A Note on Diffusion Limits of Chaotic Skew Product Flows

I. Melbourne [*]
Mathematics Department
University of Surrey
Guildford GU2 7XH, UK
and
A.M. Stuart [†]
Mathematics Institute
Warwick University
Coventry CV4 7AL, UK

September 30, 2010

Abstract

We provide an explicit rigorous derivation of a diffusion limit – a stochastic differential equation with additive noise – from a deterministic skew-product flow. This flow is assumed to exhibit time-scale separation and has the form of a slowly evolving system driven by a fast chaotic flow. Under mild assumptions on the fast flow, we prove convergence to a stochastic differential equation as the time-scale separation grows. In contrast to existing work, we do not require the flow to have good mixing properties. As a consequence, our results incorporate a large class of fast flows, including the classical Lorenz equations.

1 Introduction

There is considerable interest in understanding how stochastic behaviour emerges from deterministic systems, both in the mathematics and applications literature. In this note we provide a simple explicit construction of such emergent stochastic

[*]E-mail address: i.melbourne@surrey.ac.uk
[†]E-mail address: a.m.stuart@warwick.ac.uk.

behaviour in the setting of skew-product flows exhibiting time-scale separation. We prove a diffusion limit for the following ordinary differential equations (ODEs):

$$\dot{x}^{(\epsilon)} = \epsilon^{-1} f_0(y^{(\epsilon)}) + f(x^{(\epsilon)}, y^{(\epsilon)}), \quad x^{(\epsilon)}(0) = \xi, \tag{1.1a}$$

$$\dot{y}^{(\epsilon)} = \epsilon^{-2} g(y^{(\epsilon)}), \quad y^{(\epsilon)}(0) = \eta. \tag{1.1b}$$

Here $x^{(\epsilon)} \in \mathbb{R}^d, y^{(\epsilon)} \in \mathbb{R}^\ell$. Roughly speaking we assume that the equation for $y^{(\epsilon)}$ has a compact attractor $\Lambda \subset \mathbb{R}^\ell$ supporting an invariant measure and satisfying certain "mild chaoticity" assumptions. These conditions are stated precisely in Assumptions 1.2 below. In addition, we assume that f_0 should average to zero with respect to this measure.

Consider the stochastic differential equation (SDE)

$$X(t) = \xi + \int_0^t F(X(s))\, ds + \sqrt{\Sigma} W(t). \tag{1.2}$$

where W is unit d-dimensional Brownian motion, Σ is a $d \times d$ covariance matrix (depending on f_0 and g) and $F(x)$ is the average of $f(x, \cdot)$ with respect to the aforementioned invariant measure. The goal of the note is to prove the following limit theorem relating the solution $x^{(\epsilon)}$ of (1.1) and X of (1.2). Throughout we use $\to_w$ to denote weak convergence in the sense of probability measures [1, 2].

Theorem 1.1. *Let Assumptions 1.2 hold and let η be a random variable distributed according to the measure μ on $\mathbb{R}^\ell$ and fix any $\xi \in \mathbb{R}^d$. Then, almost surely with respect to η and W, there is a unique solution $(x^{(\epsilon)}, y^{(\epsilon)}) \in C^1([0,\infty); \mathbb{R}^d \times \mathbb{R}^\ell)$ of* (1.1) *for each $\epsilon > 0$, and a unique solution $X \in C([0,\infty); \mathbb{R}^d)$ of* (1.2)*. Furthermore $x^{(\epsilon)} \to_w X$ in $C([0,\infty), \mathbb{R}^d)$ as $\epsilon \to 0$.*

Throughout the note we make the following standing assumptions.

