$$

so that

$$\widetilde{\mathcal{L}}_\varepsilon = U_1 + \varepsilon^2 \overline{L} + \varepsilon^{2/3}\big[G\big(H(x)\big)\overline{M} + \widetilde{M}\,\big].$$

The operator $\overline{M}$ is hypoelliptic on S^1; let ν denote the corresponding invariant measure on S^1. Define

$$\gamma_0 = \int \big[\cos\theta \sin\theta + \tfrac{1}{2}\cos^2\theta - \cos^2\theta \sin^2\theta\big]\,d\nu(\theta), \tag{33}$$

then there is a smooth bounded function $R(\theta)$ so that

$$\overline{M}R(\theta) = \cos\theta \sin\theta + \tfrac{1}{2}\cos^2\theta - \cos^2\theta\sin^2\theta - \gamma_0.$$

Then

$$\begin{aligned}\widetilde{\mathcal{L}}_\varepsilon(\varepsilon^{-2/3}R)(\theta) = G\big(H(x)\big)&\big[\cos\theta\sin\theta + \tfrac{1}{2}\cos^2\theta - \cos^2\theta\sin^2\theta\big] \\ &- G\big(H(x)\big)\gamma_0 + \widetilde{M}R(\theta).\end{aligned} \tag{34}$$

Here γ_0 is a fixed constant. As in [9] we can identify it as the top Lyapunov exponent of the linear SDE

$$dv_t = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} v_t\,dt + \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} v_t\,dW_t.$$

Numerically $\gamma_0 \sim 0.29$. An exact formula for γ_0 is given by Ariaratnam and Xie [1]. Notice that our choice of the function Γ has ensured that the invariant probability ν, the constant γ_0, and the function R all do not depend on h.

Step 3. Repeating the ideas of Step 1 we calculate

$$\begin{aligned}&U_1.\big[A(x)\sin^2\theta\,R'(\theta) + B(x)\big(-\cos^3\theta\sin\theta\,R'(\theta) + \tfrac{1}{2}\cos^4\theta\,R''(\theta)\big)\big] \\ &\quad = \widetilde{M}R(\theta).\end{aligned} \tag{35}$$

Step 4. For any function on $\mathbf{R}^2$ of the form $F(H(x))$ we have

$$\overline{L}(F \circ H)(x) = c(x)F'\big(H(x)\big) + d(x)F''\big(H(x)\big).$$

Recall the definitions of $\overline{c}(h)$ and $\overline{d}(h)$ and $\mathcal{N}$ in assumption (A8). The operator $\mathcal{N}$ is the stochastically averaged operator for the "slow" motion $H_t = H(x_t)$. If there is a smooth function $\Psi(h)$ such that

$$(\mathcal{N}\Psi)(h) = G(h) - \int_{(0,\infty)} G\,d\rho$$

then we have

$$\mathcal{L}_\varepsilon(\varepsilon^{-2}\Psi \circ H)(x) = G\big(H(x)\big) - \int G\,d\rho + \widetilde{\mathcal{N}}\Psi\big(H(x)\big) \tag{36}$$

where the operator $\widetilde{\mathcal{N}}$ is given by

$$\widetilde{\mathcal{N}} = \big[c(x) - \overline{c}\big(H(x)\big)\big]\frac{\partial}{\partial h} + \big[d(x) - \overline{d}\big(H(x)\big)\big]\frac{\partial^2}{\partial h^2}.$$

Since the coefficients of $\widetilde{\mathcal{N}}$ are of mean zero with respect to m_h for each h we can find functions $C(x)$ and $D(x)$ so that

$$(U_1.C)(x) = c(x) - \overline{c}\big(H(x)\big),$$
$$(U_1.D)(x) = d(x) - \overline{d}\big(H(x)\big).$$

Then

$$U_1.\big[C(x)\Psi'\big(H(x)\big) + D(x)\Psi''\big(H(x)\big)\big] = \widetilde{\mathcal{N}}\Psi\big(H(x)\big). \tag{37}$$

Now define $f_\varepsilon(x,\theta)$ by

$$\begin{aligned} f_\varepsilon(x,\theta) = {} & A(x)\cos\theta\sin\theta + B(x)\big[\tfrac{1}{2}\cos^2\theta - \cos^2\theta\sin^2\theta\big] \\ & + \varepsilon^{-2/3}R(\theta) \\ & - \big[A(x)\sin^2\theta R'(\theta) + B(x)\big(-\cos^3\theta\sin\theta R'(\theta) + \tfrac{1}{2}\cos^4\theta R''(\theta)\big)\big] \\ & + \varepsilon^{-2}\gamma_0(\Psi\circ H)(x) \\ & - \gamma_0\big[C(x)\big(\Psi'\circ H\big)(x) + D(x)\big(\Psi''\circ H\big)(x)\big]. \end{aligned}$$

Then using equations (31), (34), (35), (36) and (37) we get

$$(\mathcal{L}_\varepsilon f_\varepsilon)(x,\theta) = Q_0(x,\theta) - \gamma_0\int G\,d\rho + \varepsilon^{2/3}\sum_{k=0}^{4}\varepsilon^{2k/3}\phi_k(x,\theta) \tag{38}$$

where the functions $\phi_0,\dots,\phi_4$ are given in the Appendix. The terms involving the $\phi_k(x,\theta)$ appear because of the action of $\mathcal{L}_\varepsilon - U_1$ or $\mathcal{L}_\varepsilon - \widetilde{\mathcal{L}}_\varepsilon$ on the various terms in f_ε. We note that the functions $A(x)$, $B(x)$, $C(x)$, $D(x)$ and $\Psi(h)$ and some of their derivatives appear in the formulas for the ϕ_k.

At this point we wish to integrate equation (38) with respect to P_ε and claim that $\int \mathcal{L}_\varepsilon f_\varepsilon\,dP_\varepsilon = 0$. However we are working on a non-compact space, and so the assertion $\int \mathcal{L}_\varepsilon f_\varepsilon\,dP_\varepsilon = 0$ may fail without some further restrictions on f_ε. For example, if L denotes the standard Ornstein–Uhlenbeck operator on the real line, with invariant probability μ there exists a smooth function f so that $Lf \equiv 1$ and then $\int Lf\,d\mu = 1 \neq 0$. The following result, quoted from Baxendale and Goukasian [9], gives a sufficient condition.

PROPOSITION 2. *Let $\{y_t : t \geq 0\}$ be a diffusion process on a σ-compact manifold N with invariant probability measure μ. Let L be an operator acting* *on* $C^2(N)$ *functions that agrees with the generator of* $\{y_t : t \geq 0\}$ *on* C^2 *functions with compact support. Let* $f \in C^2(N)$ *and* $g \in C(N)$ *be* μ*-integrable functions satisfying* $Lf = g$. *Suppose there exists a positive* $F \in C^2(N)$ *satisfying* $LF(y) \leq kF(y)$ *for some* $k < \infty$ *such that* $\sup\{|g(y)|/F(y) : y \in N\} < \infty$ *and* $f(y)/F(y) \to 0$ *as* $y \to \infty$. *Then* $\int_N g(y)\, d\mu(y) = 0$.

At this point we can complete our list of assumptions.

(A9) G is integrable with respect to ρ and there is a smooth function $\Psi(h)$ such that $(\mathcal{N}\Psi)(h) = G(h) - \int_{(0,\infty)} G\, d\rho$ for $h > 0$.
(A10) For each sufficiently small $\varepsilon > 0$, the functions f_ε and ϕ_k, $0 \leq k \leq 4$ are all integrable with respect to P_ε.
(A11) There is $K < \infty$ so that for each sufficiently small $\varepsilon > 0$ we have $|\int \phi_k\, dP_\varepsilon| \leq K$ for $0 \leq k \leq 4$.
(A12) For each sufficiently small $\varepsilon > 0$, there exists a positive $F \in C^2(M \times S^1)$ satisfying $\mathcal{L}_\varepsilon F(x,\theta) \leq kF(x,\theta)$ for some $k < \infty$ such that $\sup\{|Q_0(x,\theta)|/F(x,\theta) : (x,\theta) \in M \times S^1\} < \infty$ and $\sup\{|\phi_k(x,\theta)|/F(x,\theta) : (x,\theta) \in M \times S^1\} < \infty$ for $0 \leq k \leq 4$ and $f_\varepsilon(x,\theta)/F(x,\theta) \to 0$ as $(x,\theta) \to \infty$ in $M \times S^1$.

