Metastability of Waves, Patterns and Oscillations Subject to Spatially-Extended Noise

J. MacLaurin

June 24, 2020

Abstract

In this paper we present a general framework in which one can rigorously study the effect of spatio-temporal noise on traveling waves, stationary patterns and oscillations that are invariant under the action of a finite-dimensional set of continuous isometries (such as translation or rotation). This formalism can accommodate patterns, waves and oscillations in reaction-diffusion systems and neural field equations. To do this, we define the phase by precisely projecting the infinite-dimensional system onto the manifold of isometries. Two differing types of stochastic phase dynamics are defined: (i) a variational phase, obtained by insisting that the difference between the projection and the original solution is orthogonal to the non-decaying eigenmodes, and (ii) an isochronal phase, defined as the limiting point on manifold obtained by taking $t \to \infty$ in the absence of noise. We outline precise stochastic differential equations for both types of phase. The variational phase SDE is then used to show that the probability of the system leaving the attracting basin of the manifold after an exponentially long period of time (in $\epsilon^{-2}$, the magnitude of the noise) is exponentially unlikely. In the case that the manifold is periodic (such as for spiral waves, spatially-distributed oscillations, or neural-field phenomena on a compact domain), the isochronal phase SDE is used to determine asymptotic limits for the average occupation times of the phase as it wanders in the basin of attraction of the manifold over very long times. In particular, we find that frequently the correlation structure of the noise biases the wandering in a particular direction, such that the noise induces a slow oscillation that would not be present in the absence of noise.

1 Introduction

Spatially-extended patterns and waves are ubiquitous in the biological and physical sciences and are a key lens through which emergent phenomena are understood. Furthermore biology is typically very noisy, and thus it is of great importance to understand the effect of stochasticity on these patterns and waves [75, 60, 79]. The literature on stochastic patterns and waves includes general Turing patterns [10], the Allen-Cahn / Cahn-Hilliard equation [43], waves and patterns in the stochastic Brusselator [8, 9], patterns in neural fields [49, 38, 56, 41, 85, 2, 62, 16, 70], interfaces in the Ginzburg-Landau equation [13, 48], the stochastic burger’s equation [12] and the effect of spatially-distributed noise on traveling waves [72, 123], such as the FKPP traveling waves [29, 20], invasion waves in ecology [64], the stochastic Nagumo equation [59, 47, 39], geometric waves [90] and numerical methods for stochastic traveling waves [58]. Good reviews of the literature on the effect of noise on traveling waves can be found in [75, 79, 60].

Consider the deterministic unforced system with solution $u_t \in C_b(\mathbb{R}^d, \mathbb{R}^N)$ (the Banach space of continuous bounded $\mathbb{R}^N$-valued functions over $\mathbb{R}^d$, for $N,d \geq 1$)

$$\frac{du_t(x)}{dt} = Au_t(x) + f(u_t)(x) \quad (1.1)$$
Here $A$ is a linear operator, such as the Laplacian, and $f$ is a continuous function that is locally bounded. Frequently the fixed points of such systems possess several symmetries (typically invariance under translation and/or rotation), and there exists a smooth attracting manifold of fixed points (parameterized by $(\phi_\alpha)_{\alpha \in \mathbb{R}^m} \subset C_0(\mathbb{R}^d, \mathbb{R}^N)$) such that

$$A\phi_\alpha + f(\phi_\alpha) = 0, \forall \alpha \in \mathbb{R}^m. \quad (1.2)$$

The thrust of this article is to drive (1.1) by space-time white noise and understand how the noise shapes the dynamics in the basin of attraction of the manifold $(\phi_\alpha)_{\alpha \in \mathbb{R}^m}$. As such, we consider the following stochastic evolution equation driven by space-time noise $W_t$:

$$du_t = [Au_t + f(u_t)]dt + \varepsilon B(t, u_t)dW_t, \quad (1.3)$$

for some $\varepsilon > 0$. In [50] we developed a method of precisely projecting the stochastic dynamics onto the manifold parameterized by $(\phi_\alpha)_{\alpha \in \mathbb{R}^m}$, generalizing deterministic work on the orbital stability of a manifold of fixed points (see [88, Chapter 5] and [51, Chapter 4]) to a stochastic setting. This paper builds on the work of [50] and has three main aims: (i) to obtain an SDE yielding a precise projected phase dynamics, (ii) obtain accurate probabilistic exit-time estimates over long periods of time and (iii) to obtain long-time ‘occupation-time’ estimates for the wandering of the stochastic phase on the manifold $\mathbb{R}^m$. In more detail:

- In [50] the manifold was taken to be one-dimensional; in this paper the manifold is multidimensional.
- In [50] the stability estimates in Section 6 require the linearization of (2.2) to be self-adjoint and immediately contractive. This assumption is relaxed, and now the main requirement is that it possesses a spectral gap (except for the neutral eigenmodes tangential to the manifold in (2.2)).
- In [50] a ‘variational phase’ SDE was obtained. In this work an SDE for the variational phase is also obtained, and an additional ‘isochronal phase’ SDE is also obtained.
- In [50] the noise was additive and a ‘Q-Wiener process’ [27], meaning in effect that there is sufficient spatial correlation that the space-time noise is spatially-continuous. In this paper we consider multiplicative noise, and when the operator $A$ is sufficiently smoothing (such as when $A$ is the sum of a Laplacian and possibly first order derivatives), the noise can be ‘cylindrical’, meaning that the driving noise is spatially-decorrelated and spatially-discontinuous.
- The bound on the growth of the error in [50, Corollary 6.3] is suboptimal. This bound was greatly improved in [70] for ‘bumps’ of activity in stochastic neural fields. In this work we use a similar method to [70] to show that the probability of the system leaving the manifold of translated bump solutions over an exponentially long period of time (i.e. $T \simeq \exp(C\varepsilon^{-2})$) is exponentially unlikely. In other words, we are in the Large Deviation regime [80].

After one has shown that, with very high probability, the system spends a very long period of time in the neighborhood of the manifold, it is natural to investigate the ‘occupation time’: that is, the typical proportion of time that it spends in particular neighborhoods of the manifold over long time intervals (see [30] for a classical result on the occupation time of Markov chains). Doing this yields an understanding of how the correlation structure of the noise interacts with the geometry of manifold $(\phi_\alpha)_{\alpha \in \mathbb{R}^m}$ to shape the wandering of the stochastic phase over long periods of time. Under the assumption that the manifold is periodic, we find that the natural timescale for the induced phase dynamics on the manifold is $\varepsilon^{-2}t$ ($\varepsilon$ being the magnitude of the white noise), and we are then able to demonstrate that the occupation time converges to that of

*A cylindrical $H$-valued Wiener process.*
the invariant measure of the rescaled process (modulo the periodicity). In many cases this result implies that the noise correlation structure can induce a slow oscillation in the phase over long periods of time. To demonstrate the convergence of the occupation time, we must employ a slightly different phase that is analogous to the isochronal phase definition for stochastic oscillators \[34\]. It is worth comparing the results of this paper to those of Blomker \[11\]. He uses a rescaling of time to understand the projection of fluctuations in SPDEs onto a manifold that is weakly unstable (by contrast, this paper concerns a projection onto a stable manifold of fixed points).

To the best of this author’s knowledge, the first rigorous work on the behavior of stochastic systems near an attracting manifold is that of Katzenberger \[55\]. He determined limiting equations for a finite-dimensional jump Markov process pulled onto a manifold by a large drift. This theory has been applied in finite-dimensional stochastic models of population dynamics, including in \[23, 76, 43, 78\]. In the deterministic literature, there is a well developed literature on the orbital stability of finite-dimensional submanifolds in infinite-dimensional spaces, including work on the Ginzburg-Landau equation \[7\], traveling waves \[51, 88\], and spiral waves \[82\]: see \[88, 51\] for many more examples.

It is well-established \[11, 73, 22, 20, 23, 14\] that for traveling waves perturbed by space-time noise, one can identify the leading order diffusive flux of the wave position \(\beta_t\) over timescales of \(O(\epsilon^{-2})\) by matching the leading order terms in the equation

\[
\langle u_t - \varphi_{\beta_t}, \psi_{\beta_t} \rangle \approx 0. \tag{1.4}
\]

Here \(\psi_{\beta_t}\) is the neutral eigenvector of the adjoint of the linearization about the traveling wave positioned at \(\beta_t\). This is directly motivated by deterministic theory (see Chapter 4 of \[51\] and Chapter 5 of \[88\]). In stochastic systems, this technique has been employed in reaction-diffusion traveling fronts by \[11, 73, 20, 79\], in stochastic neural fields by Bressloff and co-workers \[14, 50, 70\] and Kilpatrick and Ermentrout \[56\]. Hamster and Hupkes \[45\] demonstrate that the probability of leaving a neighborhood of the manifold over an exponentially long period of time goes to zero as the noise strength goes to zero. The work of Cartwright and Gottwald \[23\] is interesting because the manifold they project onto also includes a non-neutral eigenmode, and there is therefore potential for their ansatz to be accurate over longer periods of time.

Despite these strengths, none of the works in the previous paragraph (with the exception of \[70\]) determine precise expressions for the quadratic variation (i.e. \(d\beta_t d\bar{\beta}_t\)) and cross-variation (i.e. \(d\beta_t du_t\)) terms that would be necessary for \[14\] to be satisfied exactly, and this creates problems once one wishes to precisely understand the phase dynamics over longer timescales (that diverge on timescales greater than \(O(\epsilon^{-2})\)). Indeed in the deterministic setting, one knows from the implicit function theorem that \[14\] can be solved for \(\beta_t\) in some neighborhood only if \(\partial/\partial \beta_t \{\langle u_t - \varphi_{\beta_t}, \psi_{\beta_t} \rangle \} \neq 0\), and if \(\partial/\partial \beta_t \{\langle u_t - \varphi_{\beta_t}, \psi_{\beta_t} \rangle \} \) asymptotes to zero as \(t\) approaches some limit, then the coefficients in the ODE for \(\beta_t\) will blow up. It was demonstrated in \[50\] that this blowup can also occur in the stochastic setting: in the SDE for the phase \(d\beta_t\), both the drift and the diffusion coefficient blowup if \(\partial/\partial \beta_t \{\langle u_t - \varphi_{\beta_t}, \psi_{\beta_t} \rangle \} \) approaches zero. One cannot \textit{a priori} rule out \(\partial/\partial \beta_t \{\langle u_t - \varphi_{\beta_t}, \psi_{\beta_t} \rangle \} \) asymptoting to zero in the stochastic setting (by definition, noisy systems exhibit a diversity of behavior, with various degrees of probability); instead one must try to derive accurate bounds on the probability of this not occurring. Indeed recent work by this author and Bressloff \[70\] has solved \[14\] exactly for a neural field equation, and used this to prove that the probability of the system leaving a neighborhood of the manifold of bump solutions after an exponentially long period of time (i.e. \(T \simeq \exp(C \epsilon^{-2})\)) is exponentially small (i.e. \(\text{Prob} \simeq \exp(-C \epsilon^{-2})\)). In other words, instead of just determining the statistics for the leading order diffusion of the wavefront, MacLaurin and Bressloff determine an exact nonlinear SDE, coupled in the same space as the driving noise, and which is accurate over very long timescales. In summary, for a rigorous and accurate long-time expression for the phase SDE, one desires (i) an exact solution to \[14\] (we also determine an ‘isochronal phase’ in Section , that agrees with \[14\] to leading order), and (ii) control over \(\partial/\partial \beta_t \{\langle u_t - \varphi_{\beta_t}, \psi_{\beta_t} \rangle \} \) going to zero (this is encapsulated in

3
the stopping time $\tau$) and (iii) strong bounds on the fluctuations of the component of the solution orthogonal to the manifold.

Prior to the work of this author in [50], Stannat and co-workers in [84-59] (for reaction-diffusion systems) and [58] (for neural field systems) approximated the dynamics of the phase $\beta_t$ by an ordinary differential equation with $\frac{d\beta_t}{dt}$ proportional to $\frac{\partial}{\partial \beta_t} \|u_t - \varphi_{\beta_t}\|^2$. They then decomposed $u_t - \varphi_{\beta_t}$ into an Ornstein-Uhlenbeck Process plus a remainder term of lower order in $\epsilon$. They rigorously proved that this ansatz is accurate in the small $\epsilon$ limit.

There exists a literature on the first-exit-time from attracting wells in infinite-dimensional stochastic systems [0-3], generalizing the classical Kramer’s Law (see [28]) to an infinite dimensional setting. Usually one obtains these estimates through proving a Large Deviations principle for the system [87, 83, 24, 61, 80, 90]. A helpful explanation of how a Large Deviations Principles usually implies that a system spends an exponentially long period of time in the neighbourhood of the attracting fixed point can be found in [28, Section 5.6]: these estimates were originally derived by Freidlin and Wentzell [37]. A recent preprint of Salins and Spiliopoulos [81] has determined the exponential asymptotics of the first exit time from the attracting basin of a fixed point in an SPDE. By contrast, the asymptotics in this paper study the first exit time from the attracting basin of a smooth manifold of fixed points.

The organization of the paper is as follows. In Section 2 we describe the general setting we consider, and outline the necessary assumptions. Section 3 then goes on to list some examples that fit into the general setting, including reaction-diffusion traveling waves, spiral waves, spatially-distributed oscillations and patterns in neural fields. In Section 4 we define two different phase equations: the variational phase, obtained by defining the position of the pattern / wave to be such that the orthogonal amplitude is precisely perpendicular to the neutral eigenvectors of the adjoint operator, and the isochronal phase, obtained by removing noise from the system and taking $t \to \infty$. In Section 5 we demonstrate that the probability of the system leaving a neighborhood of the manifold before an exponentially long period of time is exponentially unlikely. Finally in Section 6 assuming the manifold to be periodic, we determine the limiting probability distribution for the phase (modulo the periodicity), and the average shift induced in the phase by the noise over long periods of time.

Notation: As usual, $C(\mathbb{R}^d)$ and $C^\infty(\mathbb{R}^d)$ will denote the spaces of real-valued functions on $\mathbb{R}^d$ that are continuous and smooth respectively. Moreover $L^p(\mathbb{R}^d)$ ($p \geq 1$), will be the space of $p$-integrable functions with respect to the Lebesgue measure on $\mathbb{R}^d$. Finally, for general Banach spaces $E_1, E_2$, we will denote by $\mathcal{L}(E_1, E_2)$ the space of bounded linear operators $E_1 \to E_2$. $\mathcal{L}(E_1, E_2)$ is equipped with the operator norm.

Let $H := [L^2(\mathbb{R}^d)]^N$, equipped with the standard inner product denoted by $\langle \cdot, \cdot \rangle$ and norm $\| \cdot \|$. Let $\mathcal{L}_{HS}$ be the space of all linear Hilbert-Schmidt operators $H \to H$, with the norm of $B \in \mathcal{L}_{HS}$ written as

$$\|B\|_{HS} = \sum_{j=1}^{\infty} \langle e_j, Be_j \rangle,$$

(1.5)

where $\{e_j\}_{j \in \mathbb{Z}^+}$ is any orthonormal basis for $H$. The operator norm of any $U \in \mathcal{L}(H, H)$ is written as

$$\|U\|_\infty = \sup \{ \|Uz\| : \|z\| = 1 \}.$$

(1.6)

Let $E := \varphi_0 + H_0$ (i.e. $u \in E$ if and only if $u = \varphi_0 + v$ for some $v \in H_0$), endowed with the topology inherited from $H_0$. For $y, z \in E$, with $y = \varphi_0 + y_0$ and $z = \varphi_0 + z_0$, we write $\|y - z\|_0 := \|y_0 - z_0\|_0$. 

4
2 Problem Setup

We consider the system in (1.1), forced by multiplicative spatially-extended white noise. This yields an \( E \)-valued stochastic differential equation \( (E = \varphi_0 + H) \), where

\[
du_t = [Au_t + f(u_t)]dt + \varepsilon B(t, u_t)dW_t,
\]

for some \( \varepsilon > 0 \). For each \( t \geq 0 \) and \( u_t \in E, B(t, u_t) \) is a linear operator on \( H \); it specifies the multiplicative effect of the noise. We take the initial condition to be \( u_0 := \varphi_\beta \) for some constant \( \beta \in \mathbb{R}^m \) (one could easily obtain analogous results by taking \( u_0 \) to be in a small neighborhood about \( \varphi_\beta \)). \( A \) is a linear operator (such as the Laplacian, for reaction-diffusion systems) and \( f \) is a Lipschitz nonlinear function on \( E \). As noted in the introduction, we assume that \( \{\varphi_\alpha\}_{\alpha \in \mathbb{R}^m} \) constitute a manifold of fixed points, such that

\[
A\varphi_\alpha + f(\varphi_\alpha) = 0, \quad \forall \alpha \in \mathbb{R}^m.
\]

\( W_t \) is ‘space-time white noise’, such that formally, for any \( v, y \in H \),

\[
\mathbb{E}\langle [v, W_t] \rangle = 0, \quad \mathbb{E}\langle [v, W_t]\langle y, W_t \rangle \rangle = t(v, y).
\]

Some care is required to make precise sense of (i) the definition of the white noise in (2.3) and (ii) what one means by a solution to (2.1). If the white noise were to be completely spatially decorrelated, then one cannot take a spatial derivative, and one therefore cannot use standard partial differential equation theory to make sense of a solution to (2.1). Indeed one cannot even properly define an \( H \)-valued Gaussian random variable satisfying the properties in (2.3). However if \( A \) is the Laplacian, then one knows that it works to smooth functions. Indeed since \( \frac{d^2}{dx^2} \cos(ax) = -a^2 \cos(ax) \) and \( \frac{d^2}{dx^2} \sin(ax) = -a^2 \sin(ax) \), one knows that the highly fluctuating components of a solution get strongly damped by the Laplacian. Thus the solution that we are going to define can be thought of as taking the limit of (2.1) for increasingly decorrelated noise, but such that a sensible limit is obtained because high wavenumber fluctuations get damped down by the smoothing action of \( A \). The theory of stochastic partial differential equations has been developed to make precise sense of this limit [27, 67].

Formally, we define \( W_t \) to be a cylindrical \( H \)-valued Wiener process on the filtered probability space \((\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})\). The solution described in the previous paragraph can be precisely defined by employing a stochastic analog of the variation-of-constants solution in PDEs, as stated in the following proposition.

**Proposition 2.1.** Assume the assumptions of Section 2.1. Then stochastic evolution equation (2.1) has a unique mild solution, which can be decomposed (in a non-unique way) as \( u_t = \varphi_\beta + v_\beta^t \) where \((v_\beta^t)_{t \geq 0}\) is the unique weak (and mild) \( H \)-valued solution to

\[
dv_\beta^t = [Av_\beta^t + f(\varphi_\beta + v_\beta^t) - f(\varphi_\beta)]dt + \varepsilon B(t, u_t)dW_t, \quad t \geq 0,
\]

with initial condition \( v_\beta^0 = 0 \) i.e.

\[
v_\beta^t = \int_0^t P_{t-s} [f(\varphi_\beta + v_\beta^s) - f(\varphi_\beta)] ds + \varepsilon \int_0^t P_{t-s} B(s, u_s)dW_s, \quad t \geq 0.
\]

and \((P_t^A)_{t \geq 0}\) is the semigroup generated by \( A \).

**Proof.** The proof of this result is a straightforward application of [27, Theorem 7.4] using the globally Lipschitz assumption on \( f \) (Assumption (ii)), the fact that \( A \) generates a \( C_0 \)-semigroup on \( H \) (Assumption (i)) and the assumptions on \( B \) above.
Remark 1. We remark that for traveling waves, (2.1) is in the the moving coordinate frame. To illustrate what we mean by this, suppose again we are in the concrete situation of the standard neural field equation described in Section 3.2, so that there is a solution \( \hat{u}(x - ct) \) to (3.5) for some speed \( c \). The stochastic version of this equation with purely additive noise would then be

\[
du_t = [Au_t + w * F(u_t)]dt + B(t, u_t)dW_t,
\]

where as above \( Au = cu' \), \( u \in \mathcal{D}(A) \) and now \( \tilde{B}(t, u_t)w := B(\tilde{w}) \), \( \tilde{w}(x) = w(x + ct) \) for \( w \in E \).

2.1 General Assumptions

Our assumptions on the drift dynamics are intended to resemble the assumptions in the deterministic theory outlined in [51, Chapter 4] as much as possible. We thus assume that the manifold of fixed points of (2.2) can be obtained by applying a smooth isometry \( T_\alpha : E \rightarrow E \), in the following manner,

\[
\varphi_\alpha = T_\alpha \varphi_0.
\]  

We employ the following assumptions on the family of isometries \( \{T_\alpha\} \).

Assumption 1. For \( \alpha, \beta \in \mathbb{R}^m \),

\[
T_\alpha \circ T_\beta = T_{\alpha + \beta} \tag{2.5}
\]

\[
A \cdot T_\alpha = T_\alpha \cdot A \text{ and } f(T_\alpha \cdot u) = T_\alpha \cdot f(u). \tag{2.6}
\]

The following assumption on the generator is satisfied in the vast majority of interesting applications, including for \( A \) being an elliptic operator (for waves and patterns in reaction diffusion systems), or hyperbolic (for traveling / rotating waves in neural field equations, in the co-moving reference frame).

Assumption 2. The domain of \( A \), i.e. \( \mathcal{D}(A) \), is dense in \( H \), and the restriction of \( A \) to \( H \) (also denoted by \( A \)) is the generator of a \( C_0 \)-semigroup \( P_t^A \) on \( H \).

Assumption 3. Assume that the nonlinear function \( f \) acting in \( E \) is such that:

(i) \( f \) is defined on all of \( E \), and for all \( u \in E \) there exists the Frechet Derivative for perturbations in \( H \), written \( Df(u) \in L(H, H) \), i.e. such that for all \( v \in H \),

\[
\lim_{h \to 0} \frac{f(u + hv) - f(u)}{h} - Df(u) \cdot v = 0,
\]

and \( u \to DF(u) \) is continuous.

(ii) For all \( u \in E \) there exists the second Frechet Derivative \( D^{(2)}f(u) \in L(H \times H, H) \) such that for all \( v, w \in H \),

\[
\lim_{h \to 0} \frac{Df(u + hv) \cdot v - Df(u) \cdot v}{h} - D^{(2)}f(u) \cdot v \cdot w = 0,
\]

and \( u \to D^{(2)}f(u) \) is continuous.