Assumptions 1.2. *The differential equations* (1.1) *satisfy the following:*

1. *Equation* (1.1b) *with $\epsilon = 1$ has a compact invariant set Λ, $\eta \in \Lambda$, and there is an invariant probability measure μ supported on Λ; expectation with respect to this measure is denoted by $\mathbb{E}$.*

2. *The vector fields $g : \Lambda \to \mathbb{R}^\ell$ and $f_0 : \Lambda \to \mathbb{R}^d$ are locally Lipschitz, and the vector field $f : \mathbb{R}^d \times \Lambda \to \mathbb{R}^d$ is bounded and Lipschitz with uniform Lipschitz constant L.*

3. *The vector field $f_0 : \Lambda \to \mathbb{R}^d$ averages to 0 under $\mu : \mathbb{E} f_0 = 0$.*

4. *Define $W_n(t) = n^{-\frac{1}{2}} \int_0^{nt} f_0(y^{(1)}(\tau))\, d\tau$, for $t \geq 0$. Fix any $T > 0$. We assume the weak invariance principle (WIP), namely that $W_n \to_w \sqrt{\Sigma} W$ in $C([0,T], \mathbb{R}^d)$ as $n \to \infty$ for unit d-dimensional Brownian motion W and some covariance matrix Σ, independent of T.*

5. *Define*

$$F(x) = \mathbb{E} f(x, \cdot) = \int_\Lambda f(x, y) \mu(dy),$$

noting that it is globally bounded by $|f|_\infty$ *and globally Lipschitz with constant L, since* μ *is a probability measure. Let* $T > 0$*. We assume the following large deviation principle (LDP):*

$$\mu\Big(\Big|\frac{1}{T}\int_0^T f(x, y^{(1)}(\tau))\, d\tau \,-\, F(x)\Big| > a\Big) \leq b(a, T),$$

where $b(a, T)$ *is independent of* x *and* $b(a, T) \to 0$ *as* $T \to \infty$ *for all* $a > 0$*.*

Remark 1.3. *(a) The regularity conditions on* f, f_0, g *in assumption 2 guarantee global existence and uniqueness of solutions to the ODEs* (1.1) *and the SDE* (1.2) *for all positive time and all initial conditions* $\xi \in \mathbb{R}^d$, $\eta \in \Lambda$*. We note that uniformity of the Lipschitz constant for* f *is automatic in* y *since* Λ *is compact.*
(b) Assumptions 4 and 5 hold for a large class of flows. In particular, the WIP and LDP are proved in [3, 4] for flows that have a Poincaré map modelled by a Young tower [5, 6] with summable tails. This includes Anosov and Axiom A flows, nonuniformly hyperbolic flows such as Hénon-like flows (where the Poincaré map has a Hénon-like attractor), and Lorenz attractors [7] (including the case of the classical parameter values in [8]). In this class of examples, the Poincaré map has good statistical properties and limit laws such as the WIP and LDP transfer to the flow [9].

There are two main routes leading to emergent stochastic behaviour in deterministic systems. The first is through the elimination of a large number of degrees of freedom, and the reliance on the central limit theorem to provide fluctuations and the second is through time-scale separation; see [10] for an overview. The first mechanism does not require any assumption of chaotic behaviour and may be observed in large systems of oscillators; work in this area was initiated in [11] and more recent work includes [12, 13]. The second mechanism relies on the presence of some fast chaotic dynamics to induce white noise and has a long history in the applied literature; we mention, in particular, the work in [14, 15, 16, 17]. Our work provides a very simple scenario in which the second mechanism may be used, provably, to establish emergent stochastic dynamics. We anticipate that the basic ideas would apply to a far larger class of problems as indicated, for example, by the program outlined in [18]. Moreover the basic mechanism that underlies the work in this note was identified and studied in the seminal paper [19]. However the conditions in that paper can be hard to verify for specific ordinary differential equations. In contrast our construction holds for explicit systems on $\mathbb{R}^\ell$ such as the classical Lorenz equations.

An important aspect of our theory is that we require no knowledge of mixing properties of the flow, In contrast, previous rigorous results in the literature required strong assumptions on the mixing properties of the flow. See [20] for the most powerful results in this direction where it is required that the flow has stretched exponential decay of correlations. Even for Anosov flows this has been proved only in very special cases [21, 22, 23]. Superpolynomial decay has been proved for typical Anosov and Axiom A flows [24, 25] and typical nonuniformly hyperbolic flows governed by Young towers [26, 27] but only for very smooth observables; this smoothness would have to be imposed on f_0. For the Lorenz equations there are currently no results at all on rates of mixing (though superpolynomial decay holds for typical nearby flows by [27]).

