A Generalized Lyapunov Construction for Proving Stabilization by Noise

by

Tiffany N. Kolba

Department of Mathematics
Duke University

Date: ______________________
Approved:

____
Jonathan C. Mattingly, Supervisor

____
Mauro Maggioni

____
James Nolen

____
Thomas Witelski

Dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in the Department of Mathematics
in the Graduate School of Duke University
2012

Abstract

A Generalized Lyapunov Construction for Proving Stabilization by Noise

by

Tiffany N. Kolba

Department of Mathematics
Duke University

Date: ______________________

Approved:

____
Jonathan C. Mattingly, Supervisor

____
Mauro Maggioni

____
James Nolen

____
Thomas Witelski

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Mathematics in the Graduate School of Duke University
2012

Abstract

Noise-induced stabilization occurs when an unstable deterministic system is stabilized by the addition of white noise. Proving that this phenomenon occurs for a particular system is often manifested through the construction of a global Lyapunov function. However, the procedure for constructing a Lyapunov function is often quite ad hoc, involving much time and tedium. In this thesis, a systematic algorithm for the construction of a global Lyapunov function for planar systems is presented. The general methodology is to construct a sequence of local Lyapunov functions in different regions of the plane, where the regions are delineated by different behaviors of the deterministic dynamics. A priming region, where the deterministic drift is directed inward, is first identified where there is an obvious choice for a local Lyapunov function. This priming Lyapunov function is then propagated to the other regions through a series of Poisson equations. The local Lyapunov functions are lastly patched together to form one smooth global Lyapunov function.

The algorithm is applied to a model problem which displays finite time blow up in the deterministic setting in order to prove that the system exhibits noise-induced stabilization. Moreover, the Lyapunov function constructed is in fact what we define to be a super Lyapunov function. We prove that the existence of a super Lyapunov function, along with a minorization condition, implies that the corresponding system converges to a unique invariant probability measure at an exponential rate that is independent of the initial condition.

Contents

List of Figures

List of Abbreviations and Symbols

Symbols

$\mathbb{R}$	The set of real numbers
$\mathbb{C}$	The set of complex numbers
$\mathbb{N}$	The set of natural numbers
C	The set of continuous functions
C^1	The set of continuously differentiable functions
C^2	The set of twice continuously differentiable functions
C^∞	The set of infinitely continuously differentiable functions
$B_r(z)$	The open ball of radius r about the point z
B_r	The open ball of radius r about the point 0
$\mathbb{P}_{\mathbb{x}}$	Probability starting from initial condition $\mathbb{x}$
$\mathbb{E}_{\mathbb{x}}$	Expectation starting from initial condition $\mathbb{x}$
$\lambda(\cdot)$	Lebesgue measure

Abbreviations

ODE	ordinary differential equation
PDE	partial differential equation
SDE	stochastic differential equation

Acknowledgements

This thesis is the culmination of much effort over my six years of graduate study at Duke University and there are numerous people whose contributions have made this work a possibility. First and foremost, I would like to thank my advisor, Dr. Jonathan Mattingly. He has been a wonderful mentor and has helped shape the probabilistic lens through which I view mathematical problems. Many thanks also go to the other members of my dissertation committee, Drs. Mauro Maggioni, James Nolen, and Thomas Witelski.

I would like to express sincere gratitude to Avanti Athreya, whose insight and guidance were invaluable for the completion of this research. I am also greatly indebted to Jan Wehr and David Herzog for proposing the model problem which is the primary focus of much of this dissertation. And I am especially thankful to Charles Doering for suggesting an additional example of noise-induced stabilization.

In addition to my research experience, I have enjoyed my many teaching opportunities at Duke University and I am thankful for the teaching mentorship of Professors Lewis Blake and Michael Reed. I am also deeply grateful to my fellow graduate students in the Mathematics Department at Duke University for their collaboration and sense of community.

I have been privileged to be supported by several fellowships during my graduate study and I would like to thank the Duke University Graduate School for providing me with a James B. Duke fellowship, the Statistical and Applied Mathematical

Sciences Institute for a Graduate Fellowship, and the National Science Foundation for support through a Mathematical Biology Research Training Grant (NSF-DMS-09-43760), as well as support through grants NSF-DMS-04-49910 and NSF-DMS-08-54879.