(iii) The third Frechet Derivative \( D^{(3)}f(u) \in L(H \times H \times H, H) \) exists, is such that \( u \to D^{(3)}f(u) \) is continuous.
(iv) \( \sup_{u \in E} \| Df(u) \|_{L(H,H)} < \infty \) (so that \( H \ni v \mapsto f(\varphi_{\alpha} + v) \) is globally Lipschitz \( \forall \alpha \in \mathbb{R}^m \));

(v) 
\[
\sup_{u \in E, v, w \in H, \| v \|, \| w \| \leq 1} \| D^{(2)} f(u) \cdot v \cdot w \| < \infty \tag{2.7}
\]
\[
\sup_{u \in E, v, w, z \in H, \| v \|, \| w \|, \| z \| \leq 1} \| D^{(3)} f(u) \cdot v \cdot w \cdot z \| < \infty. \tag{2.8}
\]

In order that we can project the stochastic dynamics onto the manifold parameterized by \( \{ \varphi_{\alpha} \}_{\alpha \in \mathbb{R}^m} \), we require more smoothness assumptions than in the deterministic case \([51]\). The reason for this is that realizations of the stochastic process are not differentiable in time (this is why the stochastic analog of the Chain Rule - Ito's Lemma - requires second derivatives).

**Assumption 4.** Assume that the family \( \{ \varphi_{\alpha} \}_{\alpha \in \mathbb{R}^m} \) satisfies the following conditions.

(i) For \( 1 \leq i \leq m \), the derivatives \( \varphi_{\alpha,i} := [\partial/\partial \alpha_i] \varphi_{\alpha} \) and \( \varphi_{\alpha,ij} := [\partial^2/\partial \alpha_i \partial \alpha_j] \varphi_{\alpha} \) exist (the derivatives being taken in the norm of the space \( H \)) and are all in the space \( H \).

(ii) \( \varphi_{\alpha} - \varphi_{\beta} \in H \) for any \( \alpha, \beta \in \mathbb{R}^m \). (Note that \( \varphi_{\alpha} \) is not necessarily in \( H \))

(iii) \( \alpha \mapsto \varphi_{\alpha,i} \) and \( \alpha \mapsto \varphi_{\alpha,ij} \) are globally Lipschitz for all \( 1 \leq i, j \leq m \).

Let \( L_{\alpha} = A + Df(\varphi_{\alpha}) \) be the linearization of the drift in (2.1), about \( \varphi_{\alpha} \), and let \( L_{\alpha}^* \) be its adjoint. It follows from Assumption 2 that \( L_{\alpha} \) generates a continuous semigroup \( U_{\alpha}(t) \). That is, for any \( z \) in the domain of \( A \), \( U_{\alpha}(t) \cdot z := v_t \), where \( v_t \) satisfies the linear equation
\[
\frac{dv_t}{dt} = L_{\alpha}v_t, \tag{2.9}
\]
and \( v_0 := z \). Now it follows from taking any directional derivative of (2.2) with respect to \( \alpha \) that the linearized dynamics must always have a neutral eigenmode in directions tangential to the manifold \( \{ \varphi_{\alpha} \}_{\alpha \in \mathbb{R}^m} \).

The next assumption essentially means that the linearized dynamics is stable in all other directions. Without this stability, in most circumstances the noise would quickly force the system away from the manifold \( \{ \varphi_{\alpha} \}_{\alpha \in \mathbb{R}^m} \). Because \( \varphi_{\alpha} = T_\alpha \cdot \varphi_0 \), the following spectral gap property only needs to be verified for \( \alpha = 0 \).

**Assumption 5.** The spectrum \( \sigma(L_{\alpha}) \) of \( L_{\alpha} := A + Df(\varphi_{\alpha}) \) is such that
\[
\sigma(L_{\alpha}) \subset \{ \lambda \in \mathbb{C} : \Re(\lambda) \leq -b \} \cup \{0\},
\]
for some positive constant \( b \), independent of \( \alpha \). The eigenvalue \( 0 \) is assumed to have multiplicity \( m \), and the corresponding eigenvectors are spanned by \( \{ \varphi_{\alpha,j} \}_{1 \leq i \leq m} \).

The above assumption implies that the essential spectrum of \( L_{\alpha} \) lies in the subset \( \{ \lambda \in \mathbb{C} : \Re(\lambda) \leq -b \} \). This means that \( L_{\alpha} \) is Fredholm of zero index, and therefore the kernel of the adjoint operator \( L_{\alpha}^* \) is \( m \)-dimensional. One can show \([51]\) that a basis \( \{ \psi_{\alpha,i} \}_{1 \leq i \leq m} \) for the kernel of \( L_{\alpha}^* \) can be chosen such that
\[
\langle \psi_{\alpha,i}, \varphi_{\alpha,j} \rangle = 0 \text{ if } i \neq j \tag{2.10}
\]
\[
\langle \psi_{\alpha,i}, \varphi_{\alpha,i} \rangle = 1. \tag{2.11}
\]

The assumed invariance of \( A \) and \( f \) under the isometry \( T_\alpha \) implies that we can take
\[
\psi_{\alpha,i} = T_\alpha \psi_0^i. \tag{2.12}
\]
Let \( P_\alpha : H \to H \) be the following spectral projection operator that projects the kernel of \( L^* \) onto the kernel of \( L_\alpha \), i.e.

\[
P_\alpha \cdot u = \sum_{i=1}^{m} \langle \psi^i, u \rangle \varphi_{\alpha,i} \quad \text{and define} \quad \Pi_\alpha = \mathcal{I} - P_\alpha,
\]

where \( \mathcal{I} : H \to H \) is the identity operator. Define the Resolvent Operator corresponding to \( \Pi_\alpha \) to be, for \( \lambda \in \mathbb{C} \),

\[
\mathcal{R}_\alpha(\lambda) = (\lambda \mathcal{I} - \Pi_\alpha)^{-1}.
\]

**Assumption 6.** Suppose that there exists \( M > 0 \) such that for all \( \lambda \in \mathbb{C} \) with \( \text{Re}(\lambda) \geq -b \) (here \( b \) is the constant in Assumption 3), and all \( \alpha \in \mathbb{R}^m \),

\[
\| \mathcal{R}_\alpha(\lambda) \| \leq M.
\]

Note that once the above property is satisfied for \( \alpha = 0 \), it is satisfied for all \( \alpha \in \mathbb{R}^m \), thanks to the fact that the latter is obtained from the former through the application of a smooth isometry. Define

\[
V_\alpha(t) = U_\alpha(t) - P_\alpha,
\]

and note that \( V_\alpha(t) \cdot v = 0 \) for any \( v \in H \) such that

\[
\langle v, \psi^i \rangle = 0 \quad \text{for all} \quad 1 \leq i \leq m.
\]

It follows from the Gearhart-Pruss theorem (see [51, Theorem 4.1.5]) that there exists a constant \( c \geq 1 \) such that

\[
\| V_\alpha(t) \| \leq c \exp(-bt).
\]

This constant \( c \) is independent of \( \alpha \) thanks to the fact that any \( \varphi_\alpha \) can be obtained from \( \varphi_0 \) through applying an isometry.

**Assumption 7.**

(i) For \( 1 \leq i, j, k \leq m \), the derivatives \( \partial / \partial \alpha_j \psi^i_\alpha \), \( \partial^2 / \partial \alpha_j \partial \alpha_k \psi^i_\alpha \) exist (the derivatives being taken in the norm of the space \( H \)) and are all in the space \( H \). They are written as (respectively) \( \{ \psi_{i,j}^\alpha, \psi_{i,j,k}^\alpha \} \).

(ii) \( \alpha \mapsto \psi^i_\alpha, \alpha \mapsto \psi_{i,j}^\alpha \) and \( \alpha \mapsto \psi_{i,j,k}^\alpha \) are all globally Lipschitz.

(iii) Integration by parts holds i.e. \( \langle \psi^i_\alpha, \varphi_\alpha \rangle + \langle \psi_{i,j}^\alpha, \varphi_{\alpha,j} \rangle = 0 \) and \( \langle \psi_{i,j,k}^\alpha, \varphi_{\alpha,k} \rangle + \langle \psi_{i,j}^\alpha, \varphi_\alpha \rangle = 0 \).

(iv) \( \alpha \mapsto A^* \psi^i_\alpha, \alpha \mapsto A^* \psi_{i,j}^\alpha \) and \( \alpha \mapsto A^* \psi_{i,j,k}^\alpha \) are globally Lipschitz for each \( 1 \leq i, j, k \leq m \).

**Assumption 8.** The multiplicative noise operator is assumed to have the properties, for a constant \( C_B > 0 \),

\[
B : [0, \infty) \times E \to L(H, H)
\]

\[
\| B(t, x) - B(t, y) \|_{L} \leq C_B \| x - y \| \quad \text{for all} \quad t \geq 0 \quad \text{and} \quad x, y \in E
\]

\[
\| B(t, x) \|_{L} \leq C_B
\]

The following assumption is needed to obtain the exponential moment necessary for the exit-time bound in Section 5.

**Assumption 9.** For any \( t > s \) and any \( x \in E \), \( U_\alpha(t - s)B(s, x) \) is a Hilbert-Schmidt operator, with Hilbert-Schmidt norm (as defined in (1.4)) upperbounded by

\[
\sup_{\alpha \in \mathbb{R}^m} \int_0^{T_0} (t-s)^{-\frac{3}{2}} \sup_{\mathcal{E}} \| U_\alpha(t - s)B(s, x) \|_{HS}^2 \, ds < C_{HS},
\]

for some \( \beta \in (0, 1/2) \) and constant \( C_{HS} \), where

\[
T_0 = \frac{\log 4}{b}.
\]
3 Examples

This framework applies to a huge range of stochastically-forced traveling waves, stationary patterns, spiral waves, and complex spatially-distributed patterns in neural fields. We outline some examples: further examples can be found in [51, Section 4.6].

3.1 Stationary Bump in Neural Fields on $\mathbb{S}^1$

We consider a one-dimensional neural field on $\mathbb{S}^1$. This model was originally developed to model orientation selectivity in the primate visual cortex [71]. To the best of this author’s knowledge, Kilpatrick and Ermentrout were the first to prove that the bump is linearly stable [56]. They also studied its wandering under the effect of space-time noise over long timescales. This analysis was continued in [55, 18, 70]. In fact the variational phase SDE and long-time bound on the escape probability have already been determined in [70]. In some circumstances this system can be written as a gradient flow, and exact analytical formulae are available [10].

The evolution is described by the neural field equation on the ring $\mathbb{S}^1$:

$$
\tau \frac{\partial u(\theta, t)}{\partial t} = -u(\theta, t) + \int_{\pi}^{0} J(\theta - \theta') f(u(\theta', t)) d\theta'
$$

(3.1)

where $u(\theta, t)$ denotes the activity at time $t$ of a local population of cells with direction preference $\theta \in [-\pi, \pi)$, $J(\theta - \theta')$ is the strength of synaptic weights between cells with direction preference $\theta'$ and $\theta$. (Most applications of the ring model take $\theta \in [0, \pi]$ and interpret $\theta$ as the orientation preference of a population of neurons in primary visual cortex, see for example [71, 17].) The weight distribution is a $2\pi$-periodic and even function of $\theta$ and thus has the cosine series expansion

$$
J(\theta) = \sum_{n=0}^{N} J_n \cos(n\theta).
$$

(3.2)

For analytical simplicity, we assume that there are a finite number of terms in the series expansion. Finally, the firing rate function is taken to be a sigmoid $F(u) = (1 + e^{-\gamma(u - \kappa)})^{-1}$ with gain $\gamma$ and threshold $\kappa$. To fit the dynamics into the formalism of the previous section, we take $A \cdot u = -u$, and $f(u)(\theta) = \int_{\pi}^{0} J(\theta - \theta') f(u(\theta', t)) d\theta'$. Indeed the operators $A$ and $f$ are both bounded and Lipschitz over the Hilbert space $L^2(\mathbb{S}^1)$.

The dynamics in (3.1) is invariant under the translation by $\theta$ operator: $(T_{\theta} \cdot u)(\theta') := u(\theta' - \theta)$, with $\theta' - \theta$ taken modulo $\mathbb{S}^1$. Ermentrout and Kilpatrick [57] proved the existence and stability of a family of stationary bump solutions $\{U_{\theta}\}_{\theta \in \mathbb{S}^1}$, and we therefore define $\varphi_{\alpha} = U_{\alpha \mod \mathbb{S}^1}$, and take $m = 1$. Since the operator $A$ is bounded, the constant $c$ is $1$ [70].

A key difference between the above neural field equation and reaction-diffusion systems is that the neural field equation does not have a Laplacian, which works to smooth spatial irregularities. This means that in order that the stochastic equation is well defined, we require that the stochastic integral $\int_{0}^{t} B(u_s) dW_s$ belongs to the Hilbert space $H$. For this to be the case, we require that for any orthonormal basis $\{e_j\}_{j \geq 1}$ of $H$,

$$
\sup_{u \in E} \sum_{j=1}^{\infty} \langle e_j, B(u) e_j \rangle < \infty.
$$

In the terminology of [27], $B(u)$ is a trace class operator. Some neural field equations include a Laplacian, see for instance [65, 26, 73].
3.2 Traveling fronts in neural field equations

Neural field equations taking values in $\mathbb{R}$ are known as the Wilson-Cowan equations \[89\]. They take the form
\[
\partial_t u_t(x) = -u_t(x) + \int_{\mathbb{R}} w(x - y) F(u_t(y))dy, \quad t \geq 0, \ x \in \mathbb{R},
\]
(3.3)
where $w \in C(\mathbb{R}) \cap L^1(\mathbb{R})$ is the connectivity function, and $F : \mathbb{R} \to \mathbb{R}$ is a smooth and bounded sigmoid function (known as the nonlinear gain function). It is known (see \[33\] for example) that under some conditions on the functions $w$ and $F$ (in particular that there exist precisely three solutions to the equation $x = F(x)$ at $0, a$ and $1$ with $0 < a < 1$), then there exists a unique (up to translation) function $\hat{u} \in C^\infty(\mathbb{R})$ and speed $c \in \mathbb{R}$ such that $u_t(x) = \hat{u}(x - ct)$ is a solution to \[89\], where $\hat{u}$ is such that
\[
\lim_{x \to -\infty} \hat{u}(x) = 0, \quad \lim_{x \to \infty} \hat{u}(x) = 1,
\]
so that $\hat{u}$ is indeed a wave front. Note that in this case $\hat{u}$ itself is not in $L^2(\mathbb{R})$, but it can be shown that all derivatives of $\hat{u}$ are bounded and in $L^2(\mathbb{R})$.

Substituting $\hat{u}(x - ct)$ into \[89\], we see that $\hat{u}$ is such that $0 = A\hat{u} + f(\hat{u})$, where $A u := cu'$ and $f(u) = -u + w * F(u)$, and $*$ denotes convolution as usual. Moreover, due to translation invariance, we have that $\hat{u}_\alpha := \hat{u}(-c + \alpha)$ is also such that
\[
0 = A\hat{u}_\alpha + f(\hat{u}_\alpha), \quad \alpha \in \mathbb{R}.
\]
(3.4)
We must thus interpret \[89\] in the moving co-ordinate frame, i.e. writing $\zeta = x - ct$,
\[
\partial_t \hat{u}_\alpha(\zeta) = c \partial_\zeta u_t(\zeta) - u_t(\zeta) + \int_{\mathbb{R}} w(\zeta - y) F(\hat{u}_\alpha(y))dy, \quad t \geq 0, \ x \in \mathbb{R},
\]
(3.5)
The traveling front solutions are fixed points of the above equation, and the family of isometries is translation. We are thus in a specific situation of the general setup described in the previous section, with $H = L^2(\mathbb{R})$ and $\varphi_\alpha := \hat{u}_\alpha$. The spectral gap property has been proved in \[63\].

3.3 Traveling Waves in Reaction-Diffusion Systems

Consider the one-dimensional reaction-diffusion system
\[
du_t = \{\tilde{A}u_t + W'(u_t)\} dt + B(u_t) dW_t,
\]
(3.6)
with $\tilde{A}$ the Laplacian $\frac{\partial^2}{\partial x^2}$, and $W$ a potential function. In many circumstances such systems support traveling fronts. See the discussion in \[51\] Section 4.2.2. One must take care to work in the co-moving frame to apply the formalism of this paper (as discussed in Section \[5.2\] see also Remark\[1\]): then the traveling wave solutions constitute a manifold of fixed points (invariant under spatial translation). We thus write $A := \tilde{A} + c \frac{\partial}{\partial x}$, and in the moving frame the dynamics is of the form
\[
du_t = \{Au_t + W'(u_t)\} dt + \tilde{B}(t, u_t) dW_t,
\]
(3.7)
and $\tilde{B}(t, u_t) w := B(\bar{w}), \bar{w}(x) = w(x + ct)$ for $w \in E$. See \[23\] \[47\] for a discussion of the stochastic Nagumo equation.

3.4 Traveling pulses in neural fields

One can modify the classical neural field equation \[89\] to produce traveling pulse solutions in the following way. Indeed consider the system
\[
\begin{aligned}
\partial_t u_t &= -u_t + \int_{\mathbb{R}} w(-y) F(u_t(y))dy - v_t, \quad t \geq 0, \\
\partial_t v_t &= \theta u_t - \beta v_t,
\end{aligned}
\]
(3.8)
where as above $F : \mathbb{R} \rightarrow \mathbb{R}$ is a smooth and bounded sigmoid function, $w \in \mathcal{C}(\mathbb{R}) \cap L^1(\mathbb{R})$ and $\theta > 0, \beta \geq 0$ are some constants with $\theta << \beta$. This is called the neural field equation with adaptation (see for example [14, Section 3.3] for a review). This time we look for a solution to (3.8) of the form $(\hat{u}, \hat{v})$ where as above $U$ is a smooth and bounded sigmoid function, $\alpha := (\hat{u}(\cdot - ct), \hat{v}(\cdot - ct))$ for some $c \in \mathbb{R}$, such that $\hat{u}(x)$ and $\hat{v}(x)$ decay to zero as $x \rightarrow \pm \infty$. Substituting this into (3.8), we are thus looking for a solution to the equation

$$cU'(x) = \left(\begin{array}{cc}
-1 & -1 \\
\theta & -\beta
\end{array}\right) U(x) + f(U)(x), \quad x \in \mathbb{R},$$

(3.9)

where $U(x) = (\hat{u}(x), \hat{v}(x))$, and $f(U)(x) := (w * F(\hat{u})(x), 0)^T$, for all $x \in \mathbb{R}$.

It can be shown (see [77, Section 3.1] or [35]) that there exists (again under some conditions on the parameters) a smooth function $U := (\hat{u}, \hat{v}) \in [L^2(\mathbb{R})]^2$ and speed $c \in \mathbb{R}$ such that $U$ is a solution to (3.9). Moreover $\hat{u}$ and $\hat{v}$ are both smooth functions whose derivatives are all bounded and in $L^2(\mathbb{R})$. Thus, again by translation invariance we have that $U_\alpha := U(\cdot + \alpha) \in [L^2(\mathbb{R})]^2$ is a solution to

$$AU_\alpha + f(U_\alpha) = 0$$

for all $\alpha \in \mathbb{R}$, where

$$AU := AU' - \left(\begin{array}{cc}
-1 & -1 \\
\theta & -\beta
\end{array}\right) U, \quad \forall U \in [L^2(\mathbb{R})]^2.$$ 

Once again we are thus in a specific situation of the general setup described in Section 2, this time with $H = [L^2(\mathbb{R})]^2$ and $\phi_\alpha := U_\alpha$. Since $\hat{u}(x) \rightarrow 0$ as $x \rightarrow \pm \infty$, we say that the solution is a traveling pulse. The stability of the traveling pulse has been proved in [32]. Other types of neural field models also support traveling pulses, such as [54]. See [31] for the development of a phase decomposition broadly similar to the methods outlined in this paper.

### 3.5 Neural Field Patterns on Higher Dimensional Domains

In the one-dimensional neural field model of the previous section, neurons are grouped according to their orientation selectivity, which takes on values between $-\pi/2$ and $\pi/2$. More sophisticated neural field models can involve patterns with more degrees of freedom [14, 26, 22]. One example is the celebrated explanation of hallucinations using a neural field model that is invariant under three types of group action: rotation, reflection in the plane, and a shift-twist action [17]. If one were to impose space-time noise on this model (doing this has excellent biophysical motivation, because brain signals are typically very noisy), then one could easily observe a rich range of metastable phenomena over long time periods (applying the results of Section 6). Another recent example is the analysis of the wandering of bumps of neural activity over the sphere in [87, 15]: this has two degrees of freedom. The ergodic results of Section 6 can be applied to the wandering of a bump of activity over the sphere. One must use spherical polar co-ordinates $(\theta, \phi)$, and make sure to identify the points $(-\theta, \phi)$ and $(\theta, \phi)$, and one must identify $(\theta + 2k\pi, \phi + 2l\pi)$ with $(\theta, \phi)$.

### 3.6 The Scalar Viscous Conservation Law

This is the system

$$du_t = \{Au_t + \partial_x f(u_t)\}dt + B(u)dW_t,$$

(3.10)

with $A$ the Laplacian. See the discussion in [51, Section 4.4].

### 3.7 The Parametrically-Forced Nonlinear Schrodinger Equation

See [51, Section 4.5].
3.8 Spiral Waves in Reaction Diffusion Systems

Spiral waves are pervasive in non-equilibrium reaction-diffusion systems \[3, 4, 82, 86\]. Spatially-extended oscillations are also present in neural field equations \[36\]. \[7, 4\] have identified an equation for the leading order diffusive flux of the phase of oscillations in reaction-diffusion systems, a result that is consistent with this paper. Consider for examples the two-species model in \[3\]

\[\begin{align*}
\frac{\partial u}{\partial t} &= \nabla^2 u + \epsilon^{-1} u (1 - u) \{ u - (v + b)/a \} \\
\frac{\partial v}{\partial t} &= \delta \nabla^2 v + u - v.
\end{align*}\]

The domain is a circle of radius \(R\). \[3\] demonstrate that there exist stable spiral wave solutions to the above system. In a co-rotating reference frame, these solutions are fixed points. In this co-rotating frame, there is a manifold of solutions \(\{ \varphi_\theta \}_{\theta \in S^1}\). One solution can be obtained from another by applying a rotation isometry. If one imposes space-time white noise on the above system, then the resulting system will fit the requirements of Section 2.