2 Weak Invariance and Large Deviation Principles

The WIP and LDP assumed above have two useful implications which we detail in this section, and then use in the next section to prove the main theorem.

Proposition 2.1. *Let* $(x^{(\epsilon)}(t), y^{(\epsilon)}(t))$ *denote the solution to* (1.1) *with* $f \equiv 0$, $\xi = 0$ *and with* η *a random variable distributed according to the measure* μ *on* Λ. *Let* $T > 0$. *Then* $x^{(\epsilon)} \to_w \sqrt{\Sigma} W$ *in* $C([0,T], \mathbb{R}^d)$ *as* $\epsilon \to 0$. *Here,* W *is unit* d*-dimensional Brownian motion and the covariance matrix* Σ *is independent of* T.

Proof. Note that $y^{(1)}(t)$ is the solution to the IVP $\dot{y} = g(y)$, $y(0) = \eta$. Define $W_n(t) = n^{-\frac{1}{2}} \int_0^{nt} f_0(y^{(1)}(\tau))\, d\tau$, for $t \in [0,T]$. By the WIP, $W_n \to_w \sqrt{\Sigma} W$ in $C([0,T], \mathbb{R}^d)$ as $n \to \infty$.

Now $y^{(\epsilon)}(t) = y^{(1)}(t\epsilon^{-2})$. Hence

$$x^{(\epsilon)}(t) = \epsilon^{-1} \int_0^t f_0(y^{(\epsilon)}(s))\, ds = \epsilon \int_0^{t\epsilon^{-2}} f_0(y^{(1)}(\tau))\, d\tau.$$

Writing $n = \epsilon^{-2}$, we obtain $x^{(\epsilon)}(t) = W_n(t)$ and the result follows. $\square$

Proposition 2.2. *For any* $a, T > 0$, $n \geq 0$, $x \in \mathbb{R}^d$,

$$\mathbb{E}\Big| \frac{1}{T} \int_{nT}^{(n+1)T} f(x, y^{(1)}(\tau))\, d\tau \;-\; F(x) \Big| \leq a + 2|f|_\infty\, b(a,T).$$

Proof. Let $S(T,x) = |\frac{1}{T} \int_0^T f(x, y^{(1)}(\tau))\, d\tau - F(x)|$. Then, by the LDP,

$$\begin{aligned} \mathbb{E} S(T,x) &= \int_{S(T,x) \leq a} S(T,x)\, d\mu + \int_{S(T,x) > a} S(T,x)\, d\mu \\ &\leq a + |S|_\infty\, b(a,T) \leq a + 2|f|_\infty\, b(a,T). \end{aligned}$$

This proves the result for $n = 0$, and the general case follows from invariance of μ. $\square$

3 Diffusion Limit

We now prove the diffusion limit contained in Theorem 1.1. The method of proof generalizes that described in Chapter 18 of [28] for homogenization in SDEs with additive noise and a skew-product form.

Proof of Theorem 1.1 To prove weak convergence on $[0, \infty)$, it suffices to establish weak convergence on $[0,T]$ for each fixed $T > 0$.

Write $W^{(\epsilon)}(t) = \int_0^t \frac{1}{\epsilon} f_0(y^{(\epsilon)}(s))\,ds$. By integrating the $x^{(\epsilon)}$ equation we have

$$\begin{aligned} x^{(\epsilon)}(t) &= \xi + \int_0^t \frac{1}{\epsilon} f_0(y^{(\epsilon)}(s))\,ds + \int_0^t f(x^{(\epsilon)}(s), y^{(\epsilon)}(s))\,ds \\ &= \xi + W^{(\epsilon)}(t) + Z^{(\epsilon)}(t) + \int_0^t F(x^{(\epsilon)}(s))\,ds \end{aligned}$$

where

$$Z^{(\epsilon)}(t) = \int_0^t \big(f(x^{(\epsilon)}(s), y^{(\epsilon)}(s)) - F(x^{(\epsilon)}(s))\big)ds.$$