THEOREM 4. *Assume* (A1)–(A12). *Then the top Lyapunov exponent for the linearized version* (13) *of the perturbed Hamiltonian system* (12) *satisfies*

$$\overline{\lambda}(\varepsilon) = \varepsilon^{2/3}\overline{\lambda} + O(\varepsilon^{4/3}) \qquad \text{as } \varepsilon \to 0$$

where

$$\overline{\lambda} = \gamma_0 \int_{(0,\infty)} G(h)\, d\rho(h) > 0.$$

Here γ_0 *is the fixed constant* (*approximately* 0.29) *given in* (33), *and* $G(h)$ *is the positive function given by* (27)–(29) *and* ρ *is the invariant probability for the stochastically averaged* $H(x_t)$ *process with generator* $\mathcal{N}$ *given by* (22)–(24).

PROOF. Assumption (A9) allows us to obtain (38). Assumptions (A10)–(A11) together with Proposition 2 allow the passage to (20), and then (A12) gives (21). Together with Proposition 1 we have the desired result. □

REMARK 1. In the next section we shall take $H(x_1, x_2) = x_1^2/2 + x_1^4/4 + x_2^2/2$. Notice that the methods of this section do not apply to the Hamiltonian $x_1^2/2 + x_2^2/2$ because the period $T(h)$ is constant. Some of the methods may turn out to be useful for the Hamiltonian $-\alpha x_1^2/2 + x_1^4/4 + x_2^2/2$ with $\alpha > 0$, although in this case Assumption (A5) of a single critical point is false and the structure of the averaged Hamiltonian process is more complicated; see [13].

REMARK 2. In Theorem 4 we need a long list of assumptions dealing with integrability and growth because the state spaces M and $(0,\infty)$ are non-compact. This should be compared with the elegant statement of the adjoint method in Arnold, Papanicolaou and Wihstutz [3] when dealing with a compact state space. When verifying conditions (A3), (A4), (A10) and (A11), it is useful to note that the functions Q_k, Q_k^i, t_k, t_k^i, ϕ_k and f_ε are each sums of terms of the form $f(\theta)g(x)$ where each f is bounded. Thus it suffices to consider the integrals of each of the $g(x)$ with respect to the invariant probability μ_ε for the x_t process. Similarly, in verifying condition (A12), it suffices to estimate the size of the g functions in terms of a function F of x only.

REMARK 3. Recall the measure μ_0 defined in Theorem 3. From the definitions of $G(h)$ and γ_0 we see that

$$\overline{\lambda} = \int Q_0(x,\theta)\,dm_h(x)\,d\nu(\theta)\,d\rho(h) = \int Q_0(x,\theta)\,d\mu_0(x)\,d\nu(\theta)$$

and so the passage from Proposition 1 to Theorem 4 consisted of showing that

$$\int Q_0(x,\theta)\,dP_\varepsilon(x,\theta) = \int Q_0(x,\theta)\,d\mu_0(x)\,d\nu(\theta) + O(\varepsilon^{2/3}).$$

We can replace Q_0 by a general function of the form $g(x)s(\theta)$ for smooth functions g and s and apply the same sequence of 4 steps to construct a function $\tilde{f}_\varepsilon$ such that

$$(\mathcal{L}_\varepsilon \tilde{f}_\varepsilon)(x,\theta) = g(x)s(\theta) - \int g(x)\,d\mu_0(x)\int s(\theta)\,d\nu(\theta) + \varepsilon^{2/3}\sum_{k=0}^{4}\varepsilon^{2k/3}\tilde{\phi}_k(x,\theta).$$

Then, subject to growth and integrability conditions similar to those in Theorem 4, we obtain

$$\left|\int g(x)s(\theta)\,dP_\varepsilon(x,\theta) - \int g(x)s(\theta)\,d\mu_0(x)\,d\nu(\theta)\right| = O(\varepsilon^{2/3}) \qquad \text{as } \varepsilon \to 0.$$

If s is a constant then some of the terms in $\tilde{f}_\varepsilon$ vanish, and we obtain

$$\left|\int g(x)\,d\mu_\varepsilon(x) - \int g(x)\,d\mu_0(x)\right| = O(\varepsilon^2) \qquad \text{as } \varepsilon \to 0.$$

Thus, in some sense, P_ε converges to $\mu_0 \times \nu$ at rate $\varepsilon^{2/3}$ and μ_ε converges to μ_0 at rate ε^2. We will make these statements more precise in Section 3.6.

3. Stochastic nonlinear oscillator. In this section we write $(x,y) \in \mathbf{R}^2$ and $r = \sqrt{x^2+y^2}$ and consider the system

$$\begin{aligned} dx &= y\,dt, \\ dy &= (-x - x^3 + \varepsilon^2\beta y - \varepsilon^2 b x^2 y)\,dt + \varepsilon\sigma x\,dW_t. \end{aligned} \tag{39}$$

Throughout this section we assume that $\sigma > 0$, $b \geq 0$ and $4\beta + \sigma^2 > 0$. If $b = 0$ we assume also that $\beta < 0$. By Theorem 1 and the estimate (5), the condition $4\beta + \sigma^2 > 0$ ensures that the one-point motion (x_t, y_t) is positive recurrent on $\mathbf{R}^2 \setminus \{(0,0)\}$ for all sufficiently small $\varepsilon > 0$.

We have two tasks to carry out in this section. We need to show that conditions (A1)–(A12) are satisfied, and we need to obtain explicit formulas for the function $G(h)$ and the coefficients $\overline{c}(h)$ and $\overline{d}(h)$ appearing in the operator $\mathcal{N}$.

The system (39) is a perturbation of the Hamiltonian system with $H(x, y) = x^2/2 + x^4/4 + y^2/2$ by the vector fields

$$V_0(x, y) = (\beta - bx^2)\begin{bmatrix} 0 \\ y \end{bmatrix}, \qquad V_1(x, y) = \sigma \begin{bmatrix} 0 \\ x \end{bmatrix}.$$

We have

$$U_1(x, y) = \begin{bmatrix} y \\ -x - x^3 \end{bmatrix}, \qquad U_2(x, y) = \frac{1}{(x + x^3)^2 + y^2}\begin{bmatrix} x + x^3 \\ y \end{bmatrix}$$

and so

$$\alpha_0^1(x, y) = -\frac{(\beta - bx^2)y(x + x^3)}{(x + x^3)^2 + y^2},$$
$$\alpha_0^2(x, y) = (\beta - bx^2)y^2,$$
$$\alpha_1^1(x, y) = \frac{-\sigma x(x + x^3)}{(x + x^3)^2 + y^2},$$
$$\alpha_1^2(x, y) = \sigma xy$$

and

$$J(x, y) = \frac{-3x^2(y^2 - (x + x^3)^2)}{[(x + x^3)^2 + y^2]^2}.$$

Since the system has only one-dimensional noise, we shall write $Q_k^1(x, y, \theta) = Q_k(x, y, \theta)$ and $t_k^1(x, y, \theta) = t_k(x, y, \theta)$ for $k = 5, 6, 7$.

We can see immediately that H has only one fixed point, so that (A5) is satisfied. Also, since $[(U_1.\alpha_1^2)(x, y)]^2 = \sigma^2(y^2 - x^2 - x^4)^2$ we see that (A7) is satisfied.

3.1. *Uniform estimates on integrals.* The generator of the one-point motion (x_t, y_t) is

$$L_\varepsilon = y\frac{\partial}{\partial x} + \bigl(-x - x^3 + \varepsilon^2(\beta - bx^2)y\bigr)\frac{\partial}{\partial y} + \frac{1}{2}\varepsilon^2\sigma^2x^2\frac{\partial^2}{\partial y^2}.$$

LEMMA 6. *For each sufficiently small $\varepsilon > 0$ there is a unique invariant probability measure P_ε on $(\mathbf{R}^2 \setminus \{0\}) \times S^1$ for the diffusion process with generator $\mathcal{L}_\varepsilon$. Moreover there is $\gamma > 0$, depending on ε, such that*

$$\int (x^2 + y^2)^{-\gamma} \, d\mu_\varepsilon(x, y) < \infty, \tag{40}$$

where μ_ε is the marginal distribution of P_ε on $\mathbf{R}^2 \setminus \{0\}$.