Furthermore, the formalism of Section 6 could be used to determine the long-time average phase shift induced by noise correlations. These results parallel existing results for the long-time average phase shift of finite-dimensional stochastic oscillators \[40\].

4 Definition of the Stochastic Phase

In this section we outline two different phase definitions: the variational phase, and the isochronal phase. As explained in the introduction, one of the main goals of this paper is to determine elegant and useful stochastic differential equations for the phase; indeed the variational phase and isochronal phase each have particular merits. The variational phase is obtained by insisting that the amplitude is orthogonal to the neutral eigenmodes of the adjoint operator. The isochronal phase is the limiting point on the manifold \(\{ \varphi_\alpha \}_{\alpha \in \mathbb{R}^m}\) that the system would converge to in the absence of noise. The chief advantages of the variational phase are (i) its stochastic dynamics admits a more tractable analytic expression, (ii) it can be easily employed to obtain powerful exponential bounds on the probability of the system leaving a close neighborhood of the manifold (as performed in Section 5), and (iii) less regularity assumptions on \(f, A\) and \(B\) are required for the variational phase SDE than the isochronal phase SDE (note the additional assumptions at the start of Section 4.2). The chief advantages of the isochronal phase are (i) for finite-dimensional oscillators, the isochronal phase is the phase definition most preferred by experts, and so it is natural to search for its analog in our infinite-dimensional case, and (ii) it can be used to accurately predict the average occupation times of the system as it wanders close to the manifold over very long periods of time (as performed in Section 6).

4.1 Variational Phase SDE

The variational phase \(\beta_t\) is defined to be such that (i) it is continuous for all \(t < \tau\) (\(\tau\) is a stopping time defined in (4.4)) and (ii) for all \(t < \tau\), it exactly solves the identities, for \(1 \leq i \leq m\),

\[\begin{align*}
G_i(u_t, \beta_t) &= 0 \text{ where } G_i : E \times \mathbb{R}^m \to \mathbb{R} \text{ is such that } \\
G_i(z, \alpha) &:= \langle z - \varphi_\alpha, \psi_\alpha^i \rangle. \quad (4.2)
\end{align*}\]

This phase definition agrees with our definition in \[70\] for ‘stochastic neural bumps’ in terms of a weighted Hilbert space: in this paper we determined an equation of the form (4.1) by defining the phase to minimize a potential weighted by the ratio of the eigenvectors (see also \[58\]). The definition is different from our definition in \[50\] in two respects: (i) it is multi-dimensional, and (ii), in \[50\], instead of the eigenvectors \(\{ \psi_\alpha^i \}\) of the adjoint operator \(L_\alpha^*\), we have \(\varphi_\alpha, i\) (the eigenvector...
of $\mathcal{L}_\alpha$). To leading order in $\|u_t - \varphi_{\beta_t}\|^2$, the phase definitions in \[50\] and \[4.11\] are equivalent, and either could be used to obtain accurate long-time stability estimates. Upto linear order in $\epsilon$, this definition agrees with the definitions in \[28\] \[47\] \[46\].

Standard theory \[51\] Lemma 4.3.3] dictates that \[4.11\] has a unique solution $\beta_t$ as long as $u_t$ is close enough to the manifold $\{\varphi_{\alpha}\}_{\alpha \in \mathbb{R}^m}$. In Lemma \[4.11\] we will prove that $\beta_t$ is uniquely well-defined for all times up to $\tau$, and we will outline a precise stochastic differential equation for $\beta_t$. However before we do this, we start with some informal calculations to motivate the definition of the stochastic phase.

Notice first that our initial condition is such that \[4.11\] is satisfied exactly (for $t = 0$). Now define $\mathcal{M}(z, \alpha)$ to be the $m \times m$ square matrix with elements

$$
\mathcal{M}_{ij}(z, \alpha) = -\frac{\partial}{\partial \alpha_j} G_i(z, \alpha).
$$

(4.3)

It follows from the implicit function theorem that \[4.11\] is solvable for the phase in some neighborhood of $(u_t, \beta_t)$ as long as the matrix $\mathcal{M}(u_t, \beta_t)$ is invertible. We therefore define the stopping time

$$
\tau = \inf \{ t \geq 0 : \det (\mathcal{M}(u_t, \beta_t)) = 0 \},
$$

(4.4)

and we assume that $t < \tau$, so that a local solution for $\beta_t$ in terms of $u_t$ is possible.

Since we are assuming that $\langle \varphi_{\beta_t}, \psi^i_{\beta_t} \rangle$ is invariant under $\beta$, we find that

$$
\mathcal{M}_{ij}(z, \alpha) = -\langle z, \psi^i_{\alpha, j} \rangle
= -\langle z - \varphi_{\alpha, \psi^i_{\alpha, j}} \rangle - \langle \varphi_{\alpha, \psi^i_{\alpha, j}} \rangle.
$$

(4.5)

(4.6)

Our integration by parts assumption implies that $\langle \varphi_{\alpha, \psi^i_{\alpha, j}} \rangle = -\langle \varphi_{\alpha, \psi^i_{\alpha, j}} \rangle = -\delta(i, j)$, by assumption. We can thus write

$$
\mathcal{M}_{ij}(z, \alpha) = \delta(i, j) - \langle z - \varphi_{\alpha, \psi^i_{\alpha, j}} \rangle.
$$

(4.7)

In this above form, it is clear that as long as $\|u_t - \varphi_{\beta_t}\|$ is sufficiently small, $\mathcal{M}(u_t, \beta_t)$ is always invertible.

One can guess the dynamics of $\beta_t$ by first assuming that $\beta_t$ satisfies an SDE of the form

$$
d\beta_t = \mathcal{V}(u_t, \beta_t) dt + \epsilon \mathcal{Y}(t, u_t, \beta_t) dW_t,
$$

(4.8)

for functions $\mathcal{V} : \mathbb{R}^+ \times E \times \mathbb{R}^m \to \mathbb{R}^m$ and $\mathcal{Y} : \mathbb{R}^+ \times E \times \mathbb{R}^m \to \mathcal{L}(H, \mathbb{R}^m)$ to be determined below.

As explained in \[50\] \[70\], one can then formally expand out the identity $d\mathcal{G}_t(u_t, \beta_t) = 0$ and (i) insist that the stochastic terms are zero to determine $\mathcal{V}$, and then (ii) insist that the drift terms (i.e. the terms of finite variation) are zero, and thus determine $\mathcal{Y}$. To this end, using Ito’s Lemma,

$$
d\mathcal{G}_{i,t} = \langle du_t, \psi^i_{\beta_t} \rangle + \sum_{j=1}^{m} \frac{\partial \mathcal{G}_t}{\partial \beta^j_t} d\beta^j_t + \frac{1}{2} \sum_{j,k=1}^{m} \frac{\partial^2 \mathcal{G}_t}{\partial \beta^j_t \partial \beta^k_t} d\beta^j_t d\beta^k_t + \sum_{j=1}^{m} \langle du_t, \psi^i_{\alpha, j} \rangle d\beta^j_t,
$$

(4.9)

where the respective covariations of the processes are written as $d\beta^j_t d\beta^k_t$ and $du_t d\beta^j_t$. Now if $u_t$ were in the domain of $A$ and $Au_t \in H$, then using the fact that $A \varphi_{\beta_t} + f(\varphi_{\beta_t}) = 0$, it would hold that

$$
\langle du_t, \psi^i_{\beta_t} \rangle = \langle \{Au_t + f(u_t)\} dt + \epsilon B(t, u_t) dW_t, \psi^i_{\beta_t} \rangle
$$

(4.10)

$$
= \langle \{Au_t - A \varphi_{\beta_t} + f(u_t) - f(\varphi_{\beta_t})\} dt + \epsilon B(t, u_t) dW_t, \psi^i_{\beta_t} \rangle
$$

(4.11)

$$
= \langle u_t - \varphi_{\beta_t}, A^* \psi^i_{\beta_t} \rangle dt + \langle f(u_t) - f(\varphi_{\beta_t}), \psi^i_{\beta_t} \rangle dt + \epsilon \langle B(t, u_t) dW_t, \psi^i_{\beta_t} \rangle.
$$

(4.12)

In deriving the last expression, we assumed that $u_t$ is in the domain of $A$. If $u_t$ is not in the domain of $A$, then \[4.12\] is still well-defined (our assumptions dictate that $f(u_t) - f(\varphi_{\beta_t}) \in H$.
and $A^* \psi_{\beta_i}^j \in H$), and we will see in the next section that (4.12) is in fact the correct expression to use.

Matching the stochastic terms (the coefficients of $dW_t$) in (4.9), we find that

$$\epsilon \langle B(t, u_t) dW_t, \psi_{\beta_i}^i \rangle - \epsilon \sum_{j=1}^{m} \mathcal{M}_{ij}(u_t, \beta_i) \mathcal{V}_j(t, u_t, \beta_j) dW_t = 0. \quad (4.13)$$

Inverting this equation, we find that the linear operator $\mathcal{V}_j(t, u_t, \beta_i)$ must be such that for each $z \in H$,

$$\mathcal{V}_j(t, u_t, \beta_i) \cdot z = \sum_{j=1}^{m} \mathcal{N}_{ij}(u_t, \beta_i) \langle B(t, u_t) z, \psi_{\beta_i}^j \rangle \quad (4.14)$$

$$\mathcal{N}(u_t, \beta_i) = \mathcal{M}(u_t, \beta_i)^{-1} \quad (4.15)$$

noting that $\mathcal{M}(u_t, \beta_i)^{-1}$ is the matrix inverse of $\mathcal{M}(u_t, \beta_i)$. It is immediate from the definition of the stopping time that $\mathcal{M}(u_t, \beta_i)$ is invertible for $t < \tau$. We thus find that the covariance terms must have the form (using standard theory for stochastic integrals with respect to infinite-dimensional Wiener Processes [27 Chapter 4.3]),

$$d\beta_t^i d\beta_t^j = \epsilon^2 \sum_{p,q=1}^{m} \mathcal{N}_{jp}(u_t, \beta_i) \mathcal{N}_{kq}(u_t, \beta_i) \langle B^*(t, u_t) \psi_{\beta_i}^p, B^*(t, u_t) \psi_{\beta_i}^q \rangle dt \quad (4.16)$$

$$\langle d\beta_t, \psi_{\beta_i,j} \rangle d\beta_t^l = \epsilon^2 \sum_{p=1}^{m} \mathcal{N}_{jp}(u_t, \beta_i) \langle B^*(t, u_t) \psi_{\beta_i}^p, B^*(t, u_t) \psi_{\beta_i,j}^l \rangle dt. \quad (4.17)$$

The above terms do not directly depend on $\mathcal{V}(u_t, \beta_i)$, which means that we can easily solve (4.9) for $\mathcal{V}(u_t, \beta_i)$ by matching all of the coefficients of $dt$ terms. Observe that

$$\frac{\partial^2 \mathcal{G}_i}{\partial \alpha \partial \alpha^k} = \langle z, \psi_{\alpha,j,k} \rangle = \langle z - \varphi_\alpha, \psi_{\alpha,j,k} \rangle + \langle \varphi_\alpha, \psi_{\alpha,j,k} \rangle = \langle z - \varphi_\alpha, \psi_{\alpha,j,k} \rangle - \langle \varphi_\alpha, \psi_{\alpha,j,k} \rangle = (4.18)$$

using the integration by parts formula. We find that

$$- \sum_{j=1}^{m} \mathcal{M}_{ij}(u_t, \beta_i) \mathcal{V}_j(t, u_t, \beta_j) + \epsilon^2 \sum_{p=1}^{m} \mathcal{N}_{jp}(u_t, \beta_i) \langle B^*(t, u_t) \psi_{\beta_i}^p, B^*(t, u_t) \psi_{\beta_i,j}^l \rangle$$

$$+ \frac{\epsilon^2}{2} \sum_{j,k,p,q=1}^{m} \langle u_t, \psi_{\beta_i,j,k} \rangle \mathcal{N}_{jp}(u_t, \beta_i) \mathcal{N}_{kq}(u_t, \beta_i) \langle B^*(t, u_t) \psi_{\beta_i}^p, B^*(t, u_t) \psi_{\beta_i}^q \rangle$$

$$\langle u_t - \varphi_{\beta_i}, A^* \psi_{\beta_i}^j \rangle + \langle f(u_t) - f(\varphi_{\beta_i}), \psi_{\beta_i}^j \rangle = 0. \quad (4.19)$$

Inverting the matrix $\mathcal{M}(u_t, \beta_i)$, we thus find that for $1 \leq r \leq m$,

$$\mathcal{V}_r(t, u_t, \beta_i) = \sum_{i=1}^{m} \mathcal{N}_r(u_t, \beta_i) \Big\{ \epsilon^2 \sum_{p=1}^{m} \mathcal{N}_{jp}(u_t, \beta_i) \langle B^*(t, u_t) \psi_{\beta_i}^p, B^*(t, u_t) \psi_{\beta_i,j}^l \rangle$$

$$+ \frac{\epsilon^2}{2} \sum_{j,k,p,q=1}^{m} \langle u_t, \psi_{\beta_i,j,k} \rangle \mathcal{N}_{jp}(u_t, \beta_i) \mathcal{N}_{kq}(u_t, \beta_i) \langle B^*(t, u_t) \psi_{\beta_i}^p, B^*(t, u_t) \psi_{\beta_i}^q \rangle$$

$$\langle u_t - \varphi_{\beta_i}, A^* \psi_{\beta_i}^j \rangle + \langle f(u_t) - f(\varphi_{\beta_i}), \psi_{\beta_i}^j \rangle \Big\}. \quad (4.20)$$

Now

$$\langle u_t - \varphi_{\beta_i}, A^* \psi_{\beta_i}^j \rangle + \langle Df(\varphi_{\beta_i}) \cdot (u_t - \varphi_{\beta_i}), \psi_{\beta_i}^j \rangle = \langle u_t - \varphi_{\beta_i}, L_{\beta_i} \psi_{\beta_i}^j \rangle = 0,$$
since by definition $\psi_{\beta}^i$ is an eigenvector of $\mathcal{L}_{\beta_i}$. We thus find that

\[
V_r(t, u_t, \beta_t) = \sum_{i=1}^{m} N_{ri}(u_t, \beta_t) \left\{ \epsilon^2 \sum_{j,p=1}^{m} N_{jp}(u_t, \beta_t)(B^*(t, u_t)\psi_{\beta_i}^j, B^*(t, u_t)\psi_{\beta_j}^i) + \frac{\epsilon^2}{2} \sum_{j,k,p,q=1}^{m} \langle u_t, \psi_{\beta_i,jk}^j \rangle N_{jp}(u_t, \beta_t)N_{kp}(u_t, \beta_t)(B^*(t, u_t)\psi_{\beta_j}^j, B^*(t, u_t)\psi_{\beta_i}^j) \right. \\
\left. + (f(u_t) - f(\varphi_{\beta_i}) - Df(\varphi_{\beta_i}) \cdot (u_t - \varphi_{\beta_i}), \psi_{\beta_i}^i) \right\}. \tag{4.21}
\]

**Rigorous Definition of the Variational Phase SDE**

In the previous section, we guessed the form that the phase SDE should take by matching coefficients in the expression (4.1). We now rigorously prove that this informal derivation (i) defines a unique stochastic process $\beta_t$, and (ii) (4.1) is satisfied. We recall the definitions of the functions $\mathcal{V} : \mathbb{R}^+ \times E \times \mathbb{R}^m \rightarrow \mathbb{R}^m$, $\mathcal{V} = (\mathcal{V}_i)_{1 \leq i \leq m}$ and $\mathcal{Y} : \mathbb{R}^+ \times E \times \mathbb{R}^m \rightarrow \mathcal{L}(H, \mathbb{R}^m)$ in (4.14) and (4.21).

Now define $\beta_t$ to satisfy the $\mathbb{R}^m$-valued SDE

\[
d\beta_t = \mathcal{V}(t, u_t, \beta_t) dt + \epsilon \mathcal{Y}(t, u_t, \beta_t) dW_t, \tag{4.22}
\]

with initial condition $\beta_0 = \bar{\beta}$, for all times $t$ up to the stopping time $\tau$. We are going to see that this definition is consistent with our previous definition of $\beta_t$ in (4.1). Notice that the SDE for $\beta_t$ depends on the noise $W_t$ and solution $u_t$ of the original system. It is therefore essential to our argument that there exists a *strong* solution to the SDE (see [52, Chapter 5] for a definition of a strong solution). In other words we need more than just an identification of the probability law of $\beta_t$; we also require that it is coupled in the same space as $u_t$ and $W_t$.

**Lemma 4.1.** There exists a unique strong solution $\beta_t$ to the SDE in (4.22) for all times $t < \tau$. Furthermore this solution is such that, for all $t < \tau$,

\[
\mathcal{G}_t(u_t, \beta_t) = 0 \text{ where} \tag{4.23}
\]

\[
\mathcal{G}_t(z, \alpha) = (z - \varphi_{\alpha}, \psi_{\alpha}^i). \tag{4.24}
\]

**Proof.** The existence and uniqueness of the strong solution $\beta_t$ follows straightforwardly from the fact that the coefficient functions $\mathcal{V}$ and $\mathcal{Y}$ are locally Lipschitz in $\beta_t$ (see the proof in the one-dimensional case in [50]). Also, the operator $\mathcal{Y}$ is evidently Hilbert-Schmidt.

To prove (4.23), one might wish to try to find an infinite-dimensional Ito’s Lemma [27] (this is the change-of-variable formula for stochastic differential equations, analogous to the chain rule of differential calculus). However the possible unboundedness of the operator $A$ complicates any easy generalization of Ito’s Lemma to infinite dimensions. We thus instead take care to rigorously prove this; adapting the standard proof of Ito’s Lemma (see [27] and [52, Theorem 4.17]) to our setting. The trick to handling the unbounded operator $A$ is to instead work with its adjoint acting on $\psi_{\beta_i}^i$: the smoothness of the manifold $(\psi_{\alpha}^i)_{\alpha \in \mathbb{R}^m}$ ensures that this is well-behaved.

Write $u_t = u_t - \varphi_{\beta_t}$. It follows from Ito’s Lemma that, substituting the identity $A \varphi_{\beta_t} + f(\varphi_{\beta_t}) =
\[\begin{align*}
\frac{dv_t}{dt} &= (Au_t + f(u_t))dt + \epsilon B(t, u_t) dW_t - \sum_{i=1}^{m} \varphi_{\beta_i} \beta_i^i - \frac{1}{2} \sum_{j,k=1}^{m} \varphi_{\beta_{i,jk}} \beta_{i,jk}^k \\
&= (\mathcal{L}_{\beta} v_t + f(u_t) - Df(\varphi_{\beta_i}) \cdot v_t) dt + \epsilon B(t, u_t) dW_t - \sum_{i=1}^{m} \varphi_{\beta_i} \beta_i^i \\
&\quad - \frac{1}{2} \sum_{j,k=1}^{m} \varphi_{\beta_{i,jk}} \beta_{i,jk}^k
\end{align*}\]

where

\[\begin{align*}
B(s, z, \alpha) &= \mathbb{R}^+ \times H \times \mathbb{R}^m \rightarrow \mathcal{L}(H, H) \tag{4.28}
\end{align*}\]

and we have substituted the expression for \(d\beta_t^i d\beta_t^k\) in (4.16). The solution for \(v_t\), written in mild form, satisfies for \(t \in \left[t_k, t_{k+1}\right]\),

\[v_t = \bar{U}(t_k, t)v_{t_k} + \int_{t_k}^{t} \bar{U}(s, t) K_s ds + \epsilon \int_{t_k}^{t} \bar{U}(s, t) B(s, u_s, \beta_s) dW_s, \tag{4.31}\]

where \(\bar{U}(s, t)\) is the inhomogeneous semigroup generated by \(\mathcal{L}_{\beta_i}\). That is, for any \(z \in \mathcal{D}(A), \bar{U}(s, t) \cdot z := x_t\), where

\[\frac{dx_t}{dt} = \mathcal{L}_{\beta_i} x_t,\]

and \(x_t = z\). This definition can be continuously extended to all \(z \in H\). Define

\[\xi_n := \inf \left\{ t \in [0, \tau] : \det(\mathcal{M}(u_t, \beta_i)) = n^{-1} \quad \text{or} \quad \left\| \int_{0}^{t} B(s, u_s) dW_s \right\| \geq n \right\} \tag{4.32}\]

It may be seen that \((\xi_n)_{n \geq 1}\) is nondecreasing, and that \(\lim_{n \to \infty} \xi_n = \tau\) a.s. Define for any \(t \geq 0\)

\[\beta^n_t = \beta_{t \wedge \xi_n}\]

and \(v_t^n = v_{t \wedge \xi_n}\) where as above \((v_t)_{t \geq 0}\) is \(u_t - \varphi_{\beta_i}\), and \(u_t\) is the solution to the SDE in Proposition 2.4. Let \(I = \{t_i\}_{i=1}^{M}\) be a partition of \([0, t]\) for some \(t \geq 0\). For some family \(\{\theta_k\}_{k=0}^{M-1} \subset [0, 1]\) to be specified below, set \(w_k = \theta_k v^n_{t_k} + (1 - \theta_k) v^n_{t_{k+1}}\) and \(\zeta_k = \theta_k \beta^n_{t_k} + (1 - \theta_k) \beta^n_{t_{k+1}}\). Let \(X_k = (v^n_{t_{k+1}} - v^n_{t_k}, \beta^n_{t_{k+1}} - \beta^n_{t_k})\).

