Geometric Ergodicity of Some Hypo-Elliptic Diffusions for Particle Motions

ARTICLE *in* MARKOV PROCESSES AND RELATED FIELDS · JANUARY 2002

Impact Factor: 0.48

Jonathan C. Mattingly
Duke University

Geometric Ergodicity of Some Hypo-Elliptic Diffusions for Particle Motions

J.C. Mattingly[1] and A.M. Stuart[2]

Abstract

Two degenerate SDEs arising in statistical physics are studied. The first is a Langevin equation with state-dependent noise and damping. The second is the equation of motion for a particle obeying Stokes' law in a Gaussian random field; this field is chosen to mimic certain features of turbulence. Both equations are hypo-elliptic and smoothness of probability densities may be established. By developing appropriate Lyapunov functions and by studying the necessary control problems, geometric ergodicity is proved.

Key Words: Geometric Ergodicity, Stochastic Differential Equations, Langevin Equation, Synthetic Turbulence, Hypoelliptic and Degenerate Diffusions.

[1]Mathematics Department, Stanford University, Stanford, CA94305, USA. Supported by the National Science Foundation under grant DMS-9971087.

[2]Mathematics Institute, Warwick University, Coventry, CV4 7AL, England. Supported by the Engineering and Physical Sciences Research Council of the UK under grant GR/N00340.

1 Introduction

The first objective of this note is to highlight a straightforward approach to proving geometric ergodicity for Markov chains on uncountable state spaces. The techniques are based on the approaches in [10, 8, 9, 2, 3]. We follow a recent development of these ideas described in [7], tailored to the study of degenerate diffusions, and their discretizations. We make a minor extension of the theory in [7] (from state space $\mathbb{X} = \mathbb{R}^d$ to state space $\mathbb{X} = \mathbb{T}^{d_1} \times \mathbb{R}^{d_2}$) and then use this to pursue the second objective of the note: to prove geometric ergodicity of some naturally occuring models in statistical physics. As the approach to ergodicity is designed to handle degenerate diffusions it is hence ideal for problems arising in mechanics where noise arises as a force acting directly only on momenta and not on positions.

In section 2 we state the basic theory. In section 3 we give an application to the Langevin equation for the motion of a particle in a periodic potential, subject to state-dependent noise and damping obeying the fluctuation-dissipation relation. In section 4 we give an application to the random dynamical system formed by considering the motion of particles obeying Stokes' law in a (synthetic) turbulent velocity field; this velocity field may be viewed as the solution of a linear stochastic PDE.

2 Geometric Ergodicity

In this section we state Theorem 2.3, guaranteeing geometric ergodicity. The proof is a trivial modification of that in [7] (only the state space has changed) and so we omit it. In any case the statement and proof of the theorem are close to existing treatments in the literature [10, 8, 9, 2, 3] and the main advantage of our formulation is simply that it is well-adapted to the study of possibly degenerate diffusions and their time discretizations.

Consider a homogenous Markov process $x(t)$ $(t \in \mathbb{R}^+)$ or Markov chain $x(t)$ $(t \in \mathbb{Z}^+)$ on a state space $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$ where $\mathbb{X} = \mathbb{T}^{d_1} \times \mathbb{R}^{d_2}$ with d_1, d_2 non-negative integers. [1] Here $\mathcal{B}(\mathbb{X})$ denotes the Borel σ-algebra on $\mathbb{X}$. To help combine our treatment of continuous and discrete time, we set $\mathcal{T} = \mathbb{R}^+$ (resp. $\mathbb{Z}^+$) for the Markov process (resp. chain) case. Throughout the remainder of the paper $\mathcal{B}_\delta(x)$ denotes the open ball of radius δ centred at $x \in \mathbb{X}$. We denote the transition kernel of the Markov process or chain by

$$P_t(x, A) \stackrel{\text{def}}{=} \mathbb{P}(x(t) \in A | x(0) = x), \quad t \in \mathcal{T}, x \in \mathbb{X}, A \in \mathcal{B}(\mathbb{X}) .$$

Now define the Markov chain with the kernel $P(x, A) \stackrel{\text{def}}{=} P_T(x, A)$, formed by sampling at the rate $T \in \mathcal{T}$. The following assumptions will give geometric ergodicity.

Assumption 2.1 Minorization Condition *There is a choice of $T \in \mathcal{T}$, compact C, an $\eta > 0$, and a probability measure ν, with $\nu(C^c) = 0$ and $\nu(C) = 1$, such that*

$$P(x, A) \geq \eta \nu(A) \quad \forall A \in \mathcal{B}(\mathbb{X}), x \in C.$$

[1] We set $d' = d_1 + d_2$ in general, and it is convenient in the Langevin case to set $d_1 = d_2 = d$.

The minorization condition is a particular case of the general concept of small set, developed for the purposes of proving ergodicity of Markov chains on general states spaces; see [10, 8].

Let $\{x_n\}_{n\in\mathbb{Z}^+}$ be the Markov chain generated by the kernel $P(x,A)$. We use a Lyapunov function to control the return times to C. In the following $\mathcal{F}_n$ denotes the σ-algebra of events up to and including the n^{th} iteration. $\mathbb{E}$ denotes expectation induced by the Markov chain or process. We use the notation $\mathbb{E}^y$ when we wish to indicate dependence on the starting point y of the chain.

Assumption 2.2 Lyapunov Condition *There is a function* $V:\mathbb{X}\to[1,\infty)$, *with* $\lim_{\|x\|\to\infty} V(x) = \infty$, *and real numbers* $\alpha\in(0,1)$, *and* $\beta\in[0,\infty)$ *such that*

$$\mathbb{E}[V(x_{n+1})|\mathcal{F}_n] \le \alpha V(x_n) + \beta.$$

The use of such Lyapunov conditions to prove ergodicity for uniformly elliptic diffusions is a well-developed subject; see [3, 5, 8].