PROOF. For sufficiently small $\varepsilon > 0$ the Lyapunov exponent $\lambda(-1, \varepsilon^2\beta, \varepsilon\sigma)$ is positive. The existence and uniqueness of μ_ε and the integrability result (40) now follows from [7], Theorem 2.8. (For the function f in the conditions of [7], Theorem 2.8, we can take the function H_ε defined in Lemma 7.) The existence of P_ε is now immediate, using the compactness of S^1. The uniqueness of P_ε follows from a result of San Martin and Arnold [21], Theorem 5.1, using techniques of geometric control and Lie algebra considerations. Notice that for the purpose of checking the uniqueness of P_ε we can work with respect to the original Euclidean frame for the v_t process. $\square$

REMARK 4. The proof of Theorem 1 uses the same result [7], Theorem 2.8. In this case we use the function

$$f(x, y) = -\left(\frac{\alpha + \delta^2}{2}\right)x^2 + \left(\frac{a}{4} + \frac{\delta b}{12}\right)x^4 + \frac{1}{2}\left(y - (\beta + \delta)x + \frac{b}{3}x^3\right)^2$$

where $\delta > 0$. If $b = 0$ we take $\delta = -2\beta/3$. In case (ii) the invariant probability μ has the first or second behavior near 0 according as $\Lambda(-2) > 0$ or $\Lambda(-2) < 0$ where Λ is the moment Lyapunov function for equation (2). The result ([10], Corollary 2.13) allows the calculation $\Lambda(-2) = -\beta$, and Theorem 1 is proven.

LEMMA 7. *There exist $\alpha > 0$ and $\varepsilon_0 > 0$ and $K < \infty$ such that*

$$\int \exp\bigl(\alpha H(x, y)^{4/9}\bigr) \, d\mu_\varepsilon(x, y) \le K \quad \text{whenever } 0 < \varepsilon \le \varepsilon_0.$$

PROOF. We consider first the case when $b > 0$. For fixed $\delta > 0$ define the function

$$H_\varepsilon(x, y) = \left(\frac{1}{2} - \frac{\varepsilon^4\delta^2}{2}\right)x^2 + \left(\frac{1}{4} + \frac{\varepsilon^2\delta b}{12}\right)x^4 + \frac{1}{2}\left(y - \varepsilon^2(\beta + \delta)x + \frac{\varepsilon^2 b}{3}x^3\right)^2.$$

Then

$$L_\varepsilon H_\varepsilon(x, y) = \varepsilon^2\left[-\delta y^2 - \frac{b}{3}x^6 + \left(\beta + \delta - \frac{b}{3}\right)x^4 + \left(\frac{\sigma^2}{2} + \beta + \delta\right)x^2 + \varepsilon^2\beta\delta xy\right].$$

It is easy to show there exist positive constants K_1, K_2 and k so that

$$H_\varepsilon(x, y) \le K_1\big(x^6 + y^2\big),$$
$$L_\varepsilon H_\varepsilon(x, y) \le -\varepsilon^2 K_2\big(x^6 + y^2\big),$$

whenever $x^2 + y^2 \ge k$ and $0 < \varepsilon \le 1/\sqrt{\delta}$. Then on the set $x^2 + y^2 \ge k$ we have

$$\begin{aligned}
&\frac{L_\varepsilon(\exp(\alpha H_\varepsilon^{2/3}))(x, y)}{\exp(\alpha H_\varepsilon^{2/3})(x, y)} \\
&\quad = \frac{2\alpha}{3} H_\varepsilon(x, y)^{-1/3} L_\varepsilon H_\varepsilon(x, y) \\
&\qquad + \left(\frac{2\alpha^2}{9} H_\varepsilon(x, y)^{-2/3} - \frac{\alpha}{9} H_\varepsilon(x, y)^{-4/3}\right)\varepsilon^2\sigma^2 x^2 \left(\frac{\partial H_\varepsilon(x, y)}{\partial y}\right)^2 \\
&\quad \le \frac{2\alpha}{3} H_\varepsilon(x, y)^{-1/3} L_\varepsilon H_\varepsilon(x, y) + \frac{4\alpha^2}{9}\varepsilon^2\sigma^2 x^2 H_\varepsilon(x, y)^{1/3} \\
&\quad \le \varepsilon^2 \frac{2\alpha}{9}(x^6 + y^2)^{2/3}\big(-3K_1^{-1/3} K_2 + 2\alpha\sigma^2 K_1^{1/3}\big).
\end{aligned}$$

Therefore if $\alpha < (3/2\sigma^2) K_2 K_1^{-2/3}$ we obtain the estimate

$$L_\varepsilon\big(\exp\big(\alpha H_\varepsilon^{2/3}\big)\big)(x, y) \le -\varepsilon^2 K_3 \exp\big(\alpha H_\varepsilon^{2/3}\big)(x, y)$$

for $x^2 + y^2 \ge k$ for some positive constant K_3. On the compact set $x^2 + y^2 \le k$ we have an estimate $L_\varepsilon(\exp(\alpha H_\varepsilon^{2/3}))(x, y) \le \varepsilon^2 K_4$ for some positive constant K. Together we get the estimate

$$L_\varepsilon\big(\exp\big(\alpha H_\varepsilon^{2/3}\big)\big)(x, y) \le -\varepsilon^2 K_3 \exp\big(\alpha H_\varepsilon^{2/3}\big)(x, y) + \varepsilon^2 K_4$$

valid on all of $\mathbf{R}^2$. It follows by a result of Meyn and Tweedie ([19], Theorem 4.3) that

$$\int \exp\big(\alpha H_\varepsilon^{2/3}\big)\, d\mu_\varepsilon \le \frac{K_4}{K_3} \qquad \text{for all } \varepsilon \le 1/\sqrt{\delta}. \tag{41}$$

Finally notice that $(1 + \varepsilon^2\delta b/3)x^4 \le 4H_\varepsilon(x, y)$ and

$$H(x, y) \le \frac{1}{2}x^2 + \frac{1}{4}x^4 + \left(y - \varepsilon^2(\beta + \delta)x + \frac{\varepsilon^2 b}{3}x^3\right)^2 + \varepsilon^4\left(\frac{b}{3}x^3 - (\beta + \delta)x\right)^2.$$

It follows that there are positive constants c_1 and c_2 such that $H(x, y) \le c_1 H_\varepsilon(x, y) + c_2 H_\varepsilon(x, y)^{3/2}$ for all $\varepsilon \le 1\sqrt{\delta}$. Therefore the estimate (41) implies one of the form

$$\int \exp\big(\alpha H^{4/9}\big)\, d\mu_\varepsilon \le K \qquad \text{for all } \varepsilon \le 1/\sqrt{\delta}$$

for a different positive α and a finite K. This completes the proof in the case $b > 0$.