We now write \(\mathcal{G}_i(v_t, \beta_i) := \mathcal{G}_i(v_t + \varphi_{\beta_i}, \beta_i)\); in this way \(\mathcal{G}_i\) is a function on a Hilbert space, rather than the Banach space \(E\), and this simplifies the calculations. From the expressions in (4.7) and (4.18), it is clear that

\[\begin{align*}
\frac{\partial \mathcal{G}_i}{\partial \alpha_j}(z, \alpha) &= \delta(i, j) - \mathcal{M}_{ij}(z, \alpha) \\
\frac{\partial^2 \mathcal{G}_i}{\partial \alpha_j \partial \alpha_k}(z, \alpha) &= \langle z, \psi^{\beta}_{\alpha,jk} \rangle
\end{align*}\]
Thus by Taylor’s theorem,

\[
\tilde{g}_l(v^n_t, \beta^n_0) - \tilde{g}_l(v^n_0, \beta_0) = \sum_{k=1}^{M-1} \left\{ \sum_{j=1}^{m} (\delta(i, j) - M_{ij}(u^n_{p_k}, \beta_{t_k}))(\beta^{n,j}_{k+1} - \beta^{n,j}_t) + \langle v^n_{t_{k+1}} - v^n_t, \psi^i_{\beta^v_{t_{k+1}}} \rangle \right\} \\
+ \sum_{p=1}^{m} \left\{ \frac{1}{2} \sum_{q=1}^{m} \frac{\partial^2 \tilde{g}_l}{\partial \beta^p_t \partial \beta^q_t} \bigg|_{(x, w_k)} (\beta^{n,p}_{k+1} - \beta^{n,p}_t)(\beta^{n,q}_{k+1} - \beta^{n,q}_t) + \langle v^n_{t_{k+1}} - v^n_t, \psi^i_{\beta^v_{t_{k+1}}} \rangle (\beta^{n,p}_{k+1} - \beta^{n,p}_t) \right\},
\]

(4.33)

for some \( \theta_k \in [0, 1] \). Now for any \( r \geq 0 \), it must be that that

\[
\lim_{s \to 0^+} \sup_{r \leq z \leq t \leq t + r} (t - z)^{-1}(\tilde{U}(z, t)v_x, \psi^i_{\beta^v_z}) = \lim_{s \to 0^+} \sup_{r \leq z \leq t \leq t + r} (t - z)^{-1}(v_z, \tilde{U}(z, t)^*\psi^i_{\beta^v_z})
\]

by the dominated convergence theorem, and using our assumption that \( \psi^i_{\alpha} \) is in the domain of \( \mathcal{L}_\alpha \) for any \( \alpha \in \mathbb{R}^m \). We thus find that, using (4.32), and recalling that by definition, \( \mathcal{L}_{\beta^v_z} \psi^i_{\beta^v_z} = 0 \),

\[
\sum_{k=1}^{M-1} \langle v^n_{t_{k+1}} - v^n_t, \psi^i_{\beta^v_{t_{k+1}}} \rangle \\
\to \int_0^{\xi_{n} \wedge t} \langle \mathcal{L}_{\beta^v_z} \psi^i_{\beta^v_z}, v_x \rangle + \langle K, \psi^i_{\beta^v_z} \rangle ds + \epsilon \int_0^{\xi_{n} \wedge t} \langle \psi^i_{\beta^v_z}, \tilde{B}(s) dW_s \rangle
\]

(4.34)

\[
= \int_0^{\xi_{n} \wedge t} \langle \mathcal{L}_{\beta^v_z} \psi^i_{\beta^v_z}, v_x \rangle ds + \epsilon \int_0^{\xi_{n} \wedge t} \langle \psi^i_{\beta^v_z}, B(s, u_s) dW_s \rangle - \epsilon (\beta^{n,i}_{t,x} - \beta^{n,i}_0)
\]

since \( \langle \psi^i_{\beta^v_z}, \phi_{\beta^v_{t,x}} \rangle = \delta(i, j) \). Clearly,

\[
\sum_{k=1}^{M-1} \sum_{j=1}^{m} \delta(i, j)(\beta^{n,j}_{k+1} - \beta^{n,j}_t) \to (\beta^{n,i}_{t,x} - \beta^{n,i}_0).
\]

Similarly, making use of (4.32), as \( \Pi \to 0 \),

\[
\sum_{k=1}^{M-1} \sum_{j=1}^{m} M_{ij}(u^n_{p_k}, \beta_{t_k})(\beta^{n,j}_{k+1} - \beta^{n,j}_t) \to \epsilon \int_0^{t \wedge \xi_n} \langle \psi^i_{\beta^v_z}, B(s, u_s) dW_s \rangle + \int_0^{t \wedge \xi_n} Z^i_s ds,
\]

where

\[
Z^i_s = \epsilon^2 \sum_{j, p=1}^{m} N_{jp}(u_t, \beta_t)(B^*(t, u_t)\psi^p_{\beta^v_t}, B^*(t, u_t)\psi^i_{\beta^v_{t,x}}) + \frac{\epsilon^2}{2} \sum_{j, k, p, q=1}^{m} \langle u_t, \psi^i_{\beta^v_{t,x}}, N_{jp}(u_t, \beta_t)N_{kp}(u_t, \beta_t)(B^*(t, u_t)\psi^p_{\beta^v_t}, B^*(t, u_t)\psi^q_{\beta^v_t})
\]

(4.35)

+ (f(u_t) - f(\varphi_{\beta^v_t}) - Df(\varphi_{\beta^v_t}) \cdot (u_t - \varphi_{\beta^v_t}), \psi^i_{\beta^v_t}) \cdot (u_t - \varphi_{\beta^v_t}, \psi^i_{\beta^v_t}).
\]

It remains for us to deal with the second order terms in the Taylor expansion (4.33). As \( \Pi \to 0 \), we find that

\[
\frac{1}{2} \sum_{k=1}^{M-1} \sum_{p, q=1}^{m} \frac{\partial^2 \tilde{g}_l}{\partial \beta^p_t \partial \beta^q_t} \bigg|_{(x, w_k)} (\beta^{n,p}_{k+1} - \beta^{n,p}_t)(\beta^{n,q}_{k+1} - \beta^{n,q}_t)
\]

(4.34)

\[
\to \frac{\epsilon^2}{2} \sum_{j, k, p, q=1}^{m} \int_0^{t \wedge \xi_n} \langle v_x, \psi^i_{\beta^v_{t,x}}, N_{jp}(u_s, \beta_s)N_{kp}(u_s, \beta_s)(B^*(s, u_s)\psi^p_{\beta^v_s}, B^*(s, u_s)\psi^q_{\beta^v_s}) ds.
\]

(4.35)
using the expression for the covariation in \((4.16)\). Note that in the above,

\[
\langle v_s, \psi^i_{\alpha,j} \rangle = \langle u_s, \psi^i_{\beta,j} \rangle + \langle \varphi_{\alpha,j}, \psi^i_{\alpha,k} \rangle,
\]

since by the integration by parts formula, \(\langle \varphi_{\alpha,j}, \psi^i_{\alpha,k} \rangle = -\langle \varphi_{\alpha}, \psi^i_{\alpha,k} \rangle\).

Finally

\[
\lim_{M \to \infty} \sum_{k=1}^{M-1} \sum_{p=1}^{m} (v^n_{tk+1} - v^n_{tk}, \psi^i_{\zeta,k,p}) (\beta^n_{tk+1} - \beta^n_{tk})
\]

\[
= \lim_{M \to \infty} \sum_{k=1}^{M-1} \sum_{p=1}^{m} \left\{ (u^n_{tk+1} - u^n_{tk}, \psi^i_{\zeta,k,p}) (\beta^n_{tk+1} - \beta^n_{tk})
\right. \\
\left. - \sum_{j=1}^{m} \langle \varphi_{\beta,j}, (\beta^n_{tk+1} - \beta^n_{tk}), \psi^i_{\zeta,k,p} \rangle (\beta^n_{tk+1} - \beta^n_{tk}) \right\}
\]

\[
\to e^2 \int_0^{\xi_n} \left\{ \sum_{j,p=1}^{m} J_{jp}(u_s, \beta_s) \langle B^*(s, u_s) \psi^p_{\beta_s}, B^*(s, u_s) \psi^i_{\beta_s, j} \rangle 
\right. \\
\left. - \sum_{j,p,q,r=1}^{m} \langle \varphi_{\beta,j}, \psi^i_{\beta_s, r} \rangle J_{jp}(u_s, \beta_s) J_{qr}(u_s, \beta_s) \langle B^*(s, u_s) \psi^p_{\beta_s}, B^*(s, u_s) \psi^q_{\beta_s, r} \rangle \right\} ds.
\]

Adding up the above identities, we find that

\[
\widetilde{g}_t(v_t, \beta_t) = 0,
\]

for all \(t \leq \xi_n\). Since this holds for any \(n\) and \(\xi_n \uparrow \tau\) we have the result. \(\square\)

### 4.2 Isochronal Phase

In Section 4.1 we defined a slightly different phase \(\gamma_t\) - the isochronal phase - which will be necessary to obtain the ergodicity result of section 6. The reason why a different phase definition is needed is that the drift of \(d\beta_t\) is \(O(\|v_t\|^2 + \epsilon^2)\), and since for all \(t \leq \tau\), \(\|v_t\|\) is typically \(O(\epsilon)\), the \(O(\|v_t\|^2)\) terms make a non-trivial contribution to the drift dynamics. The benefit of using \(\gamma_t\) is that the leading order of the drift of \(d\gamma_t\) does not depend on the amplitude \(v_t\). \(\gamma_t\) is an analog of the isochronal phase used in the phase reduction of finite-dimensional oscillators [34][19].

We require the following additional assumptions throughout this section.

**Assumption 10.** For all \(T \geq 0\),

\[
\lim_{h \to \theta} \sup_{t \in [0,T]} h^{-1/2} \left\| P^A(h) u_t - u_t \right\| = 0,
\]

\(\mathbb{P}\)-almost-surely.

In future work, in the case that \(A\) is elliptic, the following assumption could likely be considerably relaxed.

**Assumption 11.** We assume that for some choice of orthonormal basis \(\{e_j\}_{j \geq 1}\) for \(H\),

\[
\lim_{M \to \infty} \sup_{u \in E} \sum_{j=M}^{\infty} \|B(u)e_j\|^2 = 0.
\]

(4.37)
Write $\Phi : E \to \mathbb{R}^m$ to be the phase map of Section 4.1 satisfying the implicit relationship
\[ G_i(u, \Phi(u)) = 0 \text{ for all } 1 \leq i \leq m. \] (4.38)

Let $\mathcal{R} \subset E$ be a subset of the attracting basin of the manifold $\{\varphi_\alpha\}_{\alpha \in \mathbb{R}^m}$, with
\[ \mathcal{R} = \{ u \in E : \|u - \varphi_{\Phi(u)}\| \leq \delta \}, \] (4.39)
for some $\delta$ to be determined more precisely below. We assume that $\mathcal{R}$ is sufficiently close to the manifold $\{\varphi_\alpha\}_{\alpha \in \mathbb{R}^m}$ that $\Phi$ is uniquely well-defined on $\mathcal{R}$. $\delta$ will be chosen to be small enough that if $z \in \mathcal{R}$ then there exists a unique $\eta_z \in \mathbb{R}^m$ such that
\[ \lim_{t \to \infty} u_t = \varphi_{\eta_z} \text{ where } u_t = \exp(\int_0^t f(u_s) \cdot v_s ds) \text{ and } u_0 = z. \] (4.40)

Define the corresponding map to be
\[ \Theta : \mathcal{R} \to \mathbb{R}^m : \Theta(z) := \eta_z. \] (4.42)

We require an SDE expression for $\Theta(u_t)$. To do this, we must obtain a deeper understanding of the map $\Theta$: establishing that (i) it is always well-defined for $u \in \mathcal{R}$ (as long as $\delta$ is sufficiently small), and (ii) that it is twice continuously Frechet-differentiable, which will allow us to (iii) apply Ito’s Lemma to obtain an SDE expression for $d\gamma$. We also wish to show that $\Theta$ is very close to the variational phase of the previous section, which will help us obtain a more tractable expression for the occupation time estimates in Section 6.

**Lemma 4.2.** There exists $\delta > 0$ such that for all $u \in E$ such that $\|u - \varphi_{\Phi(u)}\| \leq \delta$, $\Theta(u)$ is uniquely well-defined.

**Proof.** The proof employs a contraction mapping and is an adaptation of [SS, Chapter 5]. Define $\mathcal{X} \subset C([0, \infty), H) \times \mathbb{R}^m$ to be the Banach Space of all $(v, \alpha)$ such that the following norm is finite,
\[ \|(v, \alpha)\|_2 = \|\alpha\| + \sup_{t \geq 0} \left\{ \exp(bt) \|v_t\| \right\}. \] (4.43)

For some $(v, \alpha) \in \mathcal{X}$, define $\Gamma_u(v, \alpha) \in C([0, \infty), H)$ to be
\[ \Gamma_u(v, \alpha) = v_0 + \int_0^t \left( \alpha_i \cdot \psi_{\alpha,i} \right) ds - \sum_{i=1}^m \psi_{\alpha,i} \int_t^\infty \left\{ f(v_s + \varphi_\alpha) - f(\varphi_\alpha) - Df(\varphi_\alpha) \cdot v_s \right\} ds, \] (4.44)
recalling the definition of $V_\alpha(t)$ in (2.18). For $u \in E$, define $\Lambda_u : \mathbb{R}^m \to \mathbb{R}^m$ to be, for $1 \leq i \leq m$,
\[ \Lambda_u(\alpha, v)^i = \Phi(u)^i + \langle \psi_{\Phi(u)}^i, u - \varphi_{\Phi(u)} \rangle - \langle \psi_{\alpha,i}^i, \varphi_{\alpha} - \varphi_{\Phi(u)} \rangle - \langle \alpha_i - \Phi(u)^i, \psi_{\alpha,i} \rangle + \langle \psi_{\alpha,i}^i, \int_0^\infty \left\{ f(\varphi_\alpha + v_s) - f(\varphi_\alpha) - Df(\varphi_\alpha) \cdot v_s \right\} ds \rangle. \] (4.45)

It suffices for us to show that, for $\delta$ small enough, there exists a unique $(\eta, v)$ such that $\eta = \Lambda_u(\eta, v)$ and $v = \Gamma_u(\eta, v)$. This is because, using the fact that $\langle \psi_{\Phi(u)}^i, u - \varphi_{\Phi(u)} \rangle = 0$, and $\langle \psi_{\alpha,i}^i, \phi_{\alpha,i} \rangle = 1$, the existence of a fixed point $(\eta, \hat{v}(u))$ implies that
\[ \varphi_{\eta,i} \langle \psi_{\eta,i}^i, u - \varphi_{\eta} \rangle + \int_0^\infty \left\{ f(\varphi_{\eta} + \hat{v}_s(u)) - f(\varphi_{\eta}) - Df(\varphi_{\eta}) \cdot \hat{v}_s(u) \right\} ds = 0. \] (4.46)
Adding this to (4.44), and recalling that \( P_\eta = \sum_{i=1}^m \phi_\eta \langle \psi_i^\eta, \cdot \rangle \), we obtain that

\[
\hat{v}_t(u) = \{ V_\eta(t) + P_\eta \} \cdot (u - \varphi_\eta) + \int_0^t V_\eta(t - s) \cdot \{ f(\hat{v}_s(u) + \varphi_\eta) - f(\varphi_\eta) - Df(\varphi_\eta) \cdot \hat{v}_s(u) \} ds + P_\eta \cdot \int_0^t \{ f(\hat{v}_s(u) + \varphi_\eta) - f(\varphi_\eta) - Df(\varphi_\eta) \cdot \hat{v}_s(u) \} ds. \tag{4.47}
\]

Since \( U_\eta(t) = P_\eta + V_\eta(t) \), we obtain that

\[
\hat{v}_t(u) = U_\eta(t) \cdot (u - \varphi_\eta) + \int_0^t U_\eta(t - s) \cdot \{ f(\hat{v}_s(u) + \varphi_\eta) - f(\varphi_\eta) - Df(\varphi_\eta) \cdot \hat{v}_s(u) \} ds \tag{4.48}
\]

Writing \( \hat{u}_t(u) = \varphi_\eta + \hat{v}_t(u) \), this means that \( \hat{u}_t \) satisfies the dynamics

\[
\hat{u}_t(u) = P^A(t) \cdot u + \int_0^t P^A(t - s) \cdot f(\hat{u}_s(u)) ds. \tag{4.49}
\]

Since \( \| \hat{v}_t(u) \| \to 0 \) as \( t \to \infty \), this means that \( \eta = \Theta(u) \).

It remains for us to show that there exists a unique \((\eta, \check{v})\) such that \( \eta = \Lambda_u(\eta, \check{v}) \) and \( \check{v} = \Gamma_u(\eta, \check{v}) \). The existence and uniqueness follows from the contraction mapping theorem. The following lemma contains the identities necessary for us to be able to apply the theorem.

**Lemma 4.3.** For constant \( \delta_2 > 0 \), let \( Z_u(\delta_2) \) be all \((\alpha, v) \in X\) such that

\[
\| v \|_b := \sup_{t \geq 0} \{ \exp(bt) \| v_t \| \} \leq \delta_2 \tag{4.50}
\]

\[
\| \alpha - \Phi(u) \| \leq \delta_2^{3/2}. \tag{4.51}
\]

For all sufficiently small \( \delta_2 \), and taking \( \delta \) to be sufficiently small (recall that \( R = \{ u \in E : \| u - \Phi(u) \| \leq \delta \} \}), for all \( u \in R \),

\[
(\Lambda_u(\alpha, v), \Gamma_u(\alpha, v)) \in Z_u(\delta_2) \text{ whenever } (\alpha, v) \in Z_u(\delta_2). \tag{4.52}
\]

One can also choose \( \delta \) and \( \delta_2 \) to be such that there exists \( \rho \in (0, 1) \) such that for all \((\alpha, v), (\beta, z) \in Z_u(\delta_2) \),

\[
\| \Lambda_u(\alpha, v) - \Lambda_u(\beta, z) \| + \sup_{t \geq 0} \{ \exp(bt) \| \Gamma_u(\alpha, v)_t - \Gamma_u(\beta, z)_t \| \} \leq \rho \| \alpha - \beta \| + \rho \sup_{t \geq 0} \{ \exp(bt) \| v_t - z_t \| \}. \tag{4.53}
\]

**Proof.** We start by proving (4.52). Let \((\alpha, v) \in Z_u(\delta_2) \). Define \( F : \mathbb{R}^m \times H \to H \) to be

\[
F(\alpha, v) = f(\varphi_\alpha + v) - f(\varphi_\alpha) - Df(\varphi_\alpha) \cdot v
\]

The Fréchet differentiability of \( f \) implies that there is a constant \( C_F \) such that \( \| F(\alpha, v) \| \leq C_F \| v \|^2 \). Using the bound on \( \| V_\alpha(t) \| \) in (2.18),

\[
\| \Gamma_u(v, \alpha) \| \leq c \| u - \varphi_\alpha \| \exp(-bt) + c_{F} \int_0^t \exp(2sb-2tb) \| v_s \|^2 ds + c_{F} \sum_{i=1}^m \| \varphi_{\alpha,i} \| \int_t^{\infty} \| v_s \|^2 ds. \tag{4.54}
\]

Using the triangle inequality,

\[
\| u - \varphi_\alpha \| \leq \| u - \varphi_{\Phi(u)} \| + \| \varphi_{\Phi(u)} - \varphi_\alpha \| \leq \delta + \| \alpha - \Phi(u) \| \| D\Phi(w) \|,
\]
where $w$ is in the convex hull of $\varphi_{\Psi(u)}$ and $\varphi_\alpha$, using Taylor’s Theorem. Since, by assumption, $\sup_{\alpha \in \mathbb{R}^m, 1 \leq i \leq m} |\varphi_{\alpha,i}| < \infty$, we can choose $\delta$ and $\delta_2$ to be such that
\[
\epsilon \left\{ \delta + 2\delta_2^{1/2} \sup_{\alpha \in \mathbb{R}^m, 1 \leq i \leq m} |\varphi_{\alpha,i}| \right\} \leq \delta_2/3,
\]
and this ensures that
\[
\epsilon \|u - \varphi_\alpha\| \exp(-bt) \leq \delta_2/3.
\]
Similarly, since $(\alpha, v) \in Z(\delta_2)$,
\[
cC_F \int_0^t \exp(2sb - 2tb) \|v_s\|^2 ds \leq cC_F \delta_2^2 \leq \exp(-bt)\delta_2/3,
\]
for all sufficiently small $\delta_2$. Finally,
\[
C_F \sum_{i=1}^m \|\varphi_{\alpha,i}\| \int_t^\infty \|v_s\|^2 ds \leq C_F \sum_{i=1}^m \|\varphi_{\alpha,i}\| \delta_2^2 \int_t^\infty \exp(-2bs) ds \leq \exp(-bt)\delta_2/3,
\]
for small enough $\delta_2$. Taken together, the above equations imply that
\[
\|\Gamma_u(v, \alpha)\| \leq \exp(-bt)\delta_2,
\]
(4.55)
once $\delta_2$ is small enough. Bounding $|A(\alpha, v) - \Phi(u)|$ analogously,
\[
\left| \langle \psi_\alpha^i, \varphi_\alpha - \varphi_{\Psi(u)} \rangle \right| = O(\|\Phi(u) - \alpha\| \|u - \varphi_{\Psi(u)}\|) = O(\delta_2^2)
\]
\[
\left| \langle f(\varphi + v_s) - f(\varphi) - Df(\varphi) \cdot v_s \rangle \right| = O(\|v_s\|^2).
\]
Thus for small enough $\delta_2$, it must be that $(A_u(\alpha, v), \Gamma_u(\alpha, v)) \in Z_u(\delta_t)$, as required.