In what follows, we will use the shorthand notation $|g|\le V$ to mean $|g(x)|\le V(x)$ for all x and define

$$\mathcal{G} = \{\text{measurable}\, g:\mathbb{X}\to\mathbb{R} \text{ with } |g|\le V\}.$$

Theorem 2.3 *Let $x(t)$ denote the Markov chain or process with transition kernel $P_t(x,A)$. Let $\{x_n\}_{n\in\mathbb{Z}^+}$ denote the embedded Markov chain with transition kernel $P(x,A)=P_T(x,A)$. Assume that there is a $T>0$ for which the following holds: the Markov chain $\{x_n\}_{n\in\mathbb{Z}^+}$ satisfies the* Minorization Condition *and the* Lyapunov Condition *with C given by*

$$C = \left\{x : V(x) \le \frac{2\beta}{\gamma-\alpha}\right\} \tag{2.1}$$

for some $\gamma\in(\alpha^{\frac{1}{2}},1)$. Then there exists a unique invariant measure π. Furthermore there is $r(\gamma)\in(0,1)$ and $\kappa(\gamma)\in(0,\infty)$ such that for all measurable $g\in\mathcal{G}$

$$|\mathbb{E}^{x_0} g(x_n) - \pi(g)| \le \kappa r^n V(x_0).$$

This theorem is very similar to existing results in the literature; see Chapter 15 of [8] and the references therein. The proof uses standard coupling techniques. The main interest is that the theorem has been tailored for application to degenerate SDEs.

Throughout this note we will use Theorem 2.3 to study the following SDE:

$$dx = Y(x)dt + \Sigma(x)dW, \quad x(0)=y, \tag{2.2}$$

where $x\in\mathbb{X}$, $Y:\mathbb{X}\to\mathbb{R}^{d'}$ and W is a standard m-dimensional Brownian motion. Thus $\Sigma:\mathbb{X}\to\mathbb{R}^{d'\times m}$. To establish geometric ergodicity for this SDE we use the following approach.

Assumption 2.4 *There is a function $V:\mathbb{X}\to[1,\infty)$, with $\lim_{\|x\|\to\infty}V(x)=\infty$, and real numbers $a\in(0,\infty)$, $d\in(0,\infty)$ such that*

$$\mathcal{A}\{V(x)\} \le -a\{V(x)\} + d, \tag{2.3}$$

where $\mathcal{A}$ is the generator for (2.2) *given by*

$$\mathcal{A}g = \sum_{i=1}^{d} Y_i \frac{\partial g}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{d} \left[\Sigma\Sigma^T\right]_{ij} \frac{\partial^2 g}{\partial x_i \partial x_j}. \tag{2.4}$$

Lemma 2.5 *Let Assumption 2.4 hold. Then the* Lyapunov Condition *holds.*

Sketch Proof (see also [5], Theorem 11.9.1). This is just the infinitesimal version of Assumption 2.2. If $\mathcal{F}_s$ is the $\sigma-$algebra of all events up to time s, it follows that

$$\mathbb{E}^y\{V(x(t))|\mathcal{F}_s\} \le e^{-a(t-s)}V(x(s)) + \frac{d}{a}[1 - e^{-a(t-s)}] \ . \tag{2.5}$$

If $x_n = x(nT)$, so that $\{x_n\}_{n=0}^{\infty}$ is a Markov chain, then (2.5) shows that Assumption 2.2 holds for this Markov chain: with $\alpha = e^{-aT}$ and $\beta = d/a$. $\square$

Assumption 2.6 *The Markov process generated by* (2.2) *with transition kernel $P_t(x, A)$ satisfies, for some fixed compact set $C \in \mathcal{B}(\mathbb{X})$, the following:*

i) for some $y^* \in int(C)$ there is, for any $\delta > 0$, a $t_1 = t_1(\delta) \in \mathcal{T}$ such that

$$P_{t_1}\Big(x, \mathcal{B}_\delta(y^*)\Big) > 0 \quad \forall x \in C;$$

ii) for $t \in \mathcal{T}$ the transition kernel possesses a density $p_t(x, y)$, precisely

$$P_t(x, A) = \int_A p_t(x, y)dy \quad \forall x \in C, \ A \in \mathcal{B}(\mathbb{X}) \cap \mathcal{B}(C),$$

and $p_t(x, y)$ is jointly continuous in $(x, y) \in C \times C$.

Lemma 2.7 *Let Assumption 2.6 hold. Then the* Minorization Condition *holds for the same set C.*

This is proved in [7].

Corollary 2.8 *Let $x(t)$ denote the solution of the SDE (2.2) with transition kernel $P_t(x, A)$. Assume that there is a $T > 0$ for which the following holds: the SDE satisfies Assumptions 2.4 and 2.6 with C given by*

$$C = \left\{x : V(x) \le \frac{2\beta}{\gamma - \alpha}\right\}$$

for some $\gamma \in (\alpha^{\frac{1}{2}}, 1)$. Then there exists a unique invariant measure π. Furthermore there is $\mu(\gamma) \in (0, 1)$ and $\kappa(\gamma) \in (0, \infty)$ such that for all measurable $g \in \mathcal{G}$

$$|\mathbb{E}^y g(x(t)) - \pi(g)| \le \kappa e^{-\mu t} V(y).$$

Sketch Proof Assumptions 2.4 and 2.6 prove that an appropriately sampled chain is ergodic. To extend the proof to continuous time we use an argument from [9], also employed in [7], Theorem 3.2. $\square$

3 The Langevin Equation

In this section we prove geometric ergodicity of a Langevin equation for the motion of a particle in a periodic potential, subject to state-dependent noise and damping obeying the fluctuation-dissipation relation. This result is stated in Theorem 3.2.