We now consider the case when $b = 0$ and $\beta < 0$. We use the same function H_ε as above with $b = 0$. In this case we take $\delta = -2\beta/3 > 0$. In the calculation of $L_\varepsilon H_\varepsilon(x, y)$ the x^6 term is missing and the highest order term in x is $\varepsilon^2(\beta + \delta)x^4$. The estimates on H_ε and $L_\varepsilon H_\varepsilon$ can be done in terms of $H(x, y)$ rather than $x^6 + y^2$. We replace the function $\exp(\alpha H_\varepsilon^{2/3})$ by the function $\exp(\alpha\sqrt{H_\varepsilon})$ and use the result of Meyn and Tweedie to obtain an estimate of the form

$$\int \exp\bigl(\alpha\sqrt{H_\varepsilon}\bigr)\, d\mu_\varepsilon \le \frac{K_4}{K_3} \qquad \text{for all } \varepsilon \le 1/\sqrt{\delta}.$$

Finally we use the fact that $H_\varepsilon(x, y)/H(x, y) \to 1$ as $\varepsilon \to 0$ uniformly for $(x, y) \in \mathbf{R}^2$ to obtain a stronger version of the lemma where the power $4/9$ is replaced by $1/2$. $\square$

LEMMA 8. *For every $p, q > 0$ and $\varepsilon > 0$ there is $k < \infty$, depending on p, q and ε, such that the function $F(x, y) = H(x, y)^p + H(x, y)^{-q}$ has the property $L_\varepsilon F(x, y) \le kF(x, y)$ for all $(x, y) \in \mathbf{R}^2$.*

PROOF. By direct calculation. $\square$

3.2. *Averaging around orbits; explicit formulas.* The following calculations extend those contained in Liang and Sri Namachchivaya [18]. We use the Hamiltonian function $H(x, y) = x^2/2 + y^2/2 + x^4/4$, and evaluate integrals of certain functions with respect to the probability measure m_h along the orbit $H(x, y) = h$. Throughout we shall assume $h > 0$.

Define $Q(x, h) = 2(h - x^2/2 - x^4/4)$. Then along the top half of the orbit $H(x, y) = h$ the Hamiltonian flow $\Phi_t(x, y)$ satisfies $dx/dt = \sqrt{Q(x, h)}$. Moreover the orbit meets the x-axis at the points $x = \pm x_h$ where $x_h = \sqrt{4h + 1} - 1$. For any function $f(x, y)$ satisfying $f(x, y) = f(-x, -y)$ we have

$$\begin{aligned}\int f(x, y)\, dm_h(x, y) &= \frac{1}{T(h)} \int_0^{T(h)} f\bigl(\Phi_t(x, y)\bigr)\, dt \\ &= \frac{2}{T(h)} \int_{-x_h}^{x_h} \frac{f(x, \sqrt{Q(x, h)})}{\sqrt{Q(x, h)}}\, dx.\end{aligned}$$

The following result can be proved using elementary techniques of integration.

LEMMA 9. *Write*

$$I_n(h) = \int_{-x_h}^{x_h} \frac{x^n}{\sqrt{Q(x, h)}}\, dx.$$

Then

$$I_0(h) = \frac{2}{(1+4h)^{1/4}} K(m), \tag{42}$$

$$I_2(h) = \left[(1+4h)^{1/2}\left(\frac{2E(m)}{K(m)} - 1\right) - 1\right] I_0(h), \tag{43}$$

$$I_{2n+4}(h) = -\frac{4n+4}{2n+3} I_{2n+2}(h) + \frac{8n+4}{2n+3} h I_{2n}(h) \qquad \text{if } n \geq 0, \tag{44}$$

where

$$m^2 = \frac{1}{2}\left(1 - \frac{1}{\sqrt{4h+1}}\right)$$

and $K(m)$ *and* $E(m)$ *denote the complete elliptic integrals of the first and second kind* (*see* [25], *Section* 22.7).

COROLLARY 3. *For the system* (39) *the condition* (A6) *is satisfied and*

$$G_1(h) = -\frac{I_0'(h)}{I_0(h)},$$

$$G_2(h) = \sigma^2\left[\frac{32}{7}h^2 + \frac{16}{21}h + \frac{I_2(h)}{I_0(h)}\left(-\frac{24}{7}h - \frac{16}{21}\right)\right],$$

$$G(h) = |G_1(h)|^{2/3}[G_2(h)]^{1/3}, \tag{45}$$

$$\bar{c}(h) = \left(\frac{4\beta}{3} + \frac{4b}{15}\right)h + \frac{I_2(h)}{I_0(h)}\left(\frac{\sigma^2}{2} - \frac{\beta}{3} - \frac{4b}{15} - \frac{8b}{15}h\right), \tag{46}$$

$$\bar{d}(h) = \sigma^2\left[-\frac{2}{15}h + \frac{I_2(h)}{I_0(h)}\left(\frac{2}{5}h + \frac{2}{15}\right)\right], \tag{47}$$

where the function $I_0(h)$ *is given by* (42) *and the ratio* $I_2(h)/I_0(h)$ *is given by* (43).

PROOF. Since $T(h) = 2I_0(h) = 4(1-2m^2)^{1/2}K(m)$ we obtain

$$\begin{aligned}\frac{T'(h)}{T(h)} &= \left[\frac{K'(m)}{K(m)} - \frac{2m}{1-2m^2}\right]\frac{dm}{dh} \\ &= \left[\frac{E(m)}{m(1-m^2)K(m)} - \frac{1}{m(1-2m^2)}\right]\frac{dm}{dh}\end{aligned}$$

(see [25], Section 22.736). Since $E(m) < K(m)$ for $m \neq 0$ it follows that $T'(h) < 0$ for all $h > 0$ and hence (A6) is satisfied. The formula for $G_1(h)$ is immediate from the equality $T_0(h) = 2I_0(h)$. Now

$$\begin{aligned} G_2(h) &= \sigma^2 \int \left(y^2 - x^2 - x^4\right)^2 dm_h(x, y) \\ &= \sigma^2 \int \left(2h - 2x^2 - \frac{3x^4}{2}\right)^2 dm_h(x, y) \\ &= \sigma^2 \int \left(4h^2 - 8hx^2 + (4-6h)x^4 + 6x^6 + \frac{9x^8}{4}\right) dm_h(x, y) \\ &= \sigma^2 \left(4h^2 - 8h\frac{I_2(h)}{I_0(h)} + (4-6h)\frac{I_4(h)}{I_0(h)} + 6\frac{I_6(h)}{I_0(h)} + \frac{9I_8(h)}{4I_0(h)}\right) \end{aligned}$$

and the calculation is completed by using the reduction formula (44) to obtain expressions for $I_4(h)$ and $I_6(h)$ and $I_8(h)$ in terms of $I_2(h)$ and $I_0(h)$. The calculations for $\overline{c}(h)$ and $\overline{d}(h)$ are similar. □

We can use information about the behavior of the complete elliptic integrals $K(m)$ and $E(m)$ for m near 0 and $1/\sqrt{2}$ to obtain information about $\Gamma(h) = [G_1(h)/G_2(h)]^{1/3}$ and $G(h)$ and $\overline{c}(h)$ and $\overline{d}(h)$ for h near 0 and ∞. The following result summarizes an important intermediate step.

LEMMA 10. *Write $G_3(h) = I_2(h)/I_0(h)$ and recall $G_1(h) = -I_0'(h)/I_0(h)$.*

(i) *$G_1(h)$ and $G_3(h)$ are analytic functions of h in a neighborhood of 0. As $h \to 0$ we have*

$$G_1(h) = \frac{3}{4} - \frac{87h}{32} + \frac{657h^2}{64} + O(h^3),$$

$$G_3(h) = h - \frac{9h}{8} + \frac{39h^2}{16} + O(h^3).$$

(ii) *As $h \to \infty$ we have*

$$G_1(h) \sim \frac{1}{4h}; \qquad G_1'(h) \sim -\frac{1}{4h^2}; \qquad G_1''(h) \sim \frac{1}{2h^3};$$

and

$$G_3(h) \sim 2\gamma_1 h^{1/2}; \qquad G_3'(h) \sim \gamma_1 h^{-1/2}; \qquad G_3''(h) \sim -\frac{\gamma_1}{2} h^{-3/2};$$

where $\gamma_1 = 2E(1/\sqrt{2})/K(1/\sqrt{2}) - 1 \approx 0.457$.