It remains for us to prove (4.53). It can be shown that there exist constants $C, C_2 > 0$ such that for all $v, w \in H$, $\|v\|, \|w\| \leq C_2$,
\[
\|F(\alpha, v) - F(\beta, v)\| \leq C \|\alpha - \beta\| \|v\|
\]
\[
\|F(\alpha, v) - F(\alpha, w)\| \leq C(\|v\| + \|w\|) \|v - w\|.
\]
These identities allow one to straightforwardly prove (4.53). \hfill $\Box$

Lemma 4.12 now follows immediately from the application of a fixed point theorem to the results in Lemma 4.13. \hfill $\Box$

Now that we have defined the isochronal phase $\Theta(u)$, our next step is to establish that it is twice continuously Frechet differentiable. For any $u \in \mathcal{R}$, define the following auxiliary variables $\{\dot{u}_t(u)\}_{t \geq 0}$, such that
\[
\dot{u}_t(u) = P^A(t) \cdot u + \int_0^t P^A(t - s) f(\dot{u}_s(u)) ds.
\]
(4.56)
$\dot{u}_t(u)$ is the solution of the deterministic dynamics, started at $u$, in the absence of noise. We know that $\lim_{t \to \infty} \dot{u}_t(u) = \varphi_{\Theta(u)}$. Define $\hat{v}_t(u) = \dot{u}_t(u) - \varphi_{\Theta(u)}$. As we demonstrated in the proof of Lemma 4.3, $\hat{v}_t(u)$ satisfies the identity
\[
\dot{v}_t(u) = -\sum_{i=1}^m \varphi_{\Theta(u),i} \left\langle \psi_{\Theta(u)}^i, \int_t^\infty F(\Theta(u), v_s) ds \right\rangle +
\]
\[
V_{\Theta(u)}(t) \cdot (u - \varphi_{\Theta(u)}) + \int_0^t V_{\Theta(u)}(t - s) \cdot F(\Theta(u), \hat{v}_s(u)) ds.
\]
(4.57)
The following lemma contains useful identities on the regularity of the isochronal phase map.

\[\text{...} \]
Lemma 4.4. One can choose δ to be sufficiently small that there exists a constant Ĉ such that for all \( u \in \mathcal{R} \),

\[
\| \Phi(u) - \Theta(u) \| \leq \tilde{C} \| u - \varphi_{\Theta(u)} \|^2 \tag{4.58}
\]

\[
\| \hat{v}(u) - V_{\Theta(u)} \cdot (u - \varphi_{\Theta(u)}) \|_b \leq \tilde{C} \| u - \varphi_{\Theta(u)} \|^2. \tag{4.59}
\]

\[
\| \hat{v}(u) \|_b \leq \varepsilon \| \hat{v}_0(u) \|. \tag{4.60}
\]

\[
\| \hat{v}(u) - \hat{v}(z) \|_b \leq \tilde{C} \| u - z \| \text{ for all } u, z \in \mathcal{R}. \tag{4.61}
\]

Proof. We start by establishing (4.60). Using the triangle inequality, we find that

\[
\| \hat{v}_t(u) \| \leq C_F \sum_{i=1}^m \| \hat{\varphi}_{\Theta(u),i} \| \| \hat{\psi}_{\Theta(u),i} \| \int_t^\infty \| \hat{v}_s(u) \|^2 \, ds + \varepsilon \exp(-bt) \| \hat{v}_0(u) \| + C_F \varepsilon \int_0^t \exp(bs - bt) \| \hat{v}_s(u) \|^2 \, ds.
\]

Multiplying both sides by \( \exp(bt) \) and substituting the definition of \( \| \hat{v}(u) \|_b \), we obtain that

\[
\exp(bt) \| \hat{v}_t(u) \| \leq C_F \| \hat{v}(u) \|^2 \sum_{i=1}^m \| \hat{\varphi}_{\Theta(u),i} \| \| \hat{\psi}_{\Theta(u),i} \| \exp(bt) \int_t^\infty \exp(-2bs) \, ds + \varepsilon \| \hat{v}_0(u) \|
\]

\[
+ \varepsilon C_F \| \hat{v}(u) \|_b^2 \int_0^t \exp(-bs) \, ds.
\]

We thus find that there exists a constant \( \hat{C} \) such that

\[
0 \leq \hat{C} \| \hat{v}(u) \|_b^2 - \| \hat{v}(u) \|_b + \varepsilon \| \hat{v}_0(u) \| \tag{4.64}
\]

Finding the roots of the quadratic on the right hand side, and assuming that

\[
\delta_2 < \left( 1 + \sqrt{1 - 4\hat{C} \| \hat{v}_0(u) \|} \right)/(2\hat{C}) \text{ (recall that by definition } \| \hat{v}(u) \|_b \leq \delta_2) \text{, the above equation then implies that}
\]

\[
\| \hat{v}(u) \|_b \leq \left( 1 - \sqrt{1 - 4\hat{C} \| \hat{v}_0(u) \|} \right)/(2\hat{C}) \leq \varepsilon \| \hat{v}_0(u) \|,
\]

since the curve \( x \to \sqrt{x} \) is concave at 1, it must lie beneath its tangent line. We have thus established (4.60).

We saw in the previous lemma that \( \| F(\Theta(u), v) \| \leq C_F \| v \|^2 \). It then follows easily from (4.57) that

\[
\| \hat{v}_t(u) - V_{\Theta(u)}(t) \cdot (u - \varphi_{\Theta(u)}) \| \leq \text{Const} \times \exp(-bt) \| \hat{v}(u) \|_b^2
\]

\[
\leq \text{Const} \times \varepsilon \exp(-bt) \| u - \varphi_{\Theta(u)} \|^2,
\]

thanks to (4.60). This establishes (4.59).

We next establish (4.58). Recall that \((\Theta(u), \hat{v}(u))\) is a fixed point of (4.46), which means that \( \Theta(u) = \Lambda_u(\Theta(u), \hat{v}(u)) \). (4.46) implies that

\[
\left\| \Theta(u) - \Phi(u) - \left\{ \int_0^\infty f(\varphi_\alpha + \hat{v}_s) - f(\varphi_\alpha) - Df(\varphi_\alpha) \cdot \hat{v}_s \right\} \right\|
\]

\[
= O\left( \| \Theta(u) - \Phi(u) \| \| u - \varphi_{\Theta(u)} \| + \| \Theta(u) - \Phi(u) \|^2 \right) \tag{4.66}
\]

Furthermore, using (4.59),

\[
f(\varphi_\alpha + \hat{v}_s) - f(\varphi_\alpha) - Df(\varphi_\alpha) \cdot \hat{v}_s = D^{(2)} f(\varphi_\alpha) \cdot \hat{v}_s \cdot \hat{v}_s + O\left( \| \hat{v}_s(u) \|^3 \right)
\]

\[
= D^{(2)} f(\varphi_\alpha) \cdot (V_\alpha(s) \cdot \hat{v}_0(u)) \cdot (V_\alpha(s) \cdot \hat{v}_0(u)) + O\left( \| u - \varphi_{\Theta(u)} \|^3 \right)
\]

The proof of (4.61) is omitted. \( \square \)
Lemma 4.5. (i) $\Theta : \mathcal{R} \to \mathbb{R}^m$ is twice continuously Frechet differentiable for all $u$ in the interior of $\mathcal{R}$. The first and second Frechet derivatives at $u \in \mathcal{R}$ are written as $D\Theta(u) : H \to H$ and $D^{(2)}\Theta(u) : H \times H \to H$ (it must be emphasized that we are defining $D\Theta(u)$ and $D^{(2)}\Theta(u)$ to act on $H$, not $E$).

(ii) The function

$$u \to \sum_{j=1}^{\infty} \{D^{(2)}\Theta(u) \cdot B(u)e_j : B(u)e_j\},$$

where $\{e_j\}_{j \geq 1}$ is the orthonormal basis for $H$ of Assumption 17, is continuous over $\mathcal{R}$.

(iii) Furthermore there exists a constant $C$ such that for all $u \in \mathcal{R}$, and all $w \in H$, writing $\alpha = \Theta(u)$, for $1 \leq i \leq m$,

$$|D^{(i)}\Theta(u) \cdot w - \langle \psi_i^i, w \rangle| \leq C \|u - \varphi_\alpha\| \|w\|$$

and for all $w \in H$,

$$\|D^{(2)}\Theta(u) \cdot w \cdot w - 2 \sum_{j=1}^{m} \langle \psi_j^i, w \rangle \langle \psi_j^i, w \rangle + \sum_{j,k=1}^{m} \langle \varphi_{\alpha,j} \cdot \psi_j^i, w \rangle \langle \psi_j^i, w \rangle \langle \psi_j^i, w \rangle, \delta \hat{v}(u) \cdot z)_{\mathcal{X}}.$$

Proof. The Frechet Differentiability of $\Theta$ is established by 'implicitly differentiating' the fixed point identities (4.56) and $\Theta(u) = \Lambda_u(\Theta(u), \hat{v}(u))$ that define $(\Theta(u), \hat{v}(u))$. That is, fix $z \in H$ and define $w_n = u + n^{-1}z \in \mathcal{R}$, for large enough $n > 0$. We then find that there exists $(D\Theta(u) \cdot z, \delta \hat{v}(u) \cdot z) \in \mathcal{X}$ such that

$$\lim_{n \to \infty} \|n \delta \hat{v}(w_n) - n \hat{v}(u) - \delta \hat{v}(u) \cdot z\| = 0 \text{ where}$$

$$\delta \hat{v}(u) = \delta \hat{u}(u) - \sum_{j=1}^{m} \varphi_{\theta(u),j} \left(D\Theta(u) \cdot z\right)^j$$

For all $t \geq 0$, $\delta \hat{u}(u) \cdot z = P^A(t) \cdot z + \int_{0}^{t} P^A(t-s) \cdot Df(\hat{u}_s(u)) \cdot (\hat{v}_s(u) \cdot z) ds$

$$D\Theta(u) \cdot z = \Lambda_u(\Theta(u), \hat{v}(u)) \cdot (D\Theta(u) \cdot z, \delta \hat{v}(u) \cdot z),$$

and in this last expression, $DA_{\mathcal{U}} : \mathcal{X} \to \mathbb{R}^m$ denotes the Frechet Derivative. The second Frechet Derivative, in directions $w, z \in H$, is established analogously, as follows. We find that there exists $\delta^{(2)}\hat{v}(u) \cdot w \cdot z \in C([0, \infty), H)$ such that $\|\delta^{(2)}\hat{v}(u) \cdot w \cdot z\| < \infty$ and $D^{(2)}\Theta(u) \cdot w \cdot z \in \mathbb{R}^m$ such that $\|\delta^{(2)}\hat{v}(u) \cdot w \cdot z\| < \infty$ and for all $t \geq 0$,

$$\delta^{(2)}\hat{v}(u) \cdot w \cdot z = \int_{0}^{t} P^A(t-s) \cdot \left(Df(\hat{u}_s(u)) \cdot (\delta^{(2)}\hat{v}_s(u) \cdot w) + D^{(2)}f(\hat{u}_s(u)) \cdot (\delta \hat{v}_s(u) \cdot z) \right) ds$$

and

$$D^{(2)}\Theta(u) \cdot w \cdot z = (\Lambda_u(\Theta(u), \hat{v}(u)) \cdot (\delta \hat{v}(u) \cdot z) \cdot (\delta \hat{v}(u) \cdot z)) + \Lambda_u(\Theta(u), \hat{v}(u)) \cdot (D\Theta(u) \cdot w, \delta \hat{v}(u) \cdot z)$$

(ii) follows from Assumption 17. The proof of (4.67) and (4.68) proceeds from the approximation of $\Theta$ in (4.66). This result means that $D\Phi(u) \cdot w = D\Theta(u) \cdot w$, to leading order. Applying the implicit function theorem to
Finally, we apply the approximation in (4.66), and use the approximation using integration by parts, and we thus find that
\[ u \text{ are twice-Fr\'echet-differentiable, as noted in (4.5),} \]

\[ T > 0, \]

For any \( \epsilon > 0, \)

\[ \text{Proof. The proof is essentially a generalization of Ito's Lemma to our infinite dimensional problem: it is similar to the proof of Lemma 4.1. For some} \]

\[ \xi^n = \inf \{ t \leq \tau_i - \epsilon : \| u_t - \varphi(\gamma_t) \| = n \text{ or} \sup_{1 \leq i \leq m} |\gamma^i_t| = n \} \]

For any \( T > 0, \) we discretize \([0, T \wedge \xi_n] \) into a partition \( \{ t_i \}_{i=1}^M. \) Using a second-order Taylor expansion (which is possible because \( \Theta \) is twice-Fr\'echet-differentiable, as noted in (4.5)),

\[ \Theta(u_t) - \Theta(u_0) = \sum_{i=1}^M \{ D \Theta(u_t_i) \cdot (u_{t_i+1} - u_{t_i}) + D^{(2)} \Theta(\bar{u}_i) \cdot (u_{t_i+1} - u_{t_i}) \cdot (u_{t_i+1} - u_{t_i}) \} \]

where \( \bar{u}_i = \lambda_i u_t_i + (1 - \lambda_i) u_{t_i+1} \) for some \( \lambda_i \in [0, 1]. \) Now

\[ u_{t_i+1} - u_{t_i} = \{ P^A(t_{i+1} - t_i) - I \} \cdot u_{t_i} + \int_{t_i}^{t_{i+1}} P^A(t_{i+1} - s) f(u_s) ds + \epsilon \int_{t_i}^{t_{i+1}} P^A(t_{i+1} - s) B(u_s) dW_s \]
As the partition $\Delta \to 0$,

$$\sum_{i=1}^{M} D\Theta(u_{t_i}) \int_{t_i}^{t_{i+1}} P^A(t_{i+1} - s) B(u_s) dW_s \to \int_{0}^{T \wedge \xi_n} D\Theta(u_s) \cdot B(u_s) dW_s. \quad (4.82)$$

We now show that as $\Delta \to 0$,

$$\sum_{i=1}^{M} D\Theta(\hat{u}_{t_i}) \cdot \left\{ \{ P^A(t_{i+1} - t_i) - I \} \cdot u_{t_i} + \int_{t_i}^{t_{i+1}} P^A(t_{i+1} - s) f(u_s) ds \right\} \to 0. \quad (4.83)$$

For $t \in [t_i, t_{i+1}]$, define $\hat{u}_t$ to satisfy the deterministic flow with initial condition $u_{t_i}$, i.e. $\hat{u}_{t_i} = u_{t_i}$ and

$$\hat{u}_t = P^A(t - t_i) u_{t_i} + \int_{t_i}^{t} P^A(t - s) f(\hat{u}_s) ds. \quad (4.84)$$

Now for all $t \in [t_i, t_{i+1}]$, $\Theta(\hat{u}_t) = \Theta(u_{t_i})$, because the isochronal phase is (by definition) invariant under the deterministic flow. Since $\Theta$ is Frechet differentiable, by Taylor’s Theorem there must exist $\hat{\lambda}_i \in [0, 1]$ such that, writing $\hat{u}_t = \tilde{\lambda}_i \hat{u}_t + (1 - \hat{\lambda}_i) \tilde{u}_{t_i+1}$,

$$D\Theta(\hat{u}_t) \cdot (\tilde{u}_{t_i+1} - \hat{u}_t) = 0. \quad (4.85)$$

Applying a second Taylor expansion to the above identity, one obtains that

$$D\Theta(\hat{u}_t) \cdot (\hat{u}_{t_{i+1}} - \hat{u}_t) + D^{(2)}\Theta(\hat{u}_t) \cdot (\tilde{u}_{t_i+1} - \hat{u}_t) \cdot (\hat{u}_t - \hat{u}_t) = 0, \quad (4.86)$$

and $\tilde{u}_t$ is in the convex hull of $\hat{u}_t$ and $\hat{u}_t$. Now it follows straightforwardly from the uniformly Lipschitz property of $f$ that

$$\left\| \tilde{u}_{t_i+1} - \hat{u}_t - \{ P^A(t_{i+1} - t_i) - I \} \cdot u_{t_i} + \int_{t_i}^{t_{i+1}} P^A(t_{i+1} - s) f(u_s) ds \right\| \leq \text{Const} \times (t_{i+1} - t_i) \times \sup_{s \in [t_i, t_{i+1}]} |u_s - \hat{u}_s(u_{t_i})|. \quad (4.87)$$

One easily checks using Gronwall’s Inequality (and the Lipschitz nature of $f$) that

$$|u_s - \hat{u}_s(u_{t_i})| = O((t_{i+1} - t_i)).$$

We thus find that

$$D\Theta(\hat{u}_t) \cdot \left( P^A(t_{i+1} - t_i) - I \right) \cdot u_{t_i} + \int_{t_i}^{t_{i+1}} P^A(t_{i+1} - s) f(u_s) ds =$$

$$- D^{(2)}\Theta(\hat{u}_t) \cdot (\tilde{u}_{t_i+1} - \hat{u}_t) \cdot (\hat{u}_t - \hat{u}_t) + O((t_{i+1} - t_i)^2). \quad (4.88)$$

Now, making use of the expression in (4.84), since $\sup_{t \in [0, T]} \| P^A(t) \| < \infty$, and $\| f(u_s) \|$ is uniformly bounded,

$$\| \tilde{u}_t - \hat{u}_t \| \leq \| P^A(t - t_i) u_{t_i} - u_{t_i} \| + O(t - t_i)$$

$$= o((t_{i+1} - t_i)^{1/2}),$$

using Assumption [10]. Similarly $(\hat{u}_t - \hat{u}_t) = o((t_{i+1} - t_i)^{1/2})$. We thus find that

$$(t_{i+1} - t_i)^{-1} D^{(2)}\Theta(\hat{u}_t) \cdot (\tilde{u}_{t_i+1} - \hat{u}_t) \cdot (\hat{u}_t - \hat{u}_t) \to 0,$$

as the partition goes to zero. This means that (4.83) must hold, as required.
It remains to show that as $\Delta \to 0$, for any $1 \leq p \leq m$,
\[
\left| \sum_{i=1}^{M} D^{(2)} \Theta^p(\widetilde{u}_i) \cdot (u_{t_{i+1}} - u_{t_i}) \cdot (u_{t_{i+1}} - u_{t_i}) - \epsilon^2 \int_0^{T \wedge \xi_n} \sum_{j=1}^{\infty} \left\{ D^{(2)} \Theta^p(u_s) \cdot B(u_s)e_j \cdot B(u_s)e_j \right\} ds \right| \to 0.
\]
(4.89)

Now, substituting the decomposition of $u_t$ in (4.90), and writing
\[
Z_i = \{ P^{A}(t_{i+1} - t_i) - I \} \cdot u_{t_i} + \int_{t_i}^{t_{i+1}} P^{A}(t_{i+1} - s) f(u_s) ds,
\]
(4.90)
we obtain that
\[
\sum_{i=1}^{M} D^{(2)} \Theta^p(\widetilde{u}_i) \cdot (u_{t_{i+1}} - u_{t_i}) \cdot (u_{t_{i+1}} - u_{t_i}) = \sum_{i=1}^{M} \left\{ D^{(2)} \Theta^p(\widetilde{u}_i) \cdot Z_i \cdot Z_i + 2\epsilon D^{(2)} \Theta^p(\widetilde{u}_i) \cdot Z_i \cdot \int_{t_i}^{t_{i+1}} P^{A}(t_{i+1} - s) B(u_s) dW_s \right. \\
+ \epsilon^2 D^{(2)} \Theta^p(\widetilde{u}_i) \cdot \int_{t_i}^{t_{i+1}} P^{A}(t_{i+1} - s) B(u_s) dW_s \cdot \int_{t_i}^{t_{i+1}} P^{A}(t_{i+1} - s) B(u_s) dW_s \left. \right\}. \quad (4.91)
\]

One easily shows that, as $\Delta \to 0$, and making use of Assumption 11,
\[
\sum_{i=1}^{M} \left\{ D^{(2)} \Theta^p(\widetilde{u}_i) \cdot Z_i \cdot Z_i + 2\epsilon D^{(2)} \Theta^p(\widetilde{u}_i) \cdot Z_i \cdot \int_{t_i}^{t_{i+1}} P^{A}(t_{i+1} - s) B(u_s) dW_s \right\} \to 0.
\]

Let $\{ w_{t_j} \}_{j \geq 1}$ be independent Brownian motions. We can represent $W_t$ as $\sum_{j=1}^{\infty} w_{t_j} e_j$, where $\{ e_j \}_{j \geq 1}$ is the orthonormal basis for $H$ of Assumption 11, noting that the summation does not converge in $H$ (but only in an appropriate ambient Hilbert space [27]). We fix an integer $K > 0$ and find that
\[
\int_{t_i}^{t_{i+1} \wedge \xi_n} P^{A}(t_{i+1} - s) B(u_s) \cdot dW_s = \sum_{j=0}^{\infty} \int_{t_i}^{t_{i+1} \wedge \xi_n} P^{A}(t_{i+1} - s) B(u_s) e_j dW_s^j + X_i
\]
where $X_i = \sum_{j=1}^{K} \int_{t_i}^{t_{i+1} \wedge \xi_n} P^{A}(t_{i+1} - s) B(u_s) e_j dW_s^j$. It thus follows from Assumption 11 that for each $n \in \mathbb{Z}^+$ there must exist $K_n$ such that for all $K \geq K_n$,
\[
\mathbb{E} \left[ \left| D^{(2)} \Theta^p(\widetilde{u}_i) \cdot \int_{t_i}^{t_{i+1}} P^{A}(t_{i+1} - s) B(u_s) dW_s \cdot \int_{t_i}^{t_{i+1}} P^{A}(t_{i+1} - s) B(u_s) dW_s \right|^2 \right] \leq n^{-1} (t_{i+1} - t_i).
\]
(4.92)

However the continuity of $D^{(2)} \Theta^p$ implies that
\[
\left| \sum_{i=1}^{M} D^{(2)} \Theta^p(\widetilde{u}_i) \cdot Z_i \cdot Z_i - \epsilon^2 \int_0^{T \wedge \xi_n} \sum_{j=1}^{K_n} (D^{(2)} \Theta^p(u_t) \cdot B(u_t)e_j \cdot B(u_t)e_j) dt \right| \to 0,
\]
as $\Delta \to 0$.  □
5 Long-Time Stability

The main aim of this section is to show that the probability of the stochastic system leaving a close neighborhood of the manifold parameterized by \( \{ \varphi_a \}_{a \in \mathbb{R}^m} \) after an exponentially long period of time, is exponentially unlikely. This result is necessary for the metastable results of the next section, and is also of independent interest. The main result of this section is Theorem 5.1.

Define the ‘amplitude’ of the solution (relative to the nearest shifted pattern / wave) to be

\[
v_t = u_t - \varphi_{\beta_t},
\]

recalling the definition of the variational phase \( \beta_t \) in (4.1). Let \( \tilde{\kappa} \) be a positive constant such that \( \tilde{\kappa} \leq \frac{1}{2m} \sup_{1 \leq i,j \leq m} \| \psi_{0,j} \|^{-1} \). We require an additional condition on \( \tilde{\kappa} \) further on, in (5.20). Define the stopping time, for some \( \kappa \leq \tilde{\kappa} := \frac{1}{2m} \sup_{1 \leq i,j \leq m} \| \psi_{0,j} \|^{-1} \),

\[
\eta = \inf \{ t \geq 0 : \| v_t \| = \kappa \}. \tag{5.2}
\]

This constant has been chosen such that, if \( \| v_t \| \leq \tilde{\kappa} \), then necessarily

\[
\sum_{i,j=1}^{m} a_i a_j M_{ij}(u_t, \beta_t) - \frac{1}{2} \sum_{i=1}^{m} a_i^2 \geq 0, \tag{5.3}
\]

for all \( a \in \mathbb{R}^m \). The above identity can be inferred from the definition in (4.7). This implies that (i) the matrix \( M(u_t, \beta_t) \) can be inverted if \( t \leq \eta \), and (ii) the map \( (u_t, \beta_t) \rightarrow N(u_t, \beta_t) \) is locally Lipschitz.