Let W be a standard d-dimensional Brownian Motion, $F : \mathbb{T}^d \to \mathbb{R}$, $\sigma : \mathbb{T}^d \to \mathbb{R}^{d\times d}$, $\gamma : \mathbb{T}^d \to \mathbb{R}^{d\times d}$ and $\rho_i : \mathbb{T}^d \to \mathbb{R}^d$ the ith column of σ. Consider the Langevin SDE for $q, p \in \mathbb{R}^d$ the position and momenta of a particle of unit mass, namely

$$dq = pdt, \tag{3.1}$$
$$dp = -\gamma(q)pdt - \nabla F(q)dt + \sigma(q)dW. \tag{3.2}$$

(Throughout this section ∇ represents a gradient with respect to q). Note that the SDE is degenerate, with noise only in the momenta. However the problem is hypo-elliptic and the theory developed in the previous section will be applicable.

We assume the fluctuation-dissipation relation ($\propto$ denotes proportional upto a constant independent of (q, p))

$$\gamma(q) \propto \sigma(q)\sigma(q)^T \quad \forall q \in \mathbb{T}^d$$

which implies the existence of a canonical invariant measure for (3.1)-(3.2). In the case $d = 1, \gamma(q) = \alpha\sigma(q)^2/2$, for example, there is a known invariant measure with density

$$\rho(p, q) \propto \exp\{-\alpha[\frac{p^2}{2} + F(q)]\}.$$

Assumption 3.1 *We assume that σ and F are both C^∞, that*

$$1 \le F(q) \le F^+ < \infty \quad \forall q \in \mathbb{T}^d \tag{3.3}$$

and that γ is uniformly positive-definite: for some $\gamma^- > 0$

$$\gamma^-\|z\|^2 \le \langle z, \gamma(q)z\rangle \quad \forall q \in \mathbb{T}^d, z \in \mathbb{R}^d. \tag{3.4}$$

From this assumption it follows that $\sigma(q)$ is invertible for all $q \in \mathbb{T}^d$ and hence that the $\{\rho_i\}_{i=1}^d$ are linearly independent everywhere in $\mathbb{T}^d$.

Under these conditions it is possible to prove global in time existence and uniqueness of solutions to (3.1)–(3.2) using the fact that all vector fields are globally Lipschitz on $\mathbb{X} = \mathbb{T}^d \times \mathbb{R}^d$. It is expedient to write (3.1)–(3.2) in the abstract form (2.2) where now

$$x = \begin{pmatrix} q \\ p \end{pmatrix} \in \mathbb{T}^d \times \mathbb{R}^d,\ W = \begin{pmatrix} W_1 \\ \vdots \\ W_d \end{pmatrix} \in \mathbb{R}^d,\ Y(x) = \begin{pmatrix} p \\ -\gamma(q)p - \nabla F(q) \end{pmatrix},\ \Sigma(x) = \begin{pmatrix} O \\ \sigma(q) \end{pmatrix}. \tag{3.5}$$

Here each W_i is an independent standard one-dimensional Brownian motion and $O \in \mathbb{R}^{d\times d}$ is the zero matrix. Note that we may write

$$\Sigma(x)dW = \sum_{i=1}^d X_i(q)dW_i,\ X_i = \begin{pmatrix} 0 \\ \rho_i(q) \end{pmatrix},\ 0 \in \mathbb{R}^d. \tag{3.6}$$

For (3.1)–(3.2), it is useful to define the Lyapunov function

$$V(x) \stackrel{\text{def}}{=} \frac{1}{2}||p||^2 + F(q) \tag{3.7}$$

which is, of course, the Hamiltonian for the Langevin equation in the absence of noise and damping.

We may now define

$$\mathcal{G}_l = \{\text{measurable}\, g : \mathbb{T}^d \times \mathbb{R}^d \to \mathbb{R} \text{ with } |g| \leq V^l\}.$$

Theorem 3.2 *Let Assumption 3.1 hold. The SDE* (3.1)–(3.2) *with* $x(t) = (q(t)^T, p(t)^T)^T$ *has a unique invariant measure* π *on* $\mathbb{T}^d \times \mathbb{R}^d$. *Fix any* $l \geq 1$. *If* $x(0) = y$ *then there exists* $C = C(l) > 0$, $\lambda = \lambda(l) > 0$ *such that, for all* $g \in \mathcal{G}_l$,

$$\left|\mathbb{E}^y g(x(t)) - \pi(g)\right| \leq CV(y)^l e^{-\lambda t} \quad \textit{for all } t \geq 0. \tag{3.8}$$

Proof The result follows from an application of Theorem 2.3. First note that

$$V(x) \geq 1 + \frac{1}{2}||p||^2 \tag{3.9}$$

since F is bounded below by 1. Thus $V(x)^l \to \infty$ as $||x|| \to \infty$. Lemma 3.3 shows that if $\mathcal{A}$ is the generator of the process governed by (3.1)–(3.2), that is, in the form (2.2) with the definitions (3.5),

$$\mathcal{A}g = \sum_{i=1}^{2d} Y_i \frac{\partial g}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{2d} \left[\Sigma\Sigma^T\right]_{ij} \frac{\partial^2 g}{\partial x_i \partial x_j} \tag{3.10}$$

then

$$\mathcal{A}\{V(x)^l\} \leq -a_l\{V(x)^l\} + d_l$$

for some $a_l, d_l > 0$. Thus Assumption 2.4 (the Lyapunov condition) holds.