We show below that $Z^{(\epsilon)} \to 0$ in $L^1(C([0,T],\mathbb{R}^d);\mu)$. By Proposition 2.1, $W^{(\epsilon)} \to_w \sqrt{\Sigma}W$ in $C([0,T],\mathbb{R}^d)$. It follows that $W^{(\epsilon)} + Z^{(\epsilon)} \to_w \sqrt{\Sigma}W$ in $C([0,T],\mathbb{R}^d)$. Now consider the continuous map $\mathcal{G} : C([0,T],\mathbb{R}^d) \to C([0,T],\mathbb{R}^d)$ given by $\mathcal{G}(u) = v$ where v is the unique solution to the integral equation

$$v(t) = \xi + u(t) + \int_0^t F(v(s))\,ds.$$

Then

$$x^{(\epsilon)} = \mathcal{G}(W^{(\epsilon)} + Z^{(\epsilon)}).$$

Since continuous maps preserve weak convergence, it follows that $x^{(\epsilon)} \to_w X$ as required.

It remains to show the L^1 convergence of $Z^{(\epsilon)}$ to 0 in $C([0,T],\mathbb{R}^d)$. Define $g(x,y) = f(x,y) - F(x)$ and note that $|g|_\infty \le 2|f|_\infty$ and $\mathrm{Lip}(g) \le 2L$. Then $Z^{(\epsilon)}(t) = \int_0^t g(x^{(\epsilon)}(s), y^{(\epsilon)}(s))\,ds$. Let $N = [t/\delta]$ and write $Z^{(\epsilon)}(t) = Z^{(\epsilon)}(N\delta) + I_0$ where $I_0 = \int_{N\delta}^t g(x^{(\epsilon)}(s), y^{(\epsilon)}(s))\,ds$. We have

$$|I_0| \le (t - N\delta)|g|_\infty \le 2|f|_\infty \delta. \tag{3.1}$$

We now estimate $Z^{(\epsilon)}(N\delta)$ as follows:

$$\begin{aligned} Z^{(\epsilon)}(N\delta) &= \sum_{n=0}^{N-1} \int_{n\delta}^{(n+1)\delta} g(x^{(\epsilon)}(s), y^{(\epsilon)}(s))\,ds \\ &= \sum_{n=0}^{N-1} \int_{n\delta}^{(n+1)\delta} \big(g(x^{(\epsilon)}(s), y^{(\epsilon)}(s)) - g(x^{(\epsilon)}(n\delta), y^{(\epsilon)}(s))\big)\,ds \\ &\qquad + \sum_{n=0}^{N-1} \int_{n\delta}^{(n+1)\delta} g(x^{(\epsilon)}(n\delta), y^{(\epsilon)}(s))\,ds \\ &= I_1 + I_2. \end{aligned}$$

For $s \in [n\delta,(n+1)\delta]$, we have $|x^{(\epsilon)}(s) - x^{(\epsilon)}(n\delta)| \le (|f_0|_\infty + |f|_\infty)\delta\epsilon^{-1}$. Hence

$$|I_1| \le N\delta\,\mathrm{Lip}(g)(|f_0|_\infty + |f|_\infty)\delta\epsilon^{-1} \le 2L(|f_0|_\infty + |f|_\infty)T\delta\epsilon^{-1}. \tag{3.2}$$

Next,

$$I_2 = \sum_{n=0}^{N-1}\int_{n\delta}^{(n+1)\delta}(f(x^{(\epsilon)}(n\delta), y^{(\epsilon)}(s)) - F(x^{(\epsilon)}(n\delta))\,ds = \delta\sum_{n=0}^{N-1}J_n,$$

where

$$\begin{aligned} J_n &= \delta^{-1}\int_{n\delta}^{(n+1)\delta} f(x^{(\epsilon)}(n\delta), y^{(\epsilon)}(s))\,ds - F(x^{(\epsilon)}(n\delta)) \\ &= \delta^{-1}\epsilon^2\int_{n\delta\epsilon^{-2}}^{(n+1)\delta\epsilon^{-2}} f(x^{(\epsilon)}(n\delta), y^{(1)}(s))\,ds - F(x^{(\epsilon)}(n\delta)). \end{aligned}$$