The main result of this section is the following.

**Theorem 5.1.** Let \( p \in [0, 2) \). There exists a constant \( C > 0 \) (independent of the choice of \( p \)) and \( \epsilon(p) > 0 \) such that for all \( \epsilon \in (0, \epsilon(p)) \), and all \( \kappa \in [\epsilon(p), \tilde{\kappa}] \) and all \( T > 0 \),

\[
P\left( \sup_{t \in [0,T]} \| v_t \| > \kappa \right) \leq T \exp \left( -C \epsilon^{-2} \kappa^2 \right). \tag{5.4}
\]

If one desires an optimal value for the constant \( C \), one would need to perform a more detailed Large Deviations analysis (see for instance [30]), and this would likely require extensive computation. Even so, to the best of this author’s knowledge, the bound in Theorem 5.1 is the most optimal one in the literature for this sort of problem.

We discretize time into intervals of length \( \Delta t := b^{-1} \log(4\epsilon^{-1}) \), and we write \( t_a := a \Delta t \). Write \( u_a := u_{t_a} \) and \( \beta_a := \beta_{t_a} \). Define the event

\[
A_a = \{ \| v_a \| \leq \kappa/(2\epsilon) \} \cap \left\{ \| v_{a+1} \| > \kappa/(2\epsilon) \text{ or } \sup_{t \in [t_a, t_{a+1}]} \| v_t \| > \kappa \right\}. \tag{5.5}
\]

Write \( R = \lfloor T / \Delta t \rfloor \). Noting that \( \| v_0 \| = 0 \), a union-of-events bound implies that

\[
P\left( \sup_{t \in [0, R \Delta t]} \| v_t \| > \kappa \right) \leq \sum_{a=0}^{R} P(A_a). \tag{5.6}
\]

It follows from Ito’s Lemma that

\[
dv_t = (Au_t + f(u_t))dt + \epsilon B(t, u_t)dW_t - \sum_{i=1}^{m} \varphi_{\beta_t,i}d\beta^i_t - \frac{1}{2} \sum_{j,k=1}^{m} \varphi_{\beta_t,jk}d\beta^j_t d\beta^k_t \tag{5.7}
\]

\[
= (L_{\beta_t}v_t + f(u_t) - f(\varphi_{\beta_t}) - Df(\varphi_{\beta_t}) \cdot v_t)dt + \epsilon B(t, u_t)dW_t - \sum_{i=1}^{m} \varphi_{\beta_t,i}d\beta^i_t - \frac{1}{2} \sum_{j,k=1}^{m} \varphi_{\beta_t,jk}d\beta^j_t d\beta^k_t, \tag{5.8}
\]

\[\]

---

\(^{†}\) In general, research on the existence of Large Deviation principles is vastly more developed than research on the efficient numerical computation of Large Deviations rate functions / first exit times.
and we have substituted the identity $A\varphi_{\beta_n} + f(\varphi_{\beta_n}) = 0$. Write $U_a(t) := U_{\beta_n}(t)$ to be the semigroup generated by $L_{\beta_n}$. Taking the mild solution, and substituting the expression for $d\beta^j_t d\beta^k_t$ in (4.16), we have that

$$v_t = U_a(t - t_a) v_a + \int_{t_a}^t U_a(t - s) \mathcal{H}_s ds + \epsilon \int_{t_a}^t U_a(t - s) \tilde{B}(s, u_s, \beta_s) dW_s,$$

(5.9)

where

$$\tilde{B}(s, z, \alpha) = \mathbb{R}^+ \times H \times \mathbb{R}^m \rightarrow L(H, H)$$

(5.10)

and

$$\beta(s, z, \alpha) = B(s, z) - \sum_{j=1}^{m} \varphi_{\alpha j}(s, z, \alpha)$$

(5.11)

$$\mathcal{H}_s := f(u_s) - f(\varphi_{\beta_n}) - Df(\varphi_{\beta_n}) \cdot v_s - \sum_{i=1}^{m} \varphi_{\beta_n i} \mathcal{V}_i(s, u_s, \beta_s)$$

(5.12)

$$- \frac{c^2}{2} \sum_{j,k,p,q=1}^{m} \varphi_{\beta_n,jk} N_{jp}(u_s, \beta_s) N_{kq}(u_s, \beta_s) \langle B^*(s, u_s) \psi_{\beta_n p}, B^*(s, u_s) \psi_{\beta_n q} \rangle.$$

**Lemma 5.2.** There exist constants $C_1, C_2, C_3$ such that for all $t \in [t_a, t_{a+1} \wedge \eta]$, where $t_a = \sup\{t_b : t_b \leq t\}$,

$$\int_{t_a}^t U_a(t - s) \mathcal{H}_s ds \leq C_1 + C_3 \sup_{s \in [t_a, t]} \|v_s\|^2$$

$$+ \epsilon C_2 \sup_{s \in [t_a, t_{a+1}]} \sup_{1 \leq i \leq m} \left| \int_{t_a}^s \mathcal{V}_i(r, u_r, \beta_r) dW_r \right|$$

(5.13)

**Proof.** Using the definition of the semigroup in (2.17),

$$\|U_a(t - s)\|_L \leq \|V_{\beta_n}(t - s)\|_L + \|P_{\beta_n}\|_L$$

$$\leq \epsilon + \sum_{i=1}^{m} \|\varphi_{\beta_n i}\| \|\psi_{\beta_n i}\| = \epsilon + \sum_{i=1}^{m} \|\varphi_{\alpha i}\| \|\psi_{\alpha i}\|$$

using (2.18), and the fact that $P_{\beta_n}$ is an $m$-dimensional projection. Let $\tilde{C}$ be a constant such that

$$\sup_{t \geq 0, x \in E, \alpha \in \mathbb{R}^m : \|x - \varphi_{\alpha}\| \leq \bar{\kappa}, \bar{\kappa}(x, \alpha) = 0} \sup_{1 \leq i \leq m} \left| \mathcal{V}_i(t, x, \alpha) \right| \leq \tilde{C}$$

$$\sup_{t \geq 0, x \in E, \alpha \in \mathbb{R}^m : \|x - \varphi_{\alpha}\| \leq \bar{\kappa}, \bar{\kappa}(x, \alpha) = 0} \sup_{1 \leq i \leq m} \left| \mathcal{V}_i(t, x, \alpha) \right|_{HS} \leq \tilde{C}.$$

The constant $\tilde{C}$ exists, because as noted in (5.3), the choice of $\bar{\kappa}$ ensures that the lowest eigenvalue of $\mathcal{M}(x, \alpha)$ is greater than or equal to a half.

Employing the triangle inequality,

$$\|f(u_s) - f(\varphi_{\beta_n}) - Df(\varphi_{\beta_n}) \cdot v_s\| \leq \|f(u_s) - f(\varphi_{\beta_n}) - Df(\varphi_{\beta_n}) \cdot v_s\|$$

$$+ \|Df(\varphi_{\beta_n}) \cdot v_s - Df(\varphi_{\beta_n}) \cdot v_s + f(\varphi_{\beta_n}) - f(\varphi_{\beta_n})\|$$

Now using the second order Taylor expansion, there exists $\lambda_1 \in [0, 1]$ such that, writing $\bar{u}_a = \lambda_1 \varphi_{\beta_n} + (1 - \lambda_1) u_a$,

$$f(u_a) - f(\varphi_{\beta_n}) - Df(\varphi_{\beta_n}) \cdot v_a = D(2) f(\bar{u}_a) \cdot v_a \cdot v_a.$$
The assumed boundedness of the second derivative thus implies that for some constant \( C > 0 \),

\[
\| f(u_a) - f(\varphi_{\beta_a}) - D f(\varphi_{\beta_a}) v_a \| \leq C \| v_a \|^2. \tag{5.15}
\]

The boundedness of the first and second Frechet derivatives of \( f \) implies that there exists a constant \( C \) such that

\[
\| D f(\varphi_{\beta_a}) \cdot v_a - D f(\varphi_{\beta_a}) \cdot v_a + f(\varphi_{\beta_a}) - f(\varphi_{\beta_a}) \| \leq C \{ 1 + \sup_{s \in [t_a, \eta \land t_{a+1}]} \| v_s \| \} \times \sup_{s \in [t_a, \eta \land t_{a+1}]} \| \beta_s - \beta_a \| \tag{5.16}
\]

since by definition of \( \eta \), \( \sup_{s \in [t_a, \eta \land t_{a+1}]} \| v_s \| \leq \bar{\kappa} \). Using our SDE for \( \beta_i \) in (4.22), we find that as long as \( \sup_{s \in [t_a, t_{a+1}]} \| v_s \| \leq \bar{\kappa} \), there exists a constant \( \bar{C} \) such that

\[
\sup_{s \in [t_a, t_{a+1}], 1 \leq i \leq m} \| \beta_i^s - \beta_i^0 \| \leq \Delta t \bar{C} \epsilon^2 + \epsilon \sup_{s \in [t_a, \eta \land t_{a+1}]} \sup_{1 \leq i \leq m} \int_{t_a}^s \gamma_i(r, u_r, \beta_r) dW_r \]

\[
+ \Delta t C \| \psi_0^i \| \| u_t \|^2 \tag{5.18}
\]

and we have used the Cauchy-Schwarz Inequality to find that

\[
\| (f(u_t) - f(\varphi_{\beta_a}) - D f(\varphi_{\beta_a}) \cdot (u_t - \varphi_{\beta_a}), \psi_i^0) \| \leq \| f(u_t) - f(\varphi_{\beta_a}) - D f(\varphi_{\beta_a}) (u_t - \varphi_{\beta_a}) \| \| \psi_i^0 \| \leq C \| \psi_0^i \| \| v_t \|^2,
\]

thanks to (5.15), and since \( \| \psi_i^0 \| = \| \psi_0^i \| \).

Now we insist that \( \epsilon(p) \) (defined in the statement of the theorem) is such that

\[
0 < \epsilon_p^2 C_1 \leq \epsilon(p)/(16c), \tag{5.19}
\]

which is always possible since by assumption \( p < 2 \). We also insist that \( \bar{\kappa} \) is such that

\[
C^2 \bar{\kappa}^2 \leq \bar{\kappa}/(16c), \tag{5.20}
\]

and since \( \kappa \leq \bar{\kappa} \), is must be that \( C^2 \kappa^2 \leq \bar{\kappa}/16c \).

**Lemma 5.3.**

\[
A_a \subset \{ \| v_a \| \leq \kappa/(2c) \} \cap \{ B_a \cup C_a \} \quad \text{where}
\]

\[
B_a = \left\{ \epsilon C_2 \sup_{s \in [t_a, \eta \land t_{a+1}]} \int_{t_a}^{s \land \eta} \gamma_i(r, u_r, \beta_r) dW_r \geq \frac{\kappa}{16c} \right\} \tag{5.22}
\]

\[
C_a = \left\{ \epsilon \sup_{t \in t_{a+1} \land \eta} \int_{t_a}^{t \land \eta} U_a(t - s) \bar{B}(s, u_s, \beta_s) dW_s \geq \frac{\kappa}{16c} \right\}. \tag{5.23}
\]

**Proof.** Using the expression for \( v_i \) in (5.21) and the triangle inequality, for all \( t \in [t_a, \eta \land t_{a+1}] \),

\[
\| v_t \| \leq \| U_a(t - t_a)v_a \| + \left\| \int_{t_a}^t U_a(t - s) \mathcal{H}_s dW_s \right\| + \epsilon \left\| \int_{t_a}^t U_a(t - s) \bar{B}(s, u_s, \beta_s) dW_s \right\|
\]

\[
\leq \epsilon \exp(-b(t - t_a)) \| v_a \| + \epsilon^2 C_1 + \epsilon C_2 \sup_{s \in [t_a, \eta \land t_{a+1}]} \sup_{1 \leq i \leq m} \int_{t_a}^s \gamma_i(r, u_r, \beta_r) dW_r
\]

\[
+ C^2 \kappa^2 \epsilon + \epsilon \left\| \int_{t_a}^t U_a(t - s) \bar{B}(s, u_s, \beta_s) dW_s \right\|. \tag{5.24}
\]
after substituting the semigroup bound in (5.18) and the bound in (5.2). Since $t - t_a \leq b^{-1} \log(4c)$ and $\|v_a\| \leq \kappa/(2c)$, $c\exp(-b(t - t_a))\|v_a\| \leq \kappa/(8c)$. Also, $C_3\kappa^2 \leq \kappa/(16c)$. Furthermore, since by definition $\kappa > c_{(p)}$, the inequality in (5.19) implies that $\epsilon^2 C_1 \leq \kappa/(16c)$. We thus find that

$$\|v_{a+1}\| \leq \frac{\kappa}{4} + \epsilon C_2 \sup_{s \in [t_a, t_{a+1}]} \sup_{1 \leq i \leq m} \left( \int_{t_a}^{s} \| \gamma_i \| (r, u_r, \beta_r) \| dW_r \right) + \epsilon \left( \| U_a(t - s) \| B(s, u_s, \beta_s) \| dW_s \right).$$

Thus if $\|v_{a+1}\| > \frac{\kappa}{4}$, then it must be that either $B_a$ or $C_a$ must hold.

The remaining event in the definition of $A_a$ is $\|U_a(t - t_a)\| \leq \epsilon \|v_a\| \leq \epsilon \kappa/(2c)$.

$$\sup_{t \in [t_a, t_{a+1}]} \|v_t\| \leq \kappa/2 + \epsilon^2 C_1 + \epsilon C_2 \sup_{s \in [t_a, t_{a+1}]} \sup_{1 \leq i \leq m} \left( \int_{t_a}^{s} \| \gamma_i \| (r, u_r, \beta_r) \| dW_r \right) + \epsilon \left( \| U_a(t - s) \| B(s, u_s, \beta_s) \| dW_s \right) \leq \kappa/2 + \kappa/(8c) + \epsilon^2 C_1 + \epsilon C_2 \sup_{s \in [t_a, t_{a+1}]} \sup_{1 \leq i \leq m} \left( \int_{t_a}^{s} \| \gamma_i \| (r, u_r, \beta_r) \| dW_r \right) + \epsilon \sup_{t \in [t_a, t_{a+1}]} \left( \int_{t_a}^{t\wedge \eta} U_a(t - s) B(s, u_s, \beta_s) dW_s \right),$$

since $\epsilon^2 C_1 \leq \kappa/(16c)$, $C_3\kappa^2 \leq \kappa/(16c)$ and $\epsilon \geq 1$. We again see that if $\sup_{t \in [t_a, t_{a+1}]} \|v_t\| \geq \kappa$, then $B_a$ or $C_a$ must hold. \hfill \Box

It thus follows from (5.6) that

$$\mathbb{P}\left( \sup_{t \in [0, T]} \|v_t\| \geq \kappa \right) \leq \sum_{a=0}^{R} \{ \mathbb{P}(B_a) + \mathbb{P}(C_a) \}. \tag{5.25}$$

Recalling that $R = [T/\Delta t]$, it then follows from Lemma 5.4 that for a constant $C_4 > 0$,

$$\mathbb{P}\left( \sup_{t \in [0, T]} \|v_t\| \geq \kappa \right) \leq TC_4 \exp\left( - C\epsilon^{-2}\kappa^2 \right).$$

For small enough $\epsilon$, this implies Theorem 5.1.

**Lemma 5.4.** There exists a constant $C > 0$ such that

$$\sup_{a \geq 0} \mathbb{P}(B_a) \leq \exp\left( - C\epsilon^{-2}\kappa^2 \right) \tag{5.26}$$

$$\sup_{a \geq 0} \mathbb{P}(C_a) \leq \exp\left( - C\epsilon^{-2}\kappa^2 \right). \tag{5.27}$$

**Proof.** We prove the first result only. The bound of (5.27) can be obtained by taking an exponential moment of the stochastic integral: see (50) Section 5. Using the definition of $\tilde{B}$ in (5.11),

$$\mathbb{P}\left( \sup_{t \in [t_a, t_{a+1}]} \left. \int_{t_a}^{t\wedge \eta} U_a(t - s) \tilde{B}(s, u_s, \beta_s) dW_s \right\| \geq \frac{\kappa}{16c} \right) \leq \mathbb{P}\left( \sup_{t \in [t_a, t_{a+1}]} \left. \int_{t_a}^{t\wedge \eta} U_a(t - s) B(s, u_s) dW_s \right\| \geq \frac{\kappa}{32c} \right) + \mathbb{P}\left( \sum_{i=1}^{m} \sup_{t \in [t_a, t_{a+1}]} \left. \int_{t_a}^{t\wedge \eta} U_a(t - s) \varphi_{\beta_s,i} \gamma_i(s, u_s, \beta_s) dW_s \right\| \geq \frac{\kappa}{32c} \right). \tag{5.28}$$
Recall that Assumptions 10 and 11 are required in order that the isochronal phase \( S \) is well-defined. We can thus use our phase SDE to determine an approximate weak stimulus dominated by a weak non-noisy external stimulus approaches a Von Mises distribution \([70]\). In that work, the structure of the noise.

Rotation induced in the phase through the interaction of the manifold geometry and correlation in the phase over very long periods of time: in effect, we reach an understanding of the average particular subset of the manifold over long periods of time. We can also determine the average shift \( S \) everywhere twice-continuously differentiable. Since \( S \) expects the phase (taken modulo \( \pi \) for all \( \alpha \) in Assumption 11), \( \tilde{\bar{\gamma}} \in H_{jk} \) such that \( \{ \phi_\alpha \}_{\alpha \in S} \) and \( \{ \psi_\alpha \}_{\alpha \in S} \) constitute manifolds that are everywhere twice-continuously differentiable. Since \( S \) is compact, over long periods of time one expects the phase (taken modulo \( S \)) to continually return to any particular region (as long as the noise coefficients are not degenerate). We can thus use our phase SDE to determine an approximate expression for the expected proportion of time that the phase spends in the neighborhood of any particular subset of the manifold over long periods of time. We can also determine the average shift in the phase over very long periods of time: in effect, we reach an understanding of the average rotation induced in the phase through the interaction of the manifold geometry and correlation structure of the noise.

In earlier work we demonstrated that the wandering of a ‘neural bump’ induced by noise and a weak non-noisy external stimulus approaches a Von Mises distribution \([70]\). In that work, the weak stimulus dominated the \( O(\epsilon^2) \) covariance terms. By contrast, in this work there is no weak stimulus, and the long-time distribution of the phase is determined by the balance between the \( O(\epsilon^2) \) covariance terms in the drift, and the \( O(\epsilon) \) stochastic noise.

We assume throughout this section that \( B \) is independent of \( t \), so that we write \( B(u, \cdot) \). Also, we require Assumptions 10 and 11 in order that the isochronal phase SDE in (4.79) is well-defined. Recall that \( \{ \epsilon_j \}_{j \geq 1} \) is the orthonormal basis for \( H \) of Assumption 11. Define, for \( \alpha \in S \), functions \( \tilde{\nu}_i : \mathbb{R}^m \rightarrow \mathbb{R}^m \) and \( \tilde{\nu}_i : \mathbb{R}^m \rightarrow L(H, \mathbb{R}) \) that are (respectively) leading order approximations of \( D^{(2)} \Theta(\varphi_\alpha) \) and \( D^{(1)} \Theta(\varphi_\alpha) \). Writing \( \tilde{\nu}_i = (\tilde{\nu}_i)_i \in \mathbb{I} \leq i \leq m \) and \( \tilde{\nu}_i = (\tilde{\nu}_i)_i \in \mathbb{I} \leq i \leq m \), with \( \tilde{\nu}_i : \mathbb{R}^m \rightarrow L(H, \mathbb{R}) \),

\[
\tilde{\nu}_i(\alpha) = \sum_{j=1}^{m} \langle B^*(\varphi_\alpha) \psi_i^j, B^*(\varphi_\alpha) \psi_i^j \rangle - \frac{1}{2} \sum_{j,k=1}^{m} \langle \varphi_{\alpha,j}, \psi_{\alpha,k}^j \rangle \langle B^*(\varphi_\alpha) \psi_i^j, B^*(\varphi_\alpha) \psi_i^k \rangle
\]

\[+ \Gamma_{QV}(\alpha) \text{ where} \] (6.1)

\[
\Gamma_{QV}(\alpha) = \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \int_0^\infty \langle \psi_i^j, D^{(2)} f(\varphi_\alpha) \cdot (V_\alpha(s) B(\varphi_\alpha) e_j) \cdot (V_\alpha(s) B(\varphi_\alpha) e_j) \rangle ds
\] (6.2)

\[
\tilde{\nu}_i(\alpha) \cdot z = (B(\varphi_\alpha) z, \psi_i^j) \text{ and define} \]

\[
\tilde{\nu}_i(\alpha) = (B^*(\varphi_\alpha) \psi_i^j, B^*(\varphi_\alpha) \psi_i^k) \]. (6.3)

\[
\tilde{\nu}_i(\alpha) = (B^*(\varphi_\alpha) \psi_i^j, B^*(\varphi_\alpha) \psi_i^k). \]

(6.4)

It follows from Assumption 11 and the fact that \( \| V_\alpha(t) \| \leq c \exp(-bt) \) that \( \Gamma_{QV}(\alpha) \) is well-defined for all \( \alpha \in \mathbb{R}^m \). For any \( \xi \in S \), define \( p_t(\alpha|\xi) \) to satisfy the Fokker-Planck PDE, for \( \alpha \in S \),

\[
\frac{\partial}{\partial t} p_t(\alpha|\xi) = -\sum_{j=1}^{m} \frac{\partial}{\partial \alpha_j} \{ p_t(\alpha|\xi) \tilde{\nu}_j(\alpha) \} + \frac{1}{2} \sum_{j,k=1}^{m} \frac{\partial^2}{\partial \alpha_j \partial \alpha_k} \{ \tilde{\nu}_i(\alpha) \cdot \tilde{\nu}_i(\alpha) \}.
\] (6.5)

In the case of spiral waves in compact domains, even if one is working in a rotating reference frame, if the driving white noise \( W_t \) is cylindrical, then one does not need to co-rotate \( B \).
such that \( \lim_{t \to 0} p_t(\alpha; \xi) = \delta_\xi(\alpha) \), and with periodic boundary conditions. The solution of this PDE defines a transition probability density for an \( S \)-valued stochastic process started at \( \xi \). We assume that
\[
\inf_{\alpha \in S} \det \left( \mathcal{H}(\alpha) \right) > 0,
\] (6.6)
and since \( \tilde{V} \) and \( \mathcal{H} \) are continuously differentiable, this means that the stochastic process with Fokker-Plank equation given by (6.5) has a unique invariant density \( p_\star(\alpha) \). That is, \( p_\star(\alpha) \) is the unique solution of
\[
- \sum_{j=1}^{m} \frac{\partial}{\partial \alpha_j} \{ p_\star(\alpha) \tilde{V}_j(\alpha) \} + \frac{1}{2} \sum_{j,k=1}^{m} \frac{\partial^2}{\partial \alpha_j \partial \alpha_k} \{ \mathcal{H}_{jk}(\alpha) p_\star(\alpha) \} = 0,
\] (6.7)
and such that \( \int_S p_\star(\alpha) d\alpha = 1 \). Let \( p_\star \) be the probability measure on \( S \) with density \( p_\star \). Standard theory \( [44] \) dictates that
\[
\lim_{t \to \infty} \sup_{\alpha \in S} | p_t(\alpha; \xi) - p_\star(\alpha) | = 0.
\] (6.8)

The main result of this section is the following. Its implication is that the wandering across the manifold \( S \) of the phase SDE, over long periods of time, is indicated by the density \( p_\star \), with very high probability. For any \( \alpha \in \mathbb{R}^m \), write \( \alpha \mod S \) to be the unique member \( \eta \in S \) such that \( \eta' = \alpha' + 2\pi p_i \), \( \eta \in (-\pi, \pi] \) for integers \( p_i \). Let \( C^2_{\text{b,}S}(\mathbb{R}^m) \) be the set of all periodic twice continuously differentiable functions \( g \) on \( \mathbb{R}^m \), such that
\[
g(\alpha) = g(\alpha \mod S)
\] (6.9)
\[
|g| \leq 1, \quad \left| \frac{\partial g}{\partial \alpha^i} \right| \leq 1, \quad \left| \frac{\partial g}{\partial \alpha^i \partial \alpha^k} \right| \leq 1
\] (6.10)
for all \( 1 \leq j, k \leq m \), and \( \alpha \mapsto \frac{\partial g}{\partial \alpha^j} (\alpha) \) and \( \alpha \mapsto \frac{\partial g}{\partial \alpha^i \partial \alpha^k} (\alpha) \) are continuous on \( S \). This assumption means that, thanks to Taylor’s Theorem, for any \( \alpha, \beta \in S \),
\[
|g(\alpha) - g(\beta)| \leq m \| \alpha - \beta \|.
\] (6.11)

The significance of the first result in the following theorem is that the distribution of the stochastic phase modulo \( S \) over very long periods of time converges to \( p_\star \) as \( \epsilon \to 0 \). The second result determines the average asymptotic shift in the phase over very long periods of time. It parallels analogous results for the long-time average phase-shift of stochastic oscillators \( [40, 69] \).