Lemma 3.4, found at the end of this section, proves that, in any positive time, any open set may be reached with positive probability. Thus Assumption 2.6(i) holds with any choice of C and y^*. We show that Assumption 2.6(ii) holds below and so we have the minorization condition; hence Corollary 2.8 gives the desired result.

We now show that Assumption 2.6(ii) holds. Define

$$\mathcal{L} = Lie\{Y, X_1, \ldots, X_d\},$$

namely the Lie algebra generated by $\{Y, X_1, \ldots, X_d\}$. Let $\mathcal{L}_0$ be the ideal in $\mathcal{L}$ generated by $\{X_1, \ldots, X_d\}$. By Theorem 38.16 in [11] (or results in [1, 4]), it suffices to show that $\mathcal{L}_0$ spans $\mathbb{R}^{2d}$ to verify Assumption 2.6(ii). Note that

$$X_i = \begin{pmatrix} 0 \\ \rho_i(q) \end{pmatrix} \quad \text{and} \quad [X_i, Y] = \begin{pmatrix} \rho_i(q) \\ -\{\nabla\rho_i(q)\}p - \gamma(q)\rho_i(q) \end{pmatrix}.$$

Here $\nabla f(q)$, for $f : \mathbb{R}^d \to \mathbb{R}^d$, is the matrix with ij^{th} entry $\partial f^i(q)/\partial q_j$, with $f^i(q)$ the i^{th} component of $f(q)$. (Note that $\rho_i : \mathbb{R}^d \to \mathbb{R}^d$). Since σ has linearly independent columns σ_i,

$$\{X_1, \ldots, X_d, [X_1, Y], \ldots, [X_d, Y]\}$$

span $\mathbb{R}^{2d}$ everywhere in $\mathbb{X}$, implying the required smoothness. $\square$

The previous theorem requires the following two lemmas:

Lemma 3.3 *For every $l \geq 1$, there exists $a_l \in (0, \infty)$ and $d_l \in (0, \infty)$ such that, for equation* (3.1)–(3.2) *with $\mathcal{A}$ given by* (3.10),

$$\mathcal{A}\{V(x)^l\} \leq -a_l\{V(x)^l\} + d_l \ .$$

Proof We do the case $l = 1$ first. Let

$$\begin{aligned}
Y_i(x) &= p_i, && i = 1, \cdots, d \\
Y_i(x) &= -\{\gamma(q)p\}_i - \frac{\partial F}{\partial q_i}(q), && i = d+1, ..., 2d \\
\frac{\partial V}{\partial x_i} &= \frac{\partial F}{\partial q_i}(q), && i = 1, \ldots, d \\
\frac{\partial V}{\partial x_i} &= p_i && i = d+1, \ldots, d.
\end{aligned}$$

Thus we obtain

$$\sum_{i=1}^{2d} Y_i \frac{\partial V}{\partial x_i} = \langle p, \nabla F(q)\rangle - \langle p, \gamma(q)p\rangle - \langle p, \nabla F(q)\rangle$$

so that

$$\sum_{i=1}^{2d} Y_i \frac{\partial V}{\partial x_i} = -\langle p, \gamma(q)p\rangle \leq -\gamma^- \|p\|^2.$$

Also,

$$\Sigma\Sigma^T = \begin{pmatrix} 0 & 0 \\ 0 & \sigma\sigma^T \end{pmatrix}$$

and thus

$$\frac{1}{2}\sum_{i,j=1}^{2d} [\Sigma\Sigma^T]_{ij} \frac{\partial^2 V}{\partial x_i \partial x_j} = \frac{1}{2}\sum_{i,j=1}^{d} [\sigma\sigma^T]_{ij} \frac{\partial^2 V}{\partial x_{d+i}\partial x_{d+j}} = \frac{1}{2}\sum_{i=1}^{d} [\sigma\sigma^T]_{ii} \frac{\partial^2 V}{\partial p_i^2} \ .$$

But

$$\frac{\partial^2 V}{\partial p_i^2} = 1 \quad \& \quad \frac{1}{2}\sum_{i=1}^{d} [\sigma\sigma^T]_{ii} = \frac{1}{2}\sum_{i,j=1}^{d} \sigma_{ij}^2 = \frac{1}{2}\|\sigma\|_F^2,$$

where $\|\cdot\|_F$ is the Frobenius norm on matrices. Combining and using the uniform positivity of $\gamma(q)$, we have

$$\mathcal{A}V(x) \leq -2\gamma^- V(x) + 2\gamma^- F(q) + \frac{1}{2}\|\sigma(q)\|^2$$

which, since q is in a compact set and F, σ are smooth, gives the required result for $l = 1$.

Now we calculate $\mathcal{A}\{V(x)^l\}$. To this end, note that

$$\begin{aligned}
\frac{\partial}{\partial x_i}\left\{V(x)^l\right\} &= l\{V(x)\}^{l-1}\frac{\partial V}{\partial x_i} \\
\frac{\partial^2}{\partial x_i \partial x_j}\left\{V(x)^l\right\} &= \frac{\partial}{\partial x_j}\left\{l\{V(x)\}^{l-1}\frac{\partial V}{\partial x_i}\right\}, \\
\text{and } \mathcal{A}\{V(x)^l\} &= l\{V(x)\}^{l-1}\mathcal{A}V + \frac{1}{2}\sum_{i,j=1}^{d}[\sigma\sigma^T]_{ij}l(l-1)V(x)^{l-2}\frac{\partial V}{\partial p_i}\frac{\partial V}{\partial p_j}.
\end{aligned}$$