Let $a > 0$. Applying Proposition 2.2, $\mathbb{E}|J_n| \le a + 2|f|_\infty\, b(a, \delta\epsilon^{-2})$. It follows that

$$\mathbb{E}\sup_{t\in[0,T]}|I_2(t)| \le T(a + 2|f|_\infty\, b(a,\delta\epsilon^{-2})). \tag{3.3}$$

Finally, we set $\delta = \epsilon^{\frac{3}{2}}$. By (3.1) and (3.2), $\lim_{\epsilon\to 0}\mathbb{E}\sup_{t\in[0,T]}|I_0(t)| = \lim_{\epsilon\to 0}\mathbb{E}\sup_{t\in[0,T]}|I_1(t)| = 0$. Moreover, by (3.3) and the assumption on b in the LDP in Assumption 1.2(5), we have

$$\limsup_{\epsilon\to 0}\mathbb{E}\sup_{t\in[0,T]}|I_2(t)| \le Ta,$$

and, since $a > 0$ is arbitrary,

$$\lim_{\epsilon\to 0}\mathbb{E}\sup_{t\in[0,T]}|I_2(t)| = 0.$$

Altogether, $\lim_{\epsilon\to 0}\mathbb{E}\sup_{t\in[0,T]}|Z^{(\epsilon)}(t)| = 0$ as required.

4 Conclusions

The construction in this paper shows how some new ideas in the theory of dynamical systems can be used to prove a homogenization principle in ODEs, leading to emergent stochastic behaviour. The arguments are very straightforward, and are given only in the case of additive noise. However in the situation where the limiting SDE is one dimensional the ideas of Sussmann [29] can be used to derive a limiting SDE in which noise appears multiplicatively. Generalizing these ideas to skew product flows where the SDE is of higher dimension will require the theory of rough paths [30] and is the subject of ongoing work.

Finally a comment on the differences between [28] *homogenization* and *averaging* in ODE systems like (1.1). There is current interest [31] in the derivation of averaging principles for systems of ODEs exhibiting three time scales of order $\mathcal{O}(\epsilon^{-2}), \mathcal{O}(\epsilon^{-1})$ and $\mathcal{O}(1)$. The motivation is the construction of efficient numerical schemes for computation of the averaged solution, which is deterministic. Theorem 1.1, which also concerns the limiting behaviour of a system containing three time-scales, corresponds to a homogenization principle with a stochastic limit, rather than an averaging principle with deterministic limit. Thus our work provides an example of a three scale system for which an effective deterministic averaged equation cannot exist.

Acknolwedgements. The authors are grateful to Niklas Brännström and Matthew Nicol for helpful discussions. AMS is grateful to EPSRC and ERC for financial support. The research of IM was supported in part by EPSRC Grant EP/F031807/01

References

[1] P. Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, 1999.

[2] P. Billingsley. *Probability and measure*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, third edition, 1995. A Wiley-Interscience Publication.

[3] I. Melbourne and M. Nicol. Almost sure invariance principle for nonuniformly hyperbolic systems. *Comm. Math. Phys.*, 260:131–146, 2005.

[4] I. Melbourne and M. Nicol. Large deviations for nonuniformly hyperbolic systems. *Trans. Amer. Math. Soc.*, 360:6661–6676, 2008.

[5] L.-S. Young. Statistical properties of dynamical systems with some hyperbolicity. *Ann. of Math.*, 147:585–650, 1998.

[6] L.-S. Young. Recurrence times and rates of mixing. *Israel J. Math.*, 110:153–188, 1999.