COROLLARY 4. *As $h \to 0$,*

$$\begin{aligned} \Gamma(h) &\sim (3/8\sigma^2)^{1/3} h^{-2/3}, \\ \Gamma'(h) &\sim -(2/3)(3/8\sigma^2)^{1/3} h^{-5/3}, \\ \Gamma''(h) &\sim (10/9)(3/8\sigma^2)^{1/3} h^{-8/3}; \end{aligned}$$

and as $h \to \infty$,

$$\Gamma(h) \sim (7/128\sigma^2)^{1/3} h^{-1},$$
$$\Gamma'(h) \sim (7/128\sigma^2)^{1/3} h^{-2},$$
$$\Gamma''(h) \sim 2(7/128\sigma^2)^{1/3} h^{-3}.$$

3.3. *Derivatives of A, B, C and D.* The functions f_ε and ϕ_k involve the functions A, B, C and D, together with some of their derivatives. Consider first the function $A(x, y)$ required to be a solution of $U_1.A(x, y) = a(x, y) - \int a\, dm_h$ along the orbit $H(x, y) = h$. In order to verify conditions (A10)–(A12) we need estimates on $A(x, y)$ and $V_1.A(x, y)$ and $\overline{L}A(x, y)$, together with similar estimates on B, C and D. Our first result is valid for a general Hamiltonian system with a single critical point at $(0, 0)$.

LEMMA 11. *Assume* (A5). *Let* $\mathfrak{D}$ *denote an integral curve of the vector field* U_2 *from 0 to* ∞, *and let* f *be a smooth function on* $\mathbf{R}^2 \setminus \{0\}$. *Write* $\tilde{f}(x, y) = f(x, y) - \int f\, dm_h$ *for* $(x, y) \in H^{-1}(h)$. *Then there is a unique function* F *on* $\mathbf{R}^2 \setminus \{0\}$ *such that* $U_1.F(x, y) = \tilde{f}(x, y)$ *and* $F|_{\mathfrak{D}} \equiv 0$. *Moreover*, *we have the equality*

$$U_1.(U_2.F)(x, y) = U_2.\tilde{f}(x, y) - J(x, y)\tilde{f}(x, y) \tag{48}$$

and the inequalities

$$|F(x, y)| \le T(h) \sup\{|\tilde{f}(u, v)| : (u, v) \in H^{-1}(h)\}, \tag{49}$$

$$|(U_2.F)(x, y)| \le T(h) \sup\{|(U_2 - J).\tilde{f}(u, v)| : (u, v) \in H^{-1}(h)\}, \tag{50}$$

$$|(U_2^2.F)(x, y)| \le T(h) \sup\{|(U_2 - J)^2.\tilde{f}(u, v)| : (u, v) \in H^{-1}(h)\}, \tag{51}$$

valid for $(x, y) \in H^{-1}(h)$.

PROOF. For each $h > 0$ parameterize the orbit $H^{-1}(h)$ by the path $p(t)$ where $p(0) = p(T(h)) \in \mathfrak{D}$ and $p'(t) = U_1(p(t))$. Then we can take

$$F(p(t)) = \int_0^t \tilde{f}(p(s))\, ds$$

and clearly this is the unique solution of $U_1.F = \tilde{f}$ with $F|_{\mathfrak{D}} \equiv 0$. The first inequality (49) follows directly. The equality (48) is a direct application of Lemma 1. Since $F|_{\mathfrak{D}} \equiv 0$ then $F_2.A|_{\mathfrak{D}} \equiv 0$ and so we can apply the previous estimate with F and $\tilde{f}$ replaced by $U_2.F$ and $(U_2 - J).\tilde{f}$ to obtain the second estimate (50). Repeating this idea gives the third estimate (51). $\square$

LEMMA 12. *For the system* (39), *suppose* $U_1.F(x, y) = \tilde{f}(x, y)$ *as above*, *and that* $F(0, y) \equiv 0$. *For ease of notation write* $\|f\|_h = \sup\{|f(u, v)| : H(u, v) = h\}$.

(i) *There are constants* k *and* δ *such that if* $r = \sqrt{x^2+y^2} < \delta$ *and* $H(x,y) = h$ *then*

$$\begin{aligned} |F(x,y)| &\le k\|f\|_h, \\ |V_1.F(x,y)| &\le k\big[\|f\|_h + r^2\|U_2.f\|_h\big], \\ \big|\overline{L}F(x,y)\big| &\le k\big[\|f\|_h + \|V_1.f\| + r^2\|U_2.f\| + r^4\|U_2^2.f\|_h\big]. \end{aligned}$$

(ii) *If* f *and its first two derivatives have at most polynomial growth as* $(x,y)\to\infty$, *then so do* F *and* $V_1.F$ *and* $\overline{L}F$.

PROOF. This is a direct application of the estimates in the Lemma. In (i) we use the facts that $T(h)$ and the functions $\alpha_0^1, \alpha_1^1, V_1.\alpha_1^1$, J and $U_2.J$ are all bounded near $(0,0)$ and that α_1^2 and $V_1.\alpha_1^2$ are both of order r^2 near $(0,0)$. In (ii) we use the fact that all of the functions just mentioned have at most polynomial growth. □

COROLLARY 5. *For the functions* A, B, C *and* D *associated with the system* (39):

(i) A *and* $V_1.A$ *and* $\overline{L}A$ *and* B *and* $V_1.B$ *and* $\overline{L}B$ *are all* $O(r^{4/3})$ *as* $(x,y)\to(0,0)$ *and all have at most polynomial growth as* $(x,y)\to\infty$.

(ii) C *and* $V_1.C$ *and* $\overline{L}C$ *are all* $O(r^2)$ *as* $(x,y)\to(0,0)$ *and all have at most polynomial growth as* $(x,y)\to\infty$.

(iii) D *and* $V_1.D$ *and* $\overline{L}D$ *are all* $O(r^4)$ *as* $(x,y)\to(0,0)$ *and all have at most polynomial growth as* $(x,y)\to\infty$.

PROOF. This is a direct application of the previous result. In case (i) we use the estimates on Γ and its derivatives given in Subsection 3.2. □

3.4. *Existence and derivatives of* Ψ. We now turn attention to the diffusion on $(0,\infty)$ with generator $\mathcal{N}$ and the functions G and Ψ. The coefficients $\overline{c}(h)$ and $\overline{d}(h)$ were given in (46), (47). They yield the following asymptotics as h tends to 0 and ∞:

$$\overline{c}(h) \sim \left(\beta + \frac{\sigma^2}{2}\right)h \qquad \text{as } h\to 0,$$

$$\overline{d}(h) \sim \frac{\sigma^2}{4}h \qquad \text{as } h\to 0,$$

$$\overline{c}(h) \sim -\frac{8\gamma_1 b}{15}h^{3/2} \qquad \text{as } h\to\infty \text{ if } b>0,$$

$$\overline{c}(h) \sim \frac{4\beta}{3}h \qquad \text{as } h\to\infty \text{ if } b=0 \text{ and } \beta<0,$$

$$\overline{d}(h) \sim \frac{4\gamma_1\sigma^2}{5}h^{3/2} \qquad \text{as } h\to\infty.$$

It follows easily that there is a C^2 function F: $(0,\infty) \to (0,\infty)$ of the form $F(h) = h^{-c_1}$ for h sufficiently small and $F(h) = \exp c_2\sqrt{h}$ for h sufficiently large with the property that $\mathcal{N} F(h) \le -c_3 F(h) + c_4$. Here c_1, c_2, c_3 and c_4 are positive constants; this construction can be done for all sufficiently small c_1 and c_2.