But

$$\frac{\partial V}{\partial p_i} = p_i$$

and hence, by using (3.9), we obtain

$$\frac{1}{2}l(l-1)\sum_{i,j=1}^{d}[\sigma\sigma^T]_{ij}\frac{\partial V}{\partial p_i}\frac{\partial V}{\partial p_j} \leq \chi V(x)$$

for some $\chi > 0$. Thus

$$\mathcal{A}V(x)^l \leq lV(x)^{l-1}\mathcal{A}V(x) + \chi V(x)^{l-1}\ .$$

By the calculation for $l = 1$,

$$\begin{aligned}
\mathcal{A}V(x)^l &\leq lV(x)^{l-1}[d - aV(x)] + \chi V(x)^{l-1} \\
&= -alV(x)^l + (ld + \chi)V(x)^{l-1}.
\end{aligned}$$

By choosing $a_l < al$ and d_l sufficiently large we obtain

$$\mathcal{A}V(x)^l \leq -a_l V(x)^l + d_l$$

as required. $\square$

Lemma 3.4 *For all $x \in \mathbb{T}^d \times \mathbb{R}^d, t > 0$ and open $\mathcal{O} \subset \mathbb{T}^d \times \mathbb{R}^d$, the transition kernel for (3.1)–(3.2) satisfies $P_t(x, \mathcal{O}) > 0$.*

Proof It suffices to consider the probability of hitting an open ball of radius δ, $\mathcal{B}_\delta$, centered at y^+. Consider the associated control problem, derived from (3.1)–(3.2),

$$\frac{dX}{dt} = Y(X) + \Sigma(X)\frac{dU}{dt} \ . \tag{3.11}$$

For any $t > 0$, any $y \in \mathbb{T}^d \times \mathbb{R}^d$, and any $y^+ \in \mathbb{T}^d \times \mathbb{R}^d$, we can find smooth $U \in C^1([0,t], \mathbb{R}^d)$ such that (3.11) is satisfied and $X(0) = y$, $X(t) = y^+$. To see this set $X = (Q^T, \frac{dQ}{dt}^T)^T$ and note that

$$\frac{d^2Q}{dt^2} + \gamma(Q)\frac{dQ}{dt} + \nabla F(Q) = \sigma(Q)\frac{dU}{dt} \ .$$

Choose Q to be a C^∞ path such that, for the given $t > 0$,

$$\begin{pmatrix} Q(0) \\ \frac{dQ}{dt}(0) \end{pmatrix} = y, \quad \begin{pmatrix} Q(t) \\ \frac{dQ}{dt}(t) \end{pmatrix} = y^+.$$

Since σ is everywhere invertible, $\frac{dU}{dt}$ is defined by substitution and will be as smooth as ∇F and σ^{-1} – hence C^∞. Also $U(0)$ can be taken as 0.

Note that the event

$$\sup_{0 \le s \le t} ||W(s) - U(s)|| \le \epsilon$$

occurs with positive probability for any $\epsilon > 0$, since the Wiener measure of any such tube is positive (Theorem 4.20 of [14]). From this it is possible to deduce the required open set irreducibility; see [15], Theorem 5.2 or [7], Lemma 3.4. □

4 Particles in a Random Velocity Field

We consider the following model for particles $x \in \mathbb{T}^2$, the two-dimensional torus, moving in a random velocity field:

$$\tau\ddot{x} = v(x,t) - \dot{x},$$
$$d\eta = \nu A\eta + dW,$$

where $v = \nabla^\perp\eta$, and

$$\nabla^\perp = (-\partial_{x_2}, \partial_{x_1})^T.$$

Thus x denotes the position of a particle moving according to Stokes' law in a 2D incompressible velocity field. Here $A = -\Delta$ with $D(A) = \dot{H}_{per}(\Omega)$ (periodic functions in $W^{2,2}(\Omega)$ with mean zero) and Ω the unit square $[0,1] \times [0,1]$. $W(t)$ is an infinite dimensional Wiener process expressible as a weighted sum of the product of eigenfunctions of A with standard i.i.d Brownian motions. In this simple model these weights are chosen to ensure a prescribed energy spectrum for the velocity field v in statistical equilibrium, allowing various charicatures of turbulence to be studied. The fact that statistical equilibrium should be expected can be understood by noting that the equation for η is an infinite dimensional Ornstein-Uhlenbeck process. Under appropriate assumptions on the covariance of the noise it is indeed ergodic. For a review of the literture related to such models, and for some mathematical analysis of the model, including formulation as a random dynamical system, existence of a random global attractor and the ergodicity of η see [13].

Our aim here is to show that the ergodicity of η induces ergodicity in the particles. We assume that the noise and initial data excite only a finite number of Fourier components. A more sophisticated analysis will be required for the infinite dimensional problem. To be concrete, out finite-dimensionality assumption is to assume that

$$\eta = \sum_{k \in \mathcal{K}} y_k \cos(k \cdot x) + z_k \sin(k \cdot x)$$

where

$$\mathcal{K} = \{2\pi(k_1, k_2), k_i \in \{1, \dots, M\}\}m$$

where $M \geq 2$. The form of the model means that y_k and z_k are scalar Ornstein-Uhlenbeck processes with, for fixed k, the same parameters, but different noises:

$$\begin{aligned} dy_k &= -\alpha_k y_k dt + \sqrt{\lambda}_k dB_k^y, \\ dz_k &= -\alpha_k z_k dt + \sqrt{\lambda}_k dB_k^z \end{aligned}$$

and the families $\{B_k^y\}_{k \in \mathcal{K}}, \{B_k^z\}_{k \in \mathcal{K}}$ are mutually independent families of i.i.d standard Brownian motions. We assume that