**Theorem 6.1.** For any \( \delta > 0 \) there exist constants \( C_\delta > 0 \) and \( \epsilon_\delta > 0 \) such that for all \( \epsilon < \epsilon_\delta \) and all \( g \in C^2_{\text{b,}S}(\mathbb{R}^m) \), writing \( T_\delta = \exp \left( C_\delta \epsilon^{-2} \right) \),
\[
P \left( \left| T_\delta^{-1} \int_0^{T_\delta} g(\gamma_s) ds - \mathbb{E}^F_\star [g] \right| > \delta \right) \leq \exp \left( - C_\delta \epsilon^{-2} \right).
\] (6.12)

Also,
\[
P \left( \sup_{1 \leq i \leq m} \left| \epsilon^{-2} T_\delta^{-1} \gamma^i_s - \mathbb{E}^F_\star [\tilde{V}_i] \right| > \delta \right) \leq \exp \left( - C_\delta \epsilon^{-2} \right)
\] (6.13)

**Remark 2.** Effectively, (6.13) implies that over long time scales the isochronal phase \( \gamma^i_s \) changes at average rate \( \epsilon^2 \mathbb{E}^F_\star [\tilde{V}_i] \): that is, if \( \mathbb{E}^F_\star [\tilde{V}_i] \) is not equal to zero, there is a small average oscillation in the phase (taken modulo \( S \)) induced by the noise correlation. In general, this oscillation may only be discernible over timescales of \( O(\epsilon^{-2}) \). Often, if the noise correlation structure satisfies certain symmetries with respect to the manifold, \( \mathbb{E}^F_\star [\tilde{V}_i] \) is identically zero.
6.1 Proof of Theorem 6.1

To demonstrate the lemma, we discretize time into blocks of $\Delta t$ (to be specified more precisely further below, this definition of $\Delta t$ is different from the previous section). Write $t_a := a\Delta t$. To facilitate the proofs, we wish to define a stochastic process $\bm{v}_t^\alpha$ that does not depend on $\epsilon$: $\bm{v}_t^\alpha$ will have the useful property that under the rescaling of time $t \to \epsilon^{-2} t$, its dynamics closely approximates $\gamma_t$. In more detail, for any $\alpha \in S$, define $\bm{v}_t^\alpha$ to be the solution to the stochastic process

$$
d\bm{v}_t^\alpha = \tilde{V}(\bm{v}_t^\alpha)dt + \tilde{Y}(\bm{v}_t^\alpha)d\tilde{W}_t, \quad \text{(6.14)}$$

with initial condition $\bm{v}_0^\alpha = \alpha$. Here $\tilde{W}_t$ is any cylindrical Wiener process with identical probability law to $W_t$.

Define

$$\hat{\tau} = \inf \{ t \geq 0 : \| u_t - \varphi_{\beta_t} \| = \zeta \}, \quad \text{(6.15)}$$

for some $\zeta$ so be specified more precisely below. We assume that $\zeta \in (0, \tilde{\kappa})$, where $\tilde{\kappa}$ is the constant in Theorem 5.1. As long as $\zeta$ is sufficiently small, and since (as proved in (4.58)),

$$\| u_t - \varphi_{\gamma_t} \| = \| u_t - \varphi_{\beta_t} \| + O(\| u_t - \varphi_{\beta_t} \|^2) \quad \text{(6.16)}$$

it must hold that

$$\hat{\tau} \leq \tau_t, \quad \text{(6.17)}$$

and the SDE for $\gamma_t$ is well-defined for all $t \leq \tau$.

Lemma 6.2. For any $\bar{\zeta}, \delta > 0$, there exists $T > 0$ such that for all $t \geq T$,

$$\sup_{\alpha \in S} \mathbb{P}(|t^{-1} \int_0^t g(\bm{v}_s^\alpha)ds - \mathbb{E}^P[g]| \geq \delta/2) \leq \bar{\zeta} \quad \text{(6.18)}$$

Proof. Write $\bar{g}(\theta) = g(\theta) - \mathbb{E}^P[g]$. By Chebyshev’s Inequality,

$$\mathbb{P}(|t^{-1} \int_0^t \bar{g}(\bm{v}_s^\alpha)ds| \geq \delta/2) \leq 4\mathbb{E}[\int_0^t \bar{g}(\bm{v}_s^\alpha)ds^2]/(t\delta)^2$$

$$= \frac{4}{t^2\delta^2} \int_0^t \int_0^t \mathbb{E}[\bar{g}(\bm{v}_s^\alpha)\bar{g}(\bm{v}_r^\alpha)]dsdr$$

$$\leq \frac{4}{t^2\delta^2} \int_0^t \int_0^t \chi(|r-s| \leq T_0)\mathbb{E}[\bar{g}(\bm{v}_r^\alpha)\bar{g}(\bm{v}_s^\alpha)]dsdr$$

$$+ \frac{4}{t^2\delta^2} \int_0^t \int_0^t \chi(|r-s| > T_0)\mathbb{E}[\bar{g}(\bm{v}_r^\alpha)\bar{g}(\bm{v}_s^\alpha)]dsdr,$$

for any $T_0 > 0$. Thanks to the assumption in (6.8), we can choose $T_0$ to be such that

$$4\delta^{-2} \sup_{0 \leq s < t : |t-s| \geq T_0} \{ \mathbb{E}[\bar{g}(\bm{v}_t^\alpha)\bar{g}(\bm{v}_s^\alpha)] \} \leq \bar{\zeta}/2. \quad \text{(6.19)}$$

Also

$$\frac{4}{t^2\delta^2} \int_0^t \int_0^t \chi(|r-s| \leq T_0)\mathbb{E}[\bar{g}(\bm{v}_r^\alpha)\bar{g}(\bm{v}_s^\alpha)]dsdr = O(t^{-1}), \quad \text{(6.20)}$$

since $|\bar{g}|$ is uniformly bounded. The previous three equations imply the lemma.

For any $a \in \mathbb{Z}^+$, and assuming that $\tau > t_a$, define $\tilde{\gamma}_t$ to satisfy the $\mathbb{R}^m$-valued SDE for all $t \in [t_a, t_{a+1})$,

$$d\tilde{\gamma}_t = c^2 \tilde{V}(\tilde{\gamma}_t)dt + c\tilde{Y}(\tilde{\gamma}_t)d\tilde{W}_t, \quad \text{(6.21)}$$
and with initial condition such that \( \tilde{\tau}_{t_a} = \gamma_{t_a} \). Notice that \( \tilde{\gamma} \) is driven by the same Brownian motion \( W_t \) as \( \gamma_t \). One can easily check that a unique solution to the above SDE exists. Furthermore \( \nu^\gamma_a \) has the same probability law as \( \tilde{\gamma}_{t_{a+\epsilon}}(t-t_a) \) for \( t \in [t_a, t_a + T] \). This follows from the fact that they are both Markovian Processes, with identical infinitesimal generators.

We now fix some arbitrary \( \epsilon < 0 \), and let \( T \) be the constant of Lemma 6.2. Write \( \Delta t = \epsilon^{-2}T \). First observe that

\[
P(|T^{-1}_0 \int_0^{T_3} g(\gamma_s)ds - \mathbb{E}^{P^*}[g]| \geq \delta) \leq \mathbb{P}(\hat{\tau} < \exp(C_\delta \epsilon^{-2}))
\]

\[
+ \mathbb{P}(|T^{-1}_0 \int_0^{T_3} g(\gamma_s)ds - \mathbb{E}^{P^*}[g]| \geq \delta \text{ and } \hat{\tau} \geq T_3). \tag{6.22}
\]

Using Lemma 6.1 as long as \( C_\delta \) is small enough,

\[
P(\hat{\tau} < \exp(C_\delta \epsilon^{-2})) \leq \frac{1}{2} \exp(-C_\delta \epsilon^{-2}). \tag{6.23}
\]

It remains to show that

\[
P(|T^{-1}_0 \int_0^{T_3} g(\gamma_s)ds - \mathbb{E}^{P^*}[g]| \geq \delta \text{ and } \hat{\tau} \geq T_3) \leq \frac{1}{2} \exp(C_\delta \epsilon^{-2}). \tag{6.24}
\]

By Chebyshev’s Inequality, for some \( \kappa > 0 \),

\[
P(|T^{-1}_0 \int_0^{T_3} g(\gamma_s)ds - \mathbb{E}^{P^*}[g]| \geq \delta \text{ and } \hat{\tau} \geq T_3)
\]

\[
\leq \mathbb{E} \left[ \chi(\hat{\tau} \geq T_3) \left\{ \exp \left( \kappa \int_0^{T_3} g(\gamma_s)ds - \kappa \hat{\tau}\mathbb{E}^{P^*}[g] - \kappa \hat{\tau} \delta \right) + \exp \left( -\kappa \int_0^{T_3} g(\gamma_s)ds + \kappa T_3 \mathbb{E}^{P^*}[g] - \kappa T_3 \delta \right) \right\} \right]. \tag{6.25}
\]

We next show that, for a particular choice of \( \kappa \),

\[
\mathbb{E} \left[ \chi(\hat{\tau} \geq T_3) \exp \left( \kappa \int_0^{T_3} g(\gamma_s)ds - \kappa T_3 \mathbb{E}^{P^*}[g] - \kappa T_3 \delta \right) \right] \leq \exp \left( -\text{Const} \times \epsilon^{-2} \right) \tag{6.26}
\]

and omit the analogous proof that

\[
\mathbb{E} \left[ \exp \left( -\kappa \int_0^{T_3} g(\gamma_s)ds + \kappa T_3 \mathbb{E}^{P^*}[g] - \kappa T_3 \delta \right) \right] \leq \exp \left( -\text{Const} \times \epsilon^{-2} \right). \tag{6.27}
\]

Now define the random variable

\[
H_a = \exp \left\{ \kappa \{ \hat{\tau} \geq t_{a+1} \} \left( \int_{t_a}^{t_{a+1}} g(\gamma_s)ds - \Delta t \mathbb{E}^{P^*}[g] \right) \right\},
\]

and define \( a \) to be the random index such that \( \hat{\tau} \in [t_a, t_{a+1}) \). We thus find that

\[
\mathbb{E}[\chi(\hat{\tau} \geq T_3)] \exp \left( \kappa \int_0^{T_3} g(\gamma_s)ds - \kappa \hat{\tau} \mathbb{E}^{P^*}[g] \right)
\]

\[
= \mathbb{E}[\chi(\hat{\tau} \geq T_3)] \prod_{a=0}^{a} H_a \times \exp \left( \kappa \int_{t_a}^{\hat{\tau}} g(\gamma_s)ds - \kappa (\hat{\tau} - t_a) \mathbb{E}^{P^*}[g] \right)
\]

\[
\leq \mathbb{E}[\chi(\hat{\tau} \geq T_3)] \prod_{a=0}^{a} H_a \exp (2\kappa \Delta t),
\]

34
since $|g| \leq 1$ uniformly. Furthermore
\[
\mathbb{P}((\Delta t)^{-1} \int_{t_a}^{t_{a+1}} g(\gamma_s) ds - \mathbb{E}^{P_\tau} [g] \geq \delta \text{ and } \hat{\tau} \geq t_{a+1}) \\
\leq \mathbb{P}((\Delta t)^{-1} \int_{t_a}^{t_{a+1}} g(\gamma_s) ds - \mathbb{E}^{P_\tau} [g] \geq \delta/2) \\
+ \mathbb{P}((\Delta t)^{-1} \int_{t_a}^{t_{a+1}} |g(\gamma_s) - g(\gamma_s)| ds \geq \delta/2 \text{ and } \hat{\tau} \geq t_{a+1}).
\]

Since $v_{\epsilon^2(t-t_a)}$ has the same probability law as $\tilde{\gamma}_{t-t_a}$, it follows that for any $\bar{\epsilon} > 0$,
\[
\mathbb{P}((\Delta t)^{-1} \int_{t_a}^{t_{a+1}} g(\gamma_s) ds - \mathbb{E}^{P_\tau} [g] \geq \delta/2) = \mathbb{P}(\int_{0}^{T} g(v_{\gamma_s}) ds - \mathbb{E}^{P_\tau} [g] \geq \delta/2) \leq \bar{\epsilon},
\]
thanks to Lemma [6.2] as long as $\epsilon$ is small enough.

Furthermore since $g$ has Lipschitz constant upperbouned by $m$ (noted in [6.11]),
\[
\mathbb{P}((\Delta t)^{-1} \int_{t_a}^{t_{a+1}} |g(\gamma_s) - g(\gamma_s)| ds \geq \delta/2 \text{ and } \hat{\tau} \geq t_{a+1}) \\
\leq \mathbb{P} \left( \sup_{t \in \{t_a, t_{a+1}\}} \|\gamma_t - \tilde{\gamma}_t\| \geq \delta/(2m) \text{ and } \hat{\tau} \geq t_{a+1} \right) \leq \bar{\epsilon},
\]
using the result of Lemma [6.3]

Let $\hat{\alpha} = |(\Delta t)^{-1} T\delta| - 1$. The previous two results imply that
\[
\mathbb{E} \left[ \chi \{ \hat{\tau} \geq T\delta \} \prod_{a=0}^{n-1} H_a \right] \leq \left[ 2\bar{\epsilon} \exp(\kappa \Delta t) + (1 - 2\bar{\epsilon}) \exp(\kappa \delta) \right] \hat{\alpha} \exp\left( - \kappa \delta T\delta \right)
\]
\[
= \exp(\tilde{\alpha} \kappa \delta - \kappa \delta T\delta) \{1 + 2\bar{\epsilon}(\exp(\kappa \Delta t - \kappa \delta) - 1)\} \hat{\alpha}
\]
\[
\leq \exp \left\{ \tilde{\alpha} \kappa \delta + 2\bar{\epsilon} (\exp(\kappa \Delta t - \kappa \delta) - 1) - \kappa \delta T\delta \right\}. \tag{6.28}
\]

We substitute $\kappa = \epsilon^2$. We take $T$ to be large enough that
\[
\hat{\alpha} \leq T\delta/3, \tag{6.29}
\]
and we take $\bar{\epsilon}$ to be small enough that
\[
2\bar{\epsilon}(\exp(T - \epsilon^2 \delta) - 1) \leq \delta \epsilon^{-2} T\delta/3, \tag{6.30}
\]
which is always possible, since $T\delta = \exp(C\delta \epsilon^{-2})$. In this way we find that
\[
\exp \left\{ \tilde{\alpha} \kappa \delta + 2\bar{\epsilon} (\exp(\kappa \Delta t - \kappa \delta) - 1) - \kappa \delta T\delta \right\} \leq \exp \left\{ - \epsilon^{-2} \delta \exp(-C\delta \epsilon^{-2})/3 \right\} \leq \frac{1}{2} \exp\left( - C\delta \epsilon^{-2} \right),
\]
for small enough $\epsilon$. Substituting this into (6.25), we have proved (6.24), i.e.
\[
\mathbb{P} \left( \int_0^{T\delta} g(\gamma_s) ds - \mathbb{E}^{P_\tau} [g] \geq \delta \text{ and } \hat{\tau} \geq T\delta \right) \leq \exp(-C\delta \epsilon^{-2}) \tag{6.31}
\]
We have thus established (6.12).

We now prove (6.13).
Proof. Notice that, substituting the SDE for $\gamma_t$ in Lemma 4.6 and applying a union of events bound,

$$\mathbb{P}\left(\sup_{1 \leq i \leq m} |\epsilon^{-2}T_{\delta}^{-1}\gamma_{\delta}^i - \mathbb{E}^{\mathcal{F}_\tau}[\tilde{Y}_i]| > \delta\right) \leq \mathbb{P}(\hat{\tau} \leq T_\delta)$$

$$+ \mathbb{P}(\hat{\tau} > T_\delta \text{ and } \sup_{1 \leq i \leq m} \int_0^{T_\delta} D(\Theta_i(u_t) \cdot B(u_t)e_j \cdot B(u_t)e_j dt - \mathbb{E}^{\mathcal{F}_\tau}[\tilde{Y}_i]) > \delta/2)$$

$$+ \mathbb{P}(\hat{\tau} > T_\delta \text{ and } \sup_{1 \leq i \leq m} \int_0^{T_\delta} D\Theta_i(u_t)dW_t > \delta/2). \quad (6.32)$$

We have already seen that $\mathbb{P}(\hat{\tau} \leq T_\delta) \leq \exp(-C_{\delta}\epsilon^{-2})$. For the second term, employing a union-of-events bound, and Chernoff’s Inequality, for a constant $\kappa > 0$,

$$\mathbb{P}(\hat{\tau} > T_\delta \text{ and } \sup_{1 \leq i \leq m} \int_0^{T_\delta} D\Theta_i(u_t)dW_t > \delta/2) \leq$$

$$\sum_{1 \leq i \leq m} \mathbb{P}(\hat{\tau} \geq T_\delta \text{ and } \epsilon^{-1}T_{\delta}^{-1}\int_0^{T_\delta} D\Theta_i(u_t)dW_t > \delta/2) +$$

$$\sum_{1 \leq i \leq m} \mathbb{P}(\hat{\tau} > T_\delta \text{ and } \epsilon^{-1}T_{\delta}^{-1}\int_0^{T_\delta} D\Theta_i(u_t)dW_t < -\delta/2)$$

$$\leq \sum_{1 \leq i \leq m} \mathbb{E}[\chi_{\{\hat{\tau} > T_\delta\}} \exp \left\{ \kappa \int_0^{T_\delta} D\Theta_i(u_t)dW_t - \kappa\delta\epsilon T_\delta/2 \right\} +$$

$$\chi_{\{\hat{\tau} > T_\delta\}} \exp \left\{ -\kappa \int_0^{T_\delta} D\Theta_i(u_t)dW_t - \kappa\delta\epsilon T_\delta/2 \right\}]. \quad (6.33)$$

Using the bound in Lemma 4.5 as long as $\zeta$ is sufficiently small, $|D\Theta_i(u_t)|$ is uniformly bounded as long as $t \leq \hat{\tau}$, and we thus find that

$$\frac{1}{2} \int_0^{T_\delta} \{D\Theta_i^2(u_t)\}^2 dt \leq \bar{C}T_\delta,$$

for a constant $\bar{C} > 0$. We thus find that

$$\mathbb{E}[\chi_{\{\hat{\tau} > T_\delta\}} \exp \left\{ \kappa \int_0^{T_\delta} D\Theta_i(u_t)dW_t - \kappa\delta\epsilon T_\delta/2 \right\} \leq \mathbb{E}[\chi_{\{\hat{\tau} > T_\delta\}} \exp \left\{ \kappa \int_0^{T_\delta} D\Theta_i(u_t)dW_t - \kappa\delta\epsilon T_\delta/2 \right\}$$

$$- \frac{\kappa^2}{2} \int_0^{T_\delta} \{D\Theta_i(u_t)\}^2 dt + \kappa^2\bar{C}T_\delta - \kappa\delta\epsilon T_\delta/2 ] = \exp \left\{ T_\delta\kappa^2\bar{C} - \kappa\delta\epsilon T_\delta/2 \right\},$$

by Girsanov’s Theorem. We choose $\kappa = \delta\epsilon/(4\bar{C})$, and we obtain that

$$\mathbb{E}[\chi_{\{\hat{\tau} > T_\delta\}} \exp \left\{ \kappa \int_0^{T_\delta} D\Theta_i(u_t)dW_t - \kappa\delta\epsilon T_\delta/2 \right\} \leq \exp \left\{ -\delta^2\epsilon^2 T_\delta/(8\bar{C}) \right\}. \quad (6.34)$$

Since each of the terms in (6.33) can be bounded in the same manner as the above, we obtain that

$$\mathbb{P}\left(\sup_{1 \leq i \leq m} \int_0^{T_\delta} D\Theta_i(u_t)dW_t \geq \epsilon/2 \right) \leq 2m \exp \left\{ -\delta^2\epsilon^2 T_\delta/(8\bar{C}) \right\}. \quad (6.35)$$

Since $T_\delta = \exp(C_{\delta}\epsilon^{-2})$, in the limit as $\epsilon \to 0$, $\epsilon^2 T_\delta \gg \epsilon^{-2}$. For the remaining term in (6.32), using
a union of events bound,

\[ \mathbb{P}(\hat{\tau} > T_\delta \text{ and } \sup_{1 \leq i \leq m} |1/(2T_\delta) \sum_{j=1}^{\infty} \int_0^{T_\delta} D(2) \Theta_i(u_t) \cdot B(u_t)e_j \cdot B(u_t)e_j dt - E^{P_i}[\tilde{V}_i]| > \delta/2) \leq \]

\[ \mathbb{P}(\hat{\tau} > T_\delta \text{ and } \sup_{1 \leq i \leq m} \left\{ T_\delta^{-1} \int_0^{T_\delta} \left| \frac{1}{2} \sum_{j=1}^{\infty} D(2) \Theta_i(u_t) \cdot B(u_t)e_j \cdot B(u_t)e_j - \tilde{V}_i(\gamma_t) \right| |dt > \delta/4 \right\} \]

\[ + \mathbb{P}(\hat{\tau} > T_\delta \text{ and } \sup_{1 \leq i \leq m} |T_\delta^{-1} \int_0^{T_\delta} \tilde{V}_i(\gamma_t) dt - E^{P_i}[\tilde{V}_i]| > \delta/4) \]  \( (6.36) \)

Now for small enough \( \zeta \), thanks to Lemma 4.5 whenever \( t \leq \hat{\tau} \)

\[ \left| \frac{1}{2} \sum_{j=1}^{\infty} D(2) \Theta_i(u_t) \cdot B(u_t)e_j \cdot B(u_t)e_j - \tilde{V}_i(\gamma_t) \right| \leq \delta/4, \]

and the first term on the right hand side of (6.36) has probability zero.