[7] M. Holland and I. Melbourne. Central limit theorems and invariance principles for Lorenz attractors. *J. London Math. Soc.*, 76:345–364, 2007.

[8] E. D. Lorenz. Deterministic nonperiodic flow. *J. Atmosph. Sci.*, 20:130–141, 1963.

[9] I. Melbourne and A. Török. Statistical limit theorems for suspension flows. *Israel J. Math.*, 144:191–209, 2004.

[10] D. Givon, R. Kupferman, and A. M. Stuart. Extracting macroscopic dynamics: model problems and algorithms. *Nonlinearity*, 17(6):R55–R127, 2004.

[11] G. W. Ford, M. Kac, and P. Mazur. Statistical mechanics of assemblies of coupled oscillators. *J. Mathematical Phys.*, 6:504–515, 1965.

[12] R. Kupferman, A. M. Stuart, J. R. Terry, and P. F. Tupper. Long-term behaviour of large mechanical systems with random initial data. *Stoch. Dyn.*, 2(4):533–562, 2002.

[13] G. Ariel and E. Vanden-Eijnden. A strong limit theorem in the kac-zwanzig model. *Nonlinearity*, 22:145, 2009.

[14] C. Beck. Brownian motion from deterministic dynamics. *Phys. A*, 169(2):324–336, 1990.

[15] W. Just, H. Kantz, Ch. Rödenbeck, and M. Helm. Stochastic modelling: replacing fast degrees of freedom by noise. *J. Phys. A*, 34(15):3199–3213, 2001.

[16] A. J. Majda and I. Timofeyev. Remarkable statistical behavior for truncated Burgers-Hopf dynamics. *Proc. Natl. Acad. Sci. USA*, 97(23):12413–12417 (electronic), 2000.

[17] A. M. Majda, I. Timofeyev, and E. Vanden-Eijnden. Stochastic models for selected slow variables in large deterministic systems. *Nonlinearity*, 19(4):769–794, 2006.

[18] R. S. Mackay. Langevin equation for slow degrees of freedom of hamiltonian systems. *Understanding Complex Systems*, pages 89–102, 2010.

[19] G. C. Papanicolaou and W. Kohler. Asymptotic theory of mixing stochastic ordinary differential equations. *Comm. Pure Appl. Math.*, 27:641–668, 1974.

[20] D. Dolgopyat. Averaging and invariant measures. *Moscow Math. J.*, 5:537–576, 2005.

[21] N. I. Chernov. Markov approximations and decay of correlations for Anosov flows. *Ann. of Math.*, 147:269–324, 1998.

[22] D. Dolgopyat. On the decay of correlations in Anosov flows. *Ann. of Math.*, 147:357–390, 1998.

[23] C. Liverani. On contact Anosov flows. *Ann. of Math.*, 159:1275–1312, 2004.

[24] D. Dolgopyat. Prevalence of rapid mixing in hyperbolic flows. *Ergodic Theory Dynam. Systems*, 18:1097–1114, 1998.

[25] M. J. Field, I. Melbourne, and A. Török. Stability of mixing and rapid mixing for hyperbolic flows. *Ann. of Math.*, 166:269–291, 2007.

[26] I. Melbourne. Rapid decay of correlations for nonuniformly hyperbolic flows. *Trans. Amer. Math. Soc.*, 359:2421–2441, 2007.

[27] I. Melbourne. Decay of correlations for slowly mixing flows. *Proc. London Math. Soc.*, 98:163–190, 2009.

[28] G. A. Pavliotis and A. M. Stuart. *Multiscale Methods: Homogenization and Averaging*. Springer, 2008.

[29] H. Sussmann. On the gap between deterministic and stochastic ordinary differential equations. *Ann. Prob.*, 6:19–41, 1978.

[30] T. J. Lyons. Differential equations driven by rough signals. *Rev. Mat. Iberoamericana*, 14(2):215–310, 1998.

[31] G. Ariel, B. Engquist, and R. Tsai. Oscillatory systems with three separated time scales – analysis and computation. Technical report, UCLA Computational Applied Mathematics Report 10-56, 2010.