$$\alpha := \min_{k \in \mathcal{K}} \{\alpha_k\} > 0.$$

We set $N = M^2$ and define $u = (x^T, p^T, y^T, z^T) \in \mathbb{T}^2 \times \mathbb{R}^2 \times \mathbb{R}^N \times \mathbb{R}^N$ where the vector y (respectively z) contains the N y_k (respectively z_k). In this notation our problem becomes

$$du = Y(u)dt + \Sigma dB, \quad u(0) = v \tag{4.1}$$

where $B \in \mathbb{R}^{2N}$ is a vector of i.i.d standard Brownian motions. Here, for D an appropriate diagonal matrix made up of the α_k, we have

$$Y(u) = \left(\begin{array}{c} p \\ \frac{1}{\tau} \sum_{k \in \mathcal{K}} \{y_k \nabla^\perp \cos(k \cdot x) + z_k \nabla^\perp \sin(k \cdot x)\} - \frac{p}{\tau} \\ -Dy \\ -Dz \end{array} \right) \in \mathbb{R}^{2N+4},$$

and

$$\Sigma = \left(\begin{array}{cc} O_{2\times N} & O_{2\times N} \\ O_{2\times N} & O_{2\times N} \\ I_{N\times N} & O_{N\times N} \\ O_{N\times N} & I_{N\times N} \end{array} \right) \in \mathbb{R}^{2N+4 \times 2N}.$$

Here $O_{m\times n}$ (resp. $I_{m\times n}$) is the $m \times n$ zero (resp. identity) matrix. Note that

$$\Sigma dB = \sum_{k \in \mathcal{K}} \sqrt{\lambda}_k \{e_k^y dB_k^y + e_k^z dB_k^z\}$$

where $e_k^y = (0^T, 0^T, e_k^T, 0^T)^T, e_k^z = (0^T, 0^T, 0^T, e_k^T)^T$ and e_k is the unit vector in $\mathbb{R}^N$ with non-zero entry only at the entry corresponding to index k.

Global in time existence and uniqueness for (4.1) follows from the fact that all vector fields are globally Lipschitz on $\mathbb{X}$. For (4.1) it is useful to define the Lyapunov function

$$V(u) = 1 + \frac{1}{2}\{\|y\|^2 + \|z\|^2 + a\|p\|^2\} \tag{4.2}$$

where $a > 0$ is a constant to be determined.

Now we define

$$\mathcal{G}_l = \{\text{measurable } g : \mathbb{T}^2 \times \mathbb{R}^{2N+2} \to \mathbb{R} \text{ with } |g| \leq V^l\}.$$

Theorem 4.1 *The SDE* (4.1) *with* $u(t) = (x(t)^T, p(t)^T, y(t)^T, z(t)^T)^T$ *has a unique invariant measure* π *on* $\mathbb{T}^2 \times \mathbb{R}^{2N+2}$. *Furthermore there is an* $a > 0$ *for which the following holds. Fix any* $l \geq 1$. *There exists* $C = C(l) > 0$, $\lambda = \lambda(l) > 0$ *such that, for all* $g \in \mathcal{G}_l$, $u(0) = v$,

$$\left|\mathbb{E}^v g(u(t)) - \pi(g)\right| \leq CV(u)^l e^{-\lambda t} \quad \textit{for all } t \geq 0. \tag{4.3}$$

Proof The result follows from an application of Corollary 2.8. First note that $V(u)^l \to \infty$ as $||u|| \to \infty$. Lemma 4.2 shows that there is a choice of $a > 0$ such that, if $\mathcal{A}$ is the generator of the process governed by (4.1), then

$$\mathcal{A}\{V(x)^l\} \leq -a_l\{V(x)^l\} + d_l$$

for some $a_l, d_l > 0$. Thus Assumption 2.4 (the Lyapunov condition) holds.

Lemma 4.3 proves that, in any positive time, any open set may be reached with positive probability. Thus Assumption 2.6(i) holds with any choice of C and y^*. We prove below that Assumption 2.6(ii) holds and so we have the minorization condition by Lemma 2.7. Hence Corollary 2.8 gives the desired result.

It remains to establish Assumption 2.6(ii). We observe that

$$f_k^y := [Y, e_k^y] = \begin{pmatrix} 0 \\ \frac{1}{\tau}\nabla^\perp \cos(k \cdot x) \\ -\alpha_k e_k \\ 0 \end{pmatrix},$$

$$f_k^z := [Y, e_k^z] = \begin{pmatrix} 0 \\ \frac{1}{\tau}\nabla^\perp \sin(k \cdot x) \\ 0 \\ -\alpha_k e_k \end{pmatrix}$$

and then that

$$[Y, f_k^y] = \begin{pmatrix} \frac{1}{\tau}\nabla^\perp \cos(k \cdot x) \\ \times \\ \times \\ \times \end{pmatrix},$$

$$[Y, f_k^z] = \begin{pmatrix} \frac{1}{\tau}\nabla^\perp \sin(k \cdot x) \\ \times \\ \times \\ \times \end{pmatrix}.$$

Here $\times$ denotes entries immaterial in the following. Together

$$\bigcup_{k \in \mathcal{K}} \{e_k^y, e_k^z, [Y, e_k^y], [Y, e_k^z], [Y, f_k^y], [Y, f_k^y]\}$$

span $\mathbb{R}^{2N+4}$ because

$$\nabla^\perp \cos(k \cdot x) = \begin{pmatrix} -k_2 \\ k_1 \end{pmatrix} \sin(k \cdot x), \nabla^\perp \sin(k \cdot x) = \begin{pmatrix} k_2 \\ -k_1 \end{pmatrix} \cos(k \cdot x).$$