Condition (A8) on the existence of the invariant probability ρ follows immediately (using results of Khas'minskii [17], Theorem III.7.3 and Theorem IV.4.1). Moreover, by a result of Meyn and Tweedie ([19], Theorem 4.2), the invariant probability ρ has the property $\int_{(0,\infty)} F(h)\, d\rho(h) < \infty$. From the calculations in Subsection 3.2 we have the following asymptotics for G

$$G(h) \sim (9\sigma^2/8)^{1/3} h^{2/3} \qquad \text{as } h \to 0,$$
$$G(h) \sim (2\sigma^2/7)^{1/3} \qquad \text{as } h \to \infty.$$

Since G/F is bounded on $(0,\infty)$ it follows that $\int_{(0,\infty)} G(h)\, d\rho(h) < \infty$. Now the estimate $\mathcal{N} F(h) \le -c_3 F(h) + c_4$ implies that $\mathcal{N}$ is the generator of an F-exponentially ergodic process (see [19], Theorem 6.1). It follows that there exists a smooth function Ψ such that $\mathcal{N}\Psi(h) = G(h) - \int G\, d\rho$ and $|\Psi(h)|/F(h)$ is bounded. Thus (A9) is satisfied. More details of this sequence of arguments is given in Baxendale and Goukasian [9], Proposition 4 and Lemma 5. At this stage we have pointwise estimates on Ψ coming from the fact that $|\Psi(h)|/F(h)$ is bounded. But when we look at the formulas for f_ε and ϕ_2 we see that we will need better pointwise estimates on Ψ and also estimates on its first 4 derivatives. To obtain these we need the following result, which may be of independent interest

LEMMA 13. *Let $\{x_t : t \ge 0\}$ be a diffusion process, with generator L on a σ-compact space M. Suppose that there is a compact $K \subset M$ and a function $F \ge 1$ so that $LF \le -cF$ outside K for some $c > 0$. Define $\tau_K = \inf\{t \ge 0 : x_t \in K\}$. Then $P^x(\tau_K < \infty) = 1$ for all $x \notin K$. Suppose further there exist functions f and g so that $Lf = g$ outside K and $|g(x)|/F(x)$ is bounded on M and $f(x)/F(x) \to 0$ as $x \to \infty$. Then*

$$E^x f(x_{\tau_K}) - f(x) = E^x \int_0^{\tau_K} g(x_s)\, ds \qquad \textit{for all } x \notin K.$$

PROOF. The first assertion is due to Khas'minskii [17], Theorem III.7.1. To prove the second assertion we use the assumption about F to show that the local martingale $f(x_{t\wedge\tau_K}) - \int_0^{t\wedge\tau_K} g(x_s)\, ds$ is a uniformly integrable martingale. Define $\tau_n = \inf\{t \ge 0 : x_t \notin K_n\}$ where the compact sets K_n satisfy $K = K_0 \subset K_1 \subset K_2 \subset \cdots \nearrow N$. The condition $Lf = g$ implies that

$$E^x f(x_{t\wedge\tau_n\wedge\tau_K}) - f(x) = E^x \int_0^{t\wedge\tau_n\wedge\tau_K} g(x_s)\, ds \tag{52}$$

for all $x \notin K$. The assumption on F implies that $e^{c(t\wedge\tau_n\wedge\tau_K)}F(x_{t\wedge\tau_n\wedge\tau_K})$ and $F(x_{t\wedge\tau_n\wedge\tau_K}) + c\int_0^{t\wedge\tau_n\wedge\tau_K} F(x_s)\,ds$ are supermartingales. These yield the inequalities

$$E^x F(x_{t\wedge\tau_n\wedge\tau_K}) \le F(x)$$

and

$$E^x \int_0^{t\wedge\tau_n\wedge\tau_K} F(x_s)\,ds \le \frac{1}{c} f(x)$$

for $x \notin K$. The first inequality implies that the random variables $f(x_{t\wedge\tau_n\wedge\tau_K})$ are uniformly integrable with respect to P^x and so we can pass to the limit as $t \to \infty$ and $n \to \infty$ in the left side of (52). The second inequality allows the use of the dominated convergence theorem on the right side of (52) as $t \to \infty$ and $n \to \infty$. $\square$

We now specialize to the diffusion process with generator $\mathcal{N}$ on $(0,\infty)$. Taking one function F in the exponential ergodicity argument above to get growth estimates on Ψ, and then a different F with larger c_1 and c_2 here, we see that we can apply Lemma 13 with $f(x) = \Psi(x)$ and $g(x) = G(x) - \int G\,d\rho \equiv \widetilde{G}(x)$. Take $K = [\delta, k]$ for sufficiently small δ and large k, and $x < \delta$. We have

$$\Psi(\delta) - \Psi(x) = E^x \int_0^{\tau_\delta} \widetilde{G}(x_s)\,ds.$$

Replacing δ by an arbitrary point $y \le \delta$ we get

$$\Psi(y) - \Psi(x) = E^x \int_0^{\tau_y} \widetilde{G}(x_s)\,ds$$

whenever $0 < x < y \le \delta$. Replacing Ψ in this calculation by the function $f(x) = \log x$ we get

$$\log y - \log x = E^x \int_0^{\tau_y} \mathcal{N} f(x_s)\,ds.$$

Since $\mathcal{N} f(x) = x^{-1}\overline{c}(x) - x^{-2}\overline{d}(x) \to \beta + \sigma^2/4 = 2\lambda > 0$ as $x \to 0$ there exist $\delta_1 > 0$ and $k_1 < \infty$ such that $|\widetilde{G}(x)| \le k_1 \mathcal{N} f(x)$ whenever $0 < x < \delta_1$. Then

$$|\Psi(y) - \Psi(x)| \le E^x \int_0^{\tau_y} \bigl|\widetilde{G}(x_s)\bigr|\,ds \le k E^x \int_0^{\tau_y} \mathcal{N} f(x_s)\,ds = k_1(\log y - \log x)$$

whenever $0 < x < y \le \min(\delta, \delta_1) \equiv \delta_2$. Taking $y = \delta_2$ we get the improved pointwise estimate $|\Psi(x)| \le |\Psi(\delta_2)| + k_1 \log\delta_2 - k_1 \log x$ for $0 < x < \delta_2$. More importantly, letting $y \searrow x$ we get $|\Psi'(x)| \le k_1 x^{-1}$ for $0 < x < \delta_2$. Now substituting into

$$\overline{c}(x)\Psi'(x) + \overline{d}(x)\Psi''(x) = G(x) - \int G\,d\sigma$$

and using the asymptotics of $\overline{c}(x)$ and $\overline{d}(x)$ and $G(x)$ near 0 gives $|\Psi''(x)| \leq k_2 x^{-2}$ near 0. Differentiating the expression above (remembering that $\overline{c}$ and $\overline{d}$ and G^3 are real analytic functions in a neighborhood of 0 and using Corollary 3 and Lemma 10) we get consecutively $|\Psi'''(x)| \leq k_3 x^{-3}$ and then $|\Psi''''(x)| \leq k_4 x^{-4}$ near 0.

We can repeat this method near ∞. For sufficiently large k we have

$$\Psi(y) - \Psi(x) = E^x \int_0^{\tau_y} \widetilde{G}(x_s)\, ds$$

whenever $k \leq y < x$. If $b = 0$ and $\beta < 0$ we use the function $f(x) = -\log x$, and if $b > 0$ we use the function $f(x) = x^{-1/2}$. We use the formulas for $\overline{c}$ and $\overline{d}$ and G from Corollary 3 together with the asymptotics for G_1 and G_3 and their derivatives near ∞ from Lemma 4. In the first case we get estimates of the form $|\Psi(x)| \leq k_0 \log x$, $|\Psi'(x)| \leq k_1 x^{-1}$, $|\Psi''(x)| \leq k_2 x^{-3/2}$, $|\Psi'''(x)| \leq k_3 x^{-2}$ and $|\Psi''''(x)| \leq k_4 x^{-5/2}$ for x large. In the second case we get estimates of the form $|\Psi(x)| \leq k_0$ and $|\Psi^{(j)}(x)| \leq k_j x^{-3/2}$, $1 \leq j \leq 4$, for x large.

3.5. *Proof of Theorem* 2. Formulas for $\overline{c}(h)$ and $\overline{d}(h)$ and $G(h)$ have been obtained in Corollary 3. The conditions (A1) and (A5)–(A9) have been verified earlier in this section. The estimates on Γ in Corollary 4, together with Lemmas 6 and 7 imply that (A2) is satisfied. The estimates on A, B, C, D and Ψ and their derivatives from Subsections (3.3) and (3.4) imply that $f_\varepsilon(x, y)$ grows at most like $|\log r|$ as $(x, y) \to (0, 0)$ and has at most polynomial growth in x and y as $(x, y) \to \infty$. It follows from Lemmas 6 and 7 that f_ε is integrable with respect to P_ε.

Using the estimates on Γ given in Corollary 4, together with explicit formulas for the $\alpha_i^j(x, y)$ and $J(x, y)$, it is a routine (but lengthy) calculation to show that each of the entries in $M_0(x, y)$ and $M_1(x, y)$ and $DM_1(x, y)(V_1(x, y))$ remains bounded as $(x, y) \to (0, 0)$ and has at most polynomial growth as $(x, y) \to \infty$. This implies, using Lemma 4, that each of the $Q_k(x, y, \theta)$ and $t_k(x, y, \theta)$ remains bounded as $(x, y) \to (0, 0)$ and has at most polynomial growth as $(x, y) \to \infty$. Conditions (A3) and (A4) now follow immediately using Lemma 7.