Using the first result of the lemma, i.e. (6.12), since \( \tilde{V}_i \) is smooth and bounded,

\[ \mathbb{P}(\hat{\tau} > T_\delta \text{ and } \sup_{1 \leq i \leq m} |T_\delta^{-1} \int_0^{T_\delta} \tilde{V}_i(\gamma_t) dt - E^{P_i}[\tilde{V}_i]| > \delta/4) \leq \exp (-C_\delta X^{-2} \epsilon^{-2}), \]

where

\[ X = \sup_{1 \leq i,j,k \leq m, \alpha \in \mathbb{R}^m} \left\{ \left| \tilde{V}_i(\alpha) \right|, \left| \partial \alpha \tilde{V}_i(\alpha) \right|, \left| \partial^2 \alpha \tilde{V}_i(\alpha) \right| \right\}. \]

In summary, collecting all of the above terms, we have proved that for small enough \( \epsilon > 0 \), there is a constant such that

\[ \mathbb{P} \left( \sup_{1 \leq i \leq m} \epsilon^{-2} T_\delta^{-1} |\beta_t - E^{P_i}[\tilde{V}_i]| > \delta \right) \leq \exp (-\text{Const} \times \epsilon^{-2}). \]  \( (6.37) \)

**Lemma 6.3.** For any \( \bar{\epsilon}, \bar{\epsilon} > 0 \), there exists \( \zeta > 0 \) and \( \epsilon_0, C_\delta > 0 \) such that for all \( \epsilon < \epsilon_0 \),

\[ \mathbb{P} \left( \sup_{t \in [u,t_{u+1}]} \| \gamma_t - \tilde{\gamma}_t \| \geq \bar{\epsilon} \text{ and } \sup_{t \in [u,t_{u+1}]} \| u_t \| \leq \zeta < \bar{\epsilon} \right) \leq \frac{1}{2} \exp (-C_\delta \epsilon^{-2}). \]  \( (6.38) \)

**Proof.** It follows from Lemma 4.5 that

\[ \left\| \frac{1}{2} \sum_{j=1}^{\infty} D(2) \Theta_i(u_t) \cdot B(u_t)e_j \cdot B(u_t)e_j - \tilde{V}_i(\gamma_t) \right\| = O \left( \left\| u - \varphi_{\gamma_t} \right\| \right) = O \left( \left\| u - \varphi_{\beta_t} \right\| \right). \]  \( (6.40) \)

The rest of the proof then follows using similar techniques to the rest of the paper, and is omitted. \( \square \)

**References**

[1] J. Armero, J. Casademunt, L. Ramírez-Piscina, and J. M. Sancho, *Ballistic and diffusive corrections to front propagation in the presence of multiplicative noise*, Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics, 58 (1998), pp. 5494–5500.
[2] D. Avitable and K. C. Wedgwood, *Macroscopic coherent structures in a stochastic neural network: from interface dynamics to coarse-grained bifurcation analysis*, Journal of Mathematical Biology, 75 (2017), pp. 885–928.

[3] D. Barkley, *Linear Stability Analysis of Rotating Spiral Waves in Excitable Media*, Program in Applied and Computational Mathematics, 68 (1992), pp. 2090–2093.

[4] ———, *Euclidean symmetry and the dynamics of rotating spiral waves*, Physical Review Letters, 72 (1994), pp. 164–167.

[5] F. Barret, *Sharp asymptotics of metastable transition times for one dimensional SPDEs*, Annales de l’institut Henri Poincare (B) Probability and Statistics, 51 (2015), pp. 129–166.

[6] N. Berglund and B. Gentz, *Sharp estimates for metastable lifetimes in parabolic SPDEs: Kramers’ law and beyond*, Electronic Journal of Probability, 18 (2013).

[7] W. J. Beyn and V. Thümmler, *Freezing solutions of equivariant evolution equations*, SIAM Journal on Applied Dynamical Systems, 3 (2004), pp. 85–116.

[8] T. Biancalani, D. Fanelli, and F. Di Patti, *Stochastic turing patterns in the Brusselator model*, Physical Review E - Statistical, Nonlinear, and Soft Matter Physics, 81 (2010), pp. 1–8.

[9] T. Biancalani, T. Galla, and A. J. McKane, *Stochastic waves in a Brusselator model with nonlocal interaction*, Physical Review E - Statistical, Nonlinear, and Soft Matter Physics, 84 (2011).

[10] T. Biancalani, F. Jafarpour, and N. Goldenfeld, *Giant Amplification of Noise in Fluctuation-Induced Pattern Formation*, Physical Review Letters, 118 (2017), pp. 1–5.

[11] D. Blomker, *Amplitude Equations for Stochastic Partial Differential Equations*, World Scientific, 2007.

[12] D. Blömker, M. Hairer, and G. A. Pavliotis, *Multiscale analysis for stochastic partial differential equations with quadratic nonlinearities*, Nonlinearity, 20 (2007), pp. 1721–1744.

[13] S. Brassesco, A. De Masi, and E. Presutti, *Brownian fluctuations of the interface in the D=1 Ginzburg-Landau equation with noise*, Annales de l’I.H.P. Section B, 31 (1995), pp. 81–118.

[14] P. C. Bressloff, *Spatiotemporal dynamics of continuum neural fields*, Journal of Physics A: Mathematical and Theoretical, 45 (2012).

[15] P. C. Bressloff, *Stochastic neural field theory of wandering bumps on a sphere*, Physica D: Nonlinear Phenomena, (2019).

[16] P. C. Bressloff and S. R. Carroll, *Stochastic neural fields as gradient dynamical systems*, Physical Review E, 100 (2019), pp. 1–23.

[17] P. C. Bressloff, J. D. Cowan, M. Golubitsky, P. J. Thomas, and M. C. Wiener, *What geometric visual hallucinations tell us about the visual cortex*, Neural Computation, 14 (2002), pp. 473–489.

[18] P. C. Bressloff and Z. P. Kilpatrick, *Nonlinear Langevin equations for wandering patterns in stochastic neural fields*, SIAM Journal on Applied Dynamical Systems, 14 (2015), pp. 305–334.

[19] P. C. Bressloff and J. MacLaurin, *Synchronization of stochastic hybrid oscillators driven by a common switching environment*, Chaos: An Interdisciplinary Journal of Nonlinear Science, (2019), pp. 1–16.
[20] E. Brunet, B. Derrida, A. H. Mueller, and S. Munier, *Phenomenological theory giving the full statistics of the position of fluctuating pulled fronts*, Physical Review E - Statistical, Nonlinear, and Soft Matter Physics, 73 (2006), pp. 1–9.

[21] Z. Brzeźniak and S. Peszat, *Maximal Inequalities for Stochastic Convolutions*, in Stochastic Processes, Physics and Geometry: New Interplays. I, F. Gesztesy, H. Holden, J. Jost, S. Paycha, M. Rockner, and S. Scarlatti, eds., Canadian Mathematical Society, 1999.

[22] S. R. Carroll and P. C. Bressloff, *Phase equation for patterns of orientation selectivity in a neural field model of visual cortex*, SIAM Journal on Applied Dynamical Systems, 15 (2016), pp. 60–83.

[23] M. Cartwright and G. A. Gottwald, *A collective coordinate framework to study the dynamics of travelling waves in stochastic partial differential equations*, Physica D: Nonlinear Phenomena, 397 (2019), pp. 54–64.

[24] S. Cerrai and M. Röckner, *Large deviations for stochastic reaction-diffusion systems with multiplicative noise and non-lipschitz reaction term*, Annals of Probability, 32 (2004), pp. 1100–1139.

[25] G. W. Constable, A. J. McKane, and T. Rogers, *Stochastic dynamics on slow manifolds*, Journal of Physics A: Mathematical and Theoretical, 46 (2013).

[26] S. Coombes, P. beim Graben, R. Potthast, and J. Wright, *Neural Fields: Theory and Applications*, Springer, 2014.

[27] G. Da Prato and J. Zabczyk, *Stochastic Equations in Infinite Dimensions. Second Edition*, Cambridge University Press, 2014.

[28] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications 2nd Edition*, Springer, 1998.

[29] C. R. Doering, C. Mueller, and P. Smereka, *Interacting particles, the stochastic Fisher-Kolmogorov-Petrovsky-Piscounov equation, and duality*, Physica A: Statistical Mechanics and its Applications, 325 (2003), pp. 243–259.

[30] M. Donsker and S. Varadhan, *Asymptotic Evaluation of Certain Markov Process Expectations for Large Time. II*, Communications on Pure and Applied Mathematics, 28 (1975).

[31] K. Eichinger, M. V. Gnann, and C. Kuehn, *Multiscale analysis for traveling-pulse solutions to the stochastic FitzHugh-Nagumo equations*, Arxiv Preprint, (2020), pp. 1–41.

[32] G. B. Ermentrout, S. E. Folias, and Z. P. Kilpatrick, *Spatiotemporal Pattern Formation in Neural Fields with Linear Adaptation*, in Neural Fields, P. Greim Graben, S. Coombes, R. Potthast, and J. Wright, eds., Springer-Verlag, 2014.

[33] G. B. Ermentrout and J. B. McLeod, *Existence and uniqueness of travelling waves for a neural network*, Proceedings of the Royal Society of Edinburgh: Section A Mathematics, 123 (1993), pp. 461–478.

[34] G. B. Ermentrout and D. H. Terman, *Mathematical Foundations of Neuroscience*, Springer, 2010.

[35] G. Faye and A. Scheel, *Existence of pulses in excitable media with nonlocal coupling*, Advances in Mathematics, 270 (2015), pp. 400–456.

[36] S. E. Folias and P. C. Bressloff, *Stimulus-Locked Traveling Waves and Breathers in an Excitatory Neural Network*, SIAM Journal of Applied Mathematics, 65 (2005), pp. 2067–2092.
[37] M. I. Freidlin and A. D. Wentzell, Random Perturbations of Dynamical Systems. 3rd Edition, Springer Heidelberg, 2012.

[38] C. C. Fung, K. Y. Wong, and S. Wu, A moving bump in a continuous manifold: A comprehensive study of the tracking dynamics of continuous attractor neural networks, Neural Computation, 22 (2010), pp. 752–792.

[39] C. Geldhauser and C. Kuehn, Travelling waves for discrete stochastic bistable equations, Arxiv Preprint, (2020), pp. 1–21.

[40] G. Giacomin, C. Poquet, and A. Shapira, Small noise and long time phase diffusion in stochastic limit cycle oscillators, Journal of Differential Equations, 264 (2018), pp. 1019–1049.

[41] P. Gong, S. T. Loi, P. A. Robinson, and C. Y. Yang, Spatiotemporal pattern formation in two-dimensional neural circuits: Roles of refractoriness and noise, Biological Cybernetics, 107 (2013), pp. 1–13.

[42] G. A. Gottwald and L. Kramer, On propagation failure in one- and two-dimensional excitable media, Chaos, 14 (2004), pp. 855–863.

[43] T. Grafke and E. Vanden-Eijnden, Non-equilibrium transitions in multiscale systems with a bifurcating slow manifold, Journal of Statistical Mechanics: Theory and Experiment, 2017 (2017).

[44] M. Hairer and J. C. Mattingly, Yet another look at Harris’ ergodic theorem for Markov chains, in Seminar on Stochastic Analysis, Random Fields and Applications VI, Springer, 2011, pp. 109–117.

[45] C. Hamster and H. J. Hupkes, Stability of Travelling Waves on Exponentially Long Timescales in Stochastic Reaction-Diffusion Equations, 2 (2020), pp. 1–25.

[46] C. H. Hamster and H. J. Hupkes, Stability of Traveling Waves for Systems of Reaction-Diffusion Equations with Multiplicative Noise, SIAM Journal of Mathematical Analysis, 52 (2020), pp. 1386–1426.

[47] ———, Travelling waves for reaction–diffusion equations forced by translation invariant noise, Physica D: Nonlinear Phenomena, 401 (2020), p. 132233.

[48] J. W. Harris, S. C. Harris, and A. E. Kyprianou, Further probabilistic analysis of the Fisher-Kolmogorov-Petrovskii-Piscounov equation: One sided travelling-waves, Annales de l’institut Henri Poincare (B) Probability and Statistics, 42 (2006), pp. 125–145.

[49] A. Hutt, A. Longtin, and L. Schimansky-Geier, Additive noise-induced Turing transitions in spatial systems with application to neural fields and the Swift-Hohenberg equation, Physica D: Nonlinear Phenomena, 237 (2008), pp. 755–773.

[50] J. Inglis and J. Maclaurin, A General Framework for Stochastic Traveling Waves and Patterns, with Application to Neural Field Equations, SIAM Journal on Applied Dynamical Systems, 15 (2016), pp. 195–234.

[51] T. Kapitula and K. Promislow, Spectral and dynamical stability of nonlinear waves, vol. 185, Springer, 2013.

[52] I. Karatzas and S. Shreve, Brownian Motion and Stochastic Calculus 2nd Edition, 1991.

[53] G. Katzenberger, Solutions of a Stochastic Differential Equation Forced Onto a Manifold by a Large Drift, The Annals of Probability, 19 (1991), pp. 1587–1628.
[54] A. Keane and P. Gong, Propagating waves can explain irregular neural dynamics, Journal of Neuroscience, 35 (2015), pp. 1591–1605.

[55] Z. P. Kilpatrick, Delay stabilizes stochastic motion of bumps in layered neural fields, Physica D: Nonlinear Phenomena, 295-296 (2015), pp. 30–45.

[56] Z. P. Kilpatrick and B. Ermentrout, Wandering bumps in stochastic neural fields, SIAM Journal on Applied Dynamical Systems, 12 (2013), pp. 61–94.

[57] Z. P. Kilpatrick, B. Ermentrout, and B. Dohrón, Optimizing Working Memory with Heterogeneity of Recurrent Cortical Excitation, 33 (2013), pp. 18999–19011.

[58] J. Krüger and W. Stannat, Front Propagation in Stochastic Neural Fields: A Rigorous Mathematical Framework, SIAM Journal on Applied Dynamical Systems, 13 (2014), pp. 1293–1310.

[59] ———, A multiscale-analysis of stochastic bistable reaction–diffusion equations, Nonlinear Analysis, Theory, Methods and Applications, 162 (2017), pp. 197–223.

[60] C. Kuehn, Travelling Waves in Monostable and Bistable Stochastic Partial Differential Equations, Jahresbericht der Deutschen Mathematiker-Vereinigung, (2019).

[61] C. Kuehn and M. G. Riedler, Large deviations for nonlocal stochastic neural fields, Journal of Mathematical Neuroscience, 4 (2014), pp. 1–33.

[62] C. Kuehn and J. M. Tölle, A gradient flow formulation for the stochastic Amari neural field model, Journal of Mathematical Biology, 79 (2019), pp. 1227–1252.

[63] E. Lang and W. Stannat, L2-stability of traveling wave solutions to nonlocal evolution equations, Journal of Differential Equations, 261 (2016), pp. 4275–4297.

[64] M. A. Lewis, Spread rate for a nonlinear stochastic invasion, Journal of Mathematical Biology, 41 (2000), pp. 430–454.

[65] D. T. Liley, Neural Field Modelling of the Electroencephalogram: Physiological Insights and Practical Applications, in Neural Fields: Theory and Applications, S. Coombes, P. Potthast, P. beim Graben, and J. Wright, eds., Springer, 2014.

[66] B. Lindner, J. Garcia-Ojalvo, A. Neiman, and L. Schimansky-Geier, Effects of noise in excitable systems, Physics Reports, 392 (2004), pp. 321–424.

[67] W. Liu and M. Rockner, Stochastic Partial Differential Equations: An Introduction, Springer, 2015.

[68] G. J. Lord and V. Thummler, Freezing Stochastic Travelling Waves, SIAM Journal on Scientific Computing, 34 (2012).

[69] J. MacLaurin and A. Singh, The Effect of Stochastic Bursting on Biological Clock Precision, in Proceedings of the European Control Conference, 2020.

[70] J. N. MacLaurin and P. C. Bressloff, Wandering bumps in a stochastic neural field: A variational approach, Physica D: Nonlinear Phenomena, 406 (2020), p. 132403.

[71] N. V. Medathati, J. Rankin, A. I. Meso, P. Kornprobst, and G. S. Masson, Recurrent network dynamics reconciles visual motion segmentation and integration, Scientific Reports, 7 (2017), pp. 1–15.

[72] A. S. Mikhailov, L. Schimansky-Geier, and W. Ebeling, Stochastic motion of the propagating front in bistable media, Physics Letters A, 96 (1983), pp. 453–456.
[73] K. N. Mukta, J. N. MacLaurin, and P. A. Robinson, Theory of corticothalamic brain activity in a spherical geometry: Spectra, coherence, and correlation, Physical Review E, 96 (2017), p. 123.

[74] H. Nakao, T. Yanagita, and Y. Kawamura, Phase-Reduction Approach to Synchronization of Spatiotemporal Rhythms in Reaction-Diffusion Systems, Physical Review X, 4 (2014).

[75] D. Panja, Effects of fluctuations on propagating fronts, Physics Reports, 393 (2004), pp. 87–174.

[76] T. L. Parsons and T. Rogers, Dimension reduction for stochastic dynamical systems forced onto a manifold by large drift: A constructive approach with examples from theoretical biology, Journal of Physics A: Mathematical and Theoretical, 50 (2017).

[77] D. J. Pinto and G. B. Ermentrout, Spatially Structured Activity in Synaptically Coupled Neuronal Networks: I. Traveling Fronts and Pulses, SIAM Journal of Applied Mathematics, 62 (2001), pp. 206–225.

[78] L. Popovic and L. Peuckert, Diffusion dynamics on the coexistence subspace in a stochastic evolutionary game, Journal of Mathematical Biology, 80 (2020), pp. 1655–1682.

[79] F. Sagués, J. M. Sancho, and J. García-Ojalvo, Spatiotemporal order out of noise, Reviews of Modern Physics, 79 (2007), pp. 829–882.

[80] M. Salins, A. Budhiraja, and P. Dupuis, Uniform Large Deviation Principles for Banach Space Valued Stochastic Evolution Equations, Transactions of the American Mathematical Society, 372 (2019), pp. 8363–8421.

[81] M. Salins and K. Spiliopoulos, Metastability and exit problems for systems of stochastic reaction-diffusion equations, Arxiv Preprint, (2019), pp. 1–67.

[82] B. Sandstede, A. Scheel, and C. Wulff, Dynamics of spiral waves on unbounded domains using center-manifold reductions, Journal of Differential Equations, 141 (1997), pp. 122–149.

[83] R. Sowers, Large Deviations for a Reaction-Diffusion Equation with Non-Gaussian Perturbations, The Annals of Probability, 20 (1992), pp. 504–537.

[84] W. Stannat, Stability of travelling waves in stochastic Nagumo equations, Arxiv Preprint, 6 (2014), pp. 1–28.

[85] R. Thul, S. Coombes, and C. R. Laing, Neural Field Models with Threshold Noise, Journal of Mathematical Neuroscience, 6 (2016).

[86] V. K. Vanag and I. R. Epstein, Inwardly Rotating Spiral Waves in a Reaction-Diffusion System, Science, 294 (2019), pp. 835–837.

[87] S. Visser, R. Nicks, O. Faugeras, and S. Coombes, Standing and travelling waves in a spherical brain model: The Nunez model revisited, Physica D: Nonlinear Phenomena, 349 (2017), pp. 27–45.

[88] A. Volpert, V. Volpert, and V. Volpert, Traveling Wave Solutions of Parabolic Systems, vol. 140, American Mathematical Society, 1994.

[89] H. R. Wilson and J. D. Cowan, Excitatory and Inhibitory Interactions in Localized Populations of Model Neurons, Biophysical Journal, 12 (1972), pp. 1–24.

[90] B. Zdzislaw, B. Goldys, and N. Rana, Large Deviations for Stochastic Geometric Wave Equation, Arxiv Preprint, (2019).