Since $M \geq 2$ there are at least two distinct directions generated by the vectors k, and since at least one of $\sin(k \cdot x)$ and $\cos(k \cdot x)$ is non-zero at every point of $\mathbb{T}^2$, the result follows. $\square$

The previous theorem requires the following two lemmas:

Lemma 4.2 *For every $l \geq 1$, there exists $a_l \in (0, \infty)$ and $d_l \in (0, \infty)$ such that, for equation* (4.1) *with $\mathcal{A}$ given by* (3.10),

$$\mathcal{A}\{V(u)^l\} \leq -a_l\{V(u)^l\} + d_l \ .$$

Proof For any $\delta > 0$

$$\langle Y(u), \nabla V(u) \rangle \leq -\frac{a}{\tau}\|p\|^2 + \frac{a\kappa}{\tau}[\frac{1}{2\delta}\|y\|^2 + \frac{1}{2\delta}\|z\|^2 + \delta N\|p\|^2] - \alpha\|y\|^2 - \alpha\|z\|^2,$$

where

$$\alpha = \min_{k \in \mathcal{K}} \alpha_k > 0, \quad \kappa = \max_{k \in \mathcal{K}} \{\|\nabla^\perp \cos(k \cdot x)\|, \|\nabla^\perp \sin(k \cdot x)\|\}.$$

Choose δ so that $2\delta N\kappa = 1$ and then choose a so that $a\kappa = \alpha\tau\delta$. It follows that

$$\langle Y(u), \nabla V(u) \rangle \leq -\frac{a}{2\tau}\|p\|^2 - \frac{\alpha}{2}\|y\|^2 - \frac{\alpha}{2}\|z\|^2.$$

Using the fact that Σ is constant we deduce that

$$\mathcal{A}V(u) \leq -a_1 V(u) + d_1$$

with $a_1 = \min\{\frac{a}{\tau}, \alpha\}$.

Now we calculate $\mathcal{A}\{V(u)^l\}$. To this end, note that

$$
\begin{aligned}
\frac{\partial}{\partial u_i}\left\{V(u)^l\right\} &= l\{V(u)\}^{l-1}\frac{\partial V}{\partial u_i}\\
\frac{\partial^2}{\partial u_i \partial u_j}\left\{V(u)^l\right\} &= \frac{\partial}{\partial u_j}\left\{l\{V(u)\}^{l-1}\frac{\partial V}{\partial u_i}\right\},\\
\text{and } \mathcal{A}\{V(u)^l\} &= l\{V(u)\}^{l-1}\mathcal{A}V + \frac{1}{2}\sum_{i=1}^{N} l(l-1)V(u)^{l-2}\left(\frac{\partial V}{\partial u_{4+i}}\right)^2\\
&\quad + \frac{1}{2}\sum_{i=1}^{N} l(l-1)V(u)^{l-2}\left(\frac{\partial V}{\partial u_{4+N+i}}\right)^2.
\end{aligned}
$$

But the quadratic structure of V means that, for some $\chi > 0$,

$$\mathcal{A}V(u)^l \leq lV(u)^{l-1}\mathcal{A}V(u) + \chi V(u)^{l-1}\ .$$

By the calculation for $l = 1$,

$$
\begin{aligned}
\mathcal{A}V(u)^l &\leq lV(u)^{l-1}[d - aV(u)] + \chi V(u)^{l-1}\\
&= -alV(u)^l + (ld + \chi)V(u)^{l-1}.
\end{aligned}
$$

By choosing $a_l < al$ and d_l sufficiently large we obtain

$$\mathcal{A}V(u)^l \leq -a_l V(u)^l + d_l$$

as required. $\square$

Lemma 4.3 *For all $u \in \mathbb{T}^d \times \mathbb{R}^d, t > 0$ and open $\mathcal{O} \subset \mathbb{T}^d \times \mathbb{R}^d$, the transition kernel for (4.1) satisfies $P_t(u, \mathcal{O}) > 0$.*

Proof To start with we just consider the probability of reaching an open set in the coordinates concerned with particle position and momentum $(x, p) \in \mathbb{T}^2 \times \mathbb{R}^2$. At the end we describe the simple modification to the full-space $\mathbb{T}^2 \times \mathbb{R}^{2N+2}$.

We first study the control problem of connecting $(x(0), p(0)) = (x^-, p^-)$ with $(x(t), p(t)) = (x^+, p^+)$. Note that, for any $k \in \mathcal{K}$, either $\cos(k \cdot x) \neq 0$ or $\sin(k \cdot x) \neq 0$ for either $x = x^-$ or $x = x^+$. Let $k^1, k^2 \in \mathcal{K}$ be two non-parallel vectors with $k^i = (k^i_1, k^i_2)^T$. For simplicity we assume that $\cos(k^1 \cdot x), \cos(k^2 \cdot x) \neq 0$ for $x = x^-$ and $x = x^+$. Similar arguments to the following can be used by replacing cos by sin in the following and, possibly, by piecing together two arguments of the same type; we discuss this below.