Using the results on the t_k from the previous paragraph together with the estimates on the functions A, B, C, D and Ψ and their derivatives from Subsections (3.3) and (3.4), it is now easy to verify that each of the functions $\phi_0, \ldots, \phi_4$ (given in the Appendix) remain bounded as $(x, y) \to (0, 0)$ and have at most polynomial growth as $(x, y) \to \infty$. Notice that for the function ϕ_2 near 0, the estimates on C and D obtained in Subsection 3.3 together with obvious estimates on c and d and α_1^2 exactly counteract the possible growth in the derivatives of Ψ given in Subsection 3.4. Conditions (A10) and (A11) now follow using Lemma 7 and the condition (A12) follows using Lemma 8.

REMARK 5. Suppose that in equation (39) the multiplicative noise $\varepsilon\sigma x\,dW_t$ is replaced by additive noise $\varepsilon\sigma\,dW_t$. This has the effect of replacing the functions α_1^1 and α_1^2 by new functions $\alpha_1^1(x,y) = -\sigma(x+x^3)/[(x+x^3)^2+y^2]$ and $\alpha_1^2(x,y) = -\sigma x$. Then the four functions $U_1.\alpha_1^1$, $U_1.\alpha_1^2$, $U_2.\alpha_1^1$ and $U_2.\alpha_1^2$ are of order r^{-1}, r, r^{-3} and r^{-1} respectively near 0. Since the probability measure μ_ε is independent of ε (for additive noise) and has a smooth positive density, it follows that the entries in the matrices $M_1(x,y)$ and $(\partial/\partial y)M_1(x,y)$ fail to be integrable, and the entire method fails.

3.6. *Convergence of measures.* Here we continue the discussion started in Remark 3 concerning the convergence of the probability measures P_ε and μ_ε. For convenience we revert to notation used in Section 2 and let x denote a point in $\mathbf{R}^2$. The following result extends Theorem 3.

THEOREM 5. *Let $g(x)$ be a smooth function with compact support in $\mathbf{R}^2\setminus\{0\}$ and let $s(\theta)$ be a smooth function on S^1. Then*

$$\left|\int g(x)s(\theta)\,dP_\varepsilon(x,\theta) - \int g(x)s(\theta)\,d\mu_0(x)\,d\nu(\theta)\right| = O(\varepsilon^{2/3}) \qquad \text{as } \varepsilon\to 0$$

and

$$\left|\int g(x)\,d\mu_\varepsilon(x) - \int g(x)\,d\mu_0(x)\right| = O(\varepsilon^2) \qquad \text{as } \varepsilon\to 0.$$

In particular $P_\varepsilon \to \mu_0\times\nu$ weakly as probability measures on $\mathbf{R}^2\times S^1$, and $\mu_\varepsilon\to\mu_0$ as $\varepsilon\to 0$ weakly as probability measures on $\mathbf{R}^2$.

PROOF. Define $\overline{g}(h) = \int g(x)\,dm_h(x)$ and choose a function $E(x)$ so that $U_1.E(x) = g(x) - \overline{g}(H(x))$. Define $\overline{s} = \int s(\theta)\,d\nu(\theta)$ and choose a function $\widetilde{R}(\theta)$ so that $\overline{M}\widetilde{R}(\theta) = s(\theta) - \overline{s}$. Choose a function $\widetilde{\Psi}(h)$ so that $\mathcal{N}\widetilde{\Psi}(\theta) = \overline{g}(h) - \int\overline{g}\,d\rho$. The results in Subsection 3.4 on the existence of Ψ and the estimates on Ψ and its first four derivatives apply equally well to the function $\widetilde{\Psi}$. Now define $\tilde{f}_\varepsilon(x,\theta)$ by

$$\begin{aligned}\tilde{f}_\varepsilon(x,\theta) = E(x)s(\theta) &+ \varepsilon^{-2/3}\frac{\overline{g}(H(x))}{G(H(x))}\widetilde{R}(\theta) - \frac{\overline{g}(H(x))}{G(H(x))}A(x)\sin^2\theta\,\widetilde{R}'(\theta)\\ &- \frac{\overline{g}(H(x))}{G(H(x))}B(x)\Big(-\cos^3\theta\sin\theta\,\widetilde{R}'(\theta) + \frac{1}{2}\cos^4\theta\,\widetilde{R}''(\theta)\Big)\\ &+ \varepsilon^{-2}\overline{s}\big(\widetilde{\Psi}\circ H\big)(x) - \overline{s}\big[C(x)\big(\widetilde{\Psi}'\circ H\big)(x) + D(x)\big(\widetilde{\Psi}''\circ H\big)(x)\big].\end{aligned}$$

Then a direct calculation shows that

$$\begin{aligned}(\mathcal{L}_\varepsilon\tilde{f}_\varepsilon)(x,\theta) &= g(x)s(\theta) - \int\overline{g}(h)d\rho(h)\int s(\theta)\,d\nu(\theta)\\ &\quad + \varepsilon^{2/3}\sum_{k=0}^{4}\varepsilon^{2k/3}\tilde{\phi}_k(x,\theta)\end{aligned} \tag{53}$$

for suitable functions $\tilde{\phi}_k$. Now the assumption that g has compact support in $\mathbf{R}^2 \setminus \{0\}$ implies that $E(x)$ and $\overline{g}(H(x))$ are zero near 0 and ∞. Thus when we are checking integrability and growth conditions for the functions $\tilde{f}_\varepsilon$ and $\tilde{\phi}_k$ the only terms which can be non-zero near 0 or ∞ are ones involving the function $\widetilde{\Psi}$ and its derivatives. But $\widetilde{\Psi}$ satisfies the same estimates as Ψ, and so the calculations used to verify conditions (A10)–(A12) in the preceding subsection can be used here also. The rest of the argument used in Theorem 4 is unchanged, and we obtain the first estimate.

In the case when $s(\theta) \equiv 1$ then we can take $\widetilde{R}(\theta) \equiv 0$ and the argument proceeds as before, but in a simplified form. The $\tilde{\phi}_k$ terms vanish for $k = 0$, 1, 3 and 4, so that the estimate is of order ε^2.

Finally notice that Lemma 6 shows that $\{\mu_\varepsilon : 0 < \varepsilon \le \varepsilon_0\}$ and $\{P_\varepsilon : 0 < \varepsilon \le \varepsilon_0\}$ are both tight families of probability measures on $\mathbf{R}^2$ and $\mathbf{R}^2 \times S^1$, respectively. The estimates above prove uniqueness for any subsequential limit, so we obtain the statements about weak convergence of P_ε to $\mu_0 \times S^1$ and μ_ε to μ_0. $\square$

APPENDIX

Define

$$R_1(\theta) = \sin\theta\cos\theta - \sin^2\theta R'(\theta),$$
$$R_2(\theta) = \tfrac{1}{2}\cos^2\theta - \sin^2\theta\cos^2\theta + \sin\theta\cos^3\theta R'(\theta) - \tfrac{1}{2}\cos^4\theta R''(\theta).$$

Then

$$\begin{aligned}\phi_0(x,\theta) &= t_1(x,\theta)R'(\theta) + t_0(x,\theta)\big[A(x)R_1'(\theta) + B(x)R_2'(\theta)\big]\\ &\quad + \tfrac{1}{2}\sum_{i=1}^{r}\big[t_5^i(x,\theta)\big]^2\big[A(x)R_1''(\theta) + B(x)R_2''(\theta)\big];\end{aligned}$$

$$\begin{aligned}\phi_1(x,\theta) &= t_2(x,\theta)R'(\theta) + \tfrac{1}{2}\sum_{i=1}^{r}\big[t_6^i(x,\theta)\big]^2 R''(\theta)\\ &\quad + t_1(x,\theta)\big[A(x)R_1'(\theta) + B(x)R_2'(\theta)\big]\\ &\quad + \tfrac{1}{2}\sum_{i=1}^{r} t_5^i(x,\theta)\big[V_i.A(x)R_1'(\theta) + V_i.B(x)R_2'(\theta)\big];\end{aligned}$$

$$\begin{aligned}\phi_2(x,\theta) &= t_3(x,\theta)R'(\theta) + \overline{L}A(x)R_1(\theta) + \overline{L}B(x)R_2(\theta)\\ &\quad + t_2(x,\theta)\big[A(x)R_1'(\theta) + B(x)R_2'(\theta)\big]\\ &\quad + \tfrac{1}{2}\sum_{i=1}^{r}\big[t_6^i(x,\theta)\big]^2\big[A(x)R_1''(\theta) + B(x)R_2''(\theta)\big]\end{aligned}$$

$$
\begin{aligned}
&+ \sum_{i=1}^{r} t_6^i(x,\theta)\big[V_i.A(x)R_1'(\theta) + V_i.B(x)R_2'(\theta)\big]\\
&- \gamma_0\Bigg[\overline{L}C(x)\Psi'\big(H(x)\big) + \Bigg(\sum_{i=1}^{r}\alpha_i^2(x)V_i.C(x) + C(x)c(x)\Bigg)\Psi''\big(H(x)\big)\\
&\qquad\qquad + C(x)d(x)\Psi'''\big(H(x)\big)\Bigg]\\
&- \gamma_0\Bigg[\overline{L}D(x)\Psi''\big(H(x)\big)\\
&\qquad + \Bigg(\sum_{i=1}^{r}\alpha_i^2(x)V_i.D(x) + D(x)c(x)\Bigg)\Psi'''\big(H(x)\big)\\
&\qquad\qquad + D(x)d(x)\Psi''''\big(H(x)\big)\Bigg];
\end{aligned}
$$

$$
\begin{aligned}
\phi_3(x,\theta) &= t_4(x,\theta)R'(\theta) + \tfrac{1}{2}\sum_{i=1}^{r}\big[t_7^i(x,\theta)\big]^2R''(\theta)\\
&\quad + t_3(x,\theta)\big[A(x)R_1'(\theta) + B(x)R_2'(\theta)\big]\\
&\quad + \sum_{i=1}^{r} t_7^i(x,\theta)\big[V_i.A(x)R_1'(\theta) + V_i.B(x)R_2'(\theta)\big];\\
\phi_4(x,\theta) &= t_4(x,\theta)\big[A(x)R_1'(\theta) + B(x)R_2'(\theta)\big]\\
&\quad + \tfrac{1}{2}\sum_{i=1}^{r}\big[t_7^i(x,\theta)\big]^2\big[A(x)R_1''(\theta) + B(x)R_2''(\theta)\big].
\end{aligned}
$$

Acknowledgments. We would like to thank Richard Sowers, Peter Imkeller and N. Sri Namachchivaya for helpful and stimulating discussions during the preparation of this paper. We are also grateful to Sowers for sending us an early copy of his paper [23] which also uses the Pinsky–Wihstutz transformation in a similar setting; and to Imkeller and Sri Namachchivaya for showing us preliminary calculations dealing with perturbations of a Hamiltonian system by additive white noise.

REFERENCES

[1] ARIARATNAM, S. T. and XIE, W. C. (1990). Lyapunov exponent and rotation number of a two-dimensional nilpotent stochastic system. *Dynam. Stability Systems* **5** 1–9.

[2] ARNOLD, L. (1998). *Random Dynamical Systems.* Springer, Berlin.

[3] ARNOLD, L., PAPANICOLAOU, G. and WIHSTUTZ, V. (1986). Asymptotic analysis of the Lyapunov exponent and rotation number of the random oscillator and applications. *SIAM J. Appl. Math.* **46** 427–450.

[4] ARNOLD, L., SRI NAMACHCHIVAYA, N. and SCHENK–HOPPÉ, K. (1996). Toward an understanding of the stochastic Hopf bifurcation: a case study. *Internat. J. Bifur. Chaos Appl. Sci. Engrg.* **6** 1947–1975.

[5] AUSLENDER, E. and MIL'SHTEIN, G. (1982). Asymptotic expansions of the Liapunov index for linear stochastic systems with small noise. *J. Appl. Math. Mech.* **46** 277–283.

[6] BAXENDALE, P. H. (1991). Statistical equilibrium and two-point motion for a stochastic flow of diffeomorphisms. In *Spatial Stochastic Processes* (K. Alexander and J. Watkins, eds.) 189–218. Birkhäuser, Boston.

[7] BAXENDALE, P. H. (1994). A stochastic Hopf bifurcation. *Probab. Theory Related Fields* **99** 581–616.

[8] BAXENDALE, P. H. (1999). Stability along trajectories at a stochastic bifurcation point. In *Stochastic Dynamics* (H. Crauel and M. Gundlach, eds.) 1–25. Springer, Berlin.

[9] BAXENDALE, P. H. and GOUKASIAN, L (2001). Lyapunov exponents of nilpotent Itô systems with random coefficients. *Stochastic Process. Appl.* **95** 219–233.

[10] BAXENDALE, P. H. and STROOCK, D. W. (1988). Large deviations and stochastic flows of diffeomorphisms. *Probab. Theory Related Fields* **80** 169–215.

[11] FREIDLIN, M. and WEBER, M. (1998). Random perturbations of nonlinear oscillators. *Ann. Probab.* **26** 925–967.

[12] FREIDLIN, M. and WEBER, M. (1999). A remark on random perturbations of the nonlinear pendulum. *Ann. Appl. Probab.* **9** 611–628.

[13] FREIDLIN, M. and WENTZELL, A. (1994). *Random Perturbations of Hamiltonian Systems.* Amer. Math Soc., Providence, RI.

[14] IMKELLER, P. and LEDERER, C. (1999). An explicit description of the Lyapunov exponents of the noisy damped harmonic oscillator. *Dynam. Stability Systems* **14** 385–405.

[15] KELLER, H. and OCHS, G. (1999). Numerical approximation of random attractors. In *Stochastic Dynamics* (H. Crauel and M. Gundlach, eds.) 93–115. Springer, Berlin.

[16] KHAS'MINSKII, R. Z. (1964). The behavior of a conservative system under the action of slight friction and slight random noise. *J. Appl. Math. Mech.* **28** 1126–1130.

[17] KHAS'MINSKII, R. Z. (1980). *Stochastic Stability of Differential Equations*. Sijthoff and Noordhoff, Alphen aan den Rijn.

[18] LIANG, Y. and SRI NAMACHCHIVAYA, N. (1999). P-bifurcations in the noisy Duffing–van der Pol equation. In *Stochastic Dynamics* (H. Crauel and M. Gundlach, eds.) 49–70. Springer, Berlin.

[19] MEYN, S. and TWEEDIE, R. (1993). Stability of Markovian processes III: Foster–Lyapunov criteria for continuous time processes. *Adv. Appl. Prob.* **25** 518–548.

[20] PINSKY, M. and WIHSTUTZ, V. (1988). Lyapunov exponents of nilpotent Itô systems. *Stochastics* **25** 43–57.

[21] SAN MARTIN, L. and ARNOLD, L. (1986). A control problem related to the Lyapunov spectrum of stochastic flows. *Mat. Apl. Comput.* **5** 31–64.

[22] SCHENK-HOPPÉ, K. (1996). Bifurcation scenarios of the noisy Duffing–van der Pol oscillator. *Nonlinear Dynamics* **11** 255–274.

[23] SOWERS, R. (2001). On the tangent flow of a stochastic differential equation with fast drift. *Trans. Amer. Math. Soc.* **353** 1321–1334.

[24] WEDIG, W. (1991). Lyapunov exponents and invariant measures of equilibria and limit cycles. *Lyapunov Exponents. Lecture Notes in Math.* **1486** 308–321. Springer, Berlin.

[25] WHITTAKER, E. and WATSON, G. (1927). *A Course of Modern Analysis*, 4th ed. Cambridge Univ. Press.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF SOUTHERN CALIFORNIA
1042 WEST 36TH PLACE, DRB 155
LOS ANGELES, CALIFORNIA 90089-1113