Now let $x(s), 0 \leq s \leq t$, be any smooth path connecting (x^-, p^-) and (x^+, p^+) and avoiding the set

$$\{x \in \mathbb{T}^2 : \cos(k^1 \cdot x) = 0 \text{ or } \cos(k^2 \cdot x) = 0\}. \tag{4.4}$$

Then the control problem to find $U_1(t), U_2(t) \in C^\infty([0, t], \mathbb{R}^2)$ such that the differential equation

$$\tau\ddot{x} + \dot{x} = U_1(t)\nabla^\perp \sin(k^1 \cdot x) + U_2(t)\nabla^\perp \sin(k^2 \cdot x) \tag{4.5}$$

passes along this path, $0 \le s \le t$, is uniquely solvable. This follows because

$$\left(\begin{array}{cc} k_2^1 \cos(k^1 \cdot x) & k_2^2 \cos(k^2 \cdot x) \\ -k_1^1 \cos(k^1 \cdot x) & -k_1^2 \cos(k^2 \cdot x) \end{array} \right) \left(\begin{array}{c} U_1(t) \\ U_2(t) \end{array} \right) = \tau \ddot{x} + \dot{x}$$

and the matrix has determinant which is non-zero because it avoids the set (4.4) and because k^1 and k^2 are not parallel.

If both $\sin(k^1 \cdot x)$ and $\sin(k^2 \cdot x)$ are non-zero at both end-points then a similar construction works using the equation

$$\tau \ddot{x} + \dot{x} = U_1(t) \nabla^\perp \cos(k^1 \cdot x) + U_2(t) \nabla^\perp \cos(k^2 \cdot x)$$

and a smooth path avoiding points where $\sin(k^1 \cdot x)$ and $\sin(k^2 \cdot x)$ disappear. Similar arguments work if $\sin(k^1 \cdot x)$ and $\cos(k^2 \cdot x)$ are non-zero at both end-points or if $\sin(k^2 \cdot x)$ and $\cos(k^1 \cdot x)$ are non-zero at both end-points. However, if the non-zero quantities differ at the two ends of the path then it is necessary to break the controls $U_1(s), U_2(s)$ into two smooth paths, one for $s \in [0, t/2]$ and one for $s \in [t/2, t]$. Choosing the smooth path x so that none of

$$\sin(k^1 \cdot x), \sin(k^2 \cdot x), \cos(k^1 \cdot x), \cos(k^2 \cdot x)$$

disappear at $x = x(t/2)$, and then using two arguments of the preceeding type, shows that a piecewise smooth path may be chosen controlling x from (x^-, p^-) to (x^+, p^+).

Ornstein-Uhlenbeck processes can be forced to lie within a small tubular neighbourhood of any smooth curve, using the analogous property for Brownian motion. This completes the open set irreducibility in $\mathbb{T}^2 \times \mathbb{R}^2$. However, to extend the result to $\mathbb{T}^2 \times \mathbb{R}^{2N+2}$ we need to allow the Ornstein-Uhlenbeck processes themselves to arrive into any open set at time t. For this it is necessary to modify the paths $U_1(s)$ and $U_2(s)$ in the neighbourhood of $s = t$ so that they end at a prescribed point. By making this change on a sufficiently small interval near t, and using continuity of (4.5) or related control problems in $U_1(t), U_2(t)$, we can still ensure the required reachability in $\mathbb{T}^2 \times \mathbb{R}^2$, as well as in the y and z variables. Theorem 5.2 of [15] gives the desired result, noting that for this problem the Ito and Stratonovich forms of the SDE are the same. $\square$

Acknowledgements: We would like to thank Des Higham, and Wilhelm Huisinga for helpful input.

References

[1] D. R. Bell. *The Malliavin Calculus.* Longman Scientific & Technical, Harlow, 1987.

[2] R. Durrett, *Stochastic Calculus, A Practical Introduction.* CRC Press, 1996.

[3] R. Z. Has'minskii. *Stochastic Stability of Differential Equations.* Sijthoff and Noordhoff, 1980.

[4] H. Kunita. Supports of diffusion processes and controllability problems. In *Proceedings of the International Symposium on Stochastic Differential Equations (Res. Inst. Math. Sci., Kyoto Univ., Kyoto, 1976)*, pages 163–185, New York, 1978. Wiley.

[5] A. Lasota and M. Mackey, *Chaos, Fractals and Noise.* Springer-Verlag, New York, 1994.

[6] X. Mao, *Stochastic Differential Equations and Applications.* Horwood, Chichester, 1997.

[7] J. Mattingly, A.M. Stuart and D.J. Higham, *Ergodicity for SDEs and approximations: locally lipschitz vector fields and degenerate noise.* Submitted to Stoch. Proc. and Applics.

[8] S.P. Meyn and R.L. Tweedie, *Stochastic Stability of Markov Chains.* Springer-Verlag, New York, 1992.

[9] S.P. Meyn and R.L. Tweedie, *Stability of Markovian processes, I, II and III* Adv. Appl. Prob. **24**(1992), 542–574, **25**(1993), 487–517 and **25**(1933), 518–548.

[10] E. Nummelin, *General Irreducible Markov Chains and Noise.* Cambridge University Press, 1984.

[11] L.C.G. Rogers and D.Williams, *Diffusions, Markov processes and Martingales, Volume 2.* Cambridge University Press, reprinted second edition, 2000.

[12] T. Shardlow and A.M. Stuart, *A perturbation theory for ergodic Markov chains with application to numerical approximation.* SIAM J. Num. Anal.**37**(2000), 1120–1137.

[13] H. Sigurgeirsson and A.M. Stuart. *Particles Obeying Stokes' Law in a Random Velocity Field.* In preparation.

[14] D. W. Stroock. *Lectures on Topics in Stochastic Differential Equations.* Tata Institute of Fundamental Research, Bombay, 1982. With notes by Satyajit Karmakar.

[15] D.W.Strook and S.R.S. Varadhan, *On the support of diffusion processes with applications to the strong maximum principle.* Proc. Sixth Berkeley Symp. on Math. Stat. and Prob., **III**, 333–360.

[16] M.M. Tropper, *Ergodic properties and quasideterministic properties of finite-dimensional stochastic systems.* J. Stat. Phys. **17**(1977), 491–509.