Finally, I would like to thank my family for their unconditional love and support. My wonderful husband, Mark, has been a constant source of encouragement and has stood by me through all the challenges of completing this thesis. I am, in particular, grateful for his willingness to drive me back and forth between East and West campuses (without which I would have been late to many classes) and for his exceptional Matlab skills (without which many of the figures in this dissertation would not exist). I also need to thank my daughter Kayla and my adorable kitten Mystery for keeping me company during the numerous hours of research and writing (and for forcing me to take much-needed breaks).

I am deeply grateful for my parents and my brother, David. As mathematicians themselves, my parents sparked my initial interest in mathematics and have aided me throughout most of my mathematical career, from quizzing me on my times tables in elementary school, to checking my Calculus homework in high school, and to giving me tips on Topology in college. Lastly, I would like to thank God, who has blessed me abundantly with all of the gifts and skills which have enabled me to complete this dissertation.

Introduction

Stabilization by noise is an intriguing and surprising phenomenon. One would usually suspect that the addition of noise to an unstable system would further destabilize it. However, for particular systems, the dynamics are such that the addition of noise creates a stabilizing effect. In the classical example of the inverted pendulum, the addition of noise opens up a small neighborhood of local stability around a deterministically unstable fixed point [2, 17]. The examples presented in this thesis are of a markedly different nature, where instead of considering the stability of isolated fixed points, the global stability of the system is the primary focus. There are many different notions of stability in the literature; hence, the definition of stability used in this thesis will be made precise in Section 1.2. All of the deterministic systems considered in this thesis are defined as solutions to two-dimensional systems of ordinary differential equations (ODEs) and the noise is modeled as classical white noise; the precise mathematical setting will be further described in Section 1.1.

In [31], Scheutzow was the first to define a specific example of a two-dimensional unstable deterministic system and prove that it is stabilized by the addition of additive white noise. However, the dynamics of his deterministic system are quite complicated. In two recent works [10, 16], more straightforward examples of systems which exhibit the noise-induced stabilization phenomenon are presented. In all three papers, the approach to proving stabilization is to build a global Lyapunov function. The global Lyapunov function is constructed by patching together functions which are locally Lyapunov in a collection of regions whose union covers all of the possible routes to infinity. However, the local constructions have mainly the flavor of "guess-and-check," with some information of the presumed overall structure of the transport in phase space.

One of the main contributions of this thesis is a systematic algorithm for the construction of a global Lyapunov function in order to prove noise-induced stabilization. This algorithm was first presented in the joint paper with Jonathan Mattingly and Avanti Athreya [3], and this paper is closely related to much of the content of this thesis. The systematic algorithm is highly advantageous because it greatly reduces the amount of time and tedium that is often required to construct a Lyapunov function. Moreover, we define a particular type of Lyapunov function, which we call a super Lyapunov function, and we show that its existence implies a strong form of convergence to stationarity. What follows is a brief outline of the remainder of this thesis.

In chapter 1, background information on noise-induced stabilization is presented, including precise descriptions of the mathematical setting under consideration and the notion of stability used throughout the thesis. We also give an overview of the theory of Lyapunov functions, which can be used to prove stabilization. In chapter 2, we describe in detail the systematic algorithm for the construction of a global Lyapunov function, while in chapter 3 we prove key consequences of super Lyapunov functions. In particular, we show that the existence of a super Lyapunov function, along with a minorization condition, implies that the system converges to a unique invariant probability measure at a uniform, exponential rate.

In chapter 4, we apply the systematic algorithm to a model problem and show that the global Lyapunov function constructed is in fact super Lyapunov. The model problem has the property that in the deterministic setting, the process blows up in finite time for certain initial conditions. However, the region of instability is isolated, and thus, the addition of noise is able to stabilize the system by allowing the process to flow from the unstable region to stable regions.

In chapter 5, we summarize the results proven in the thesis and explore avenues for future research. In particular, we describe an additional example of a system which exhibits noise-induced stabilization, to which we would like to apply our generalized Lyapunov construction. This system is more complicated than the model problem in that there is not just one region of isolated instability, but rather multiple regions of instability. The addition of noise creates a stabilizing effect in this system by connecting the regions of instability and forming a periodic motion.

In the two appendices, we present additional results for the model problem, namely an analysis in polar coordinates and an analysis of the positivity of the density and invariant probability measure in the case of degenerate noise.

1

Background on Noise-Induced Stabilization

In this chapter, we present background information on the phenomenon of noise-induced stabilization. We first present a detailed description of the mathematical setting under consideration and the precise notion of stability used in this thesis. We then describe the theory of Lyapunov functions, which can be used to prove stabilization.

1.1 Mathematical Setting

The deterministic systems studied in this thesis will be restricted to those that are solutions to two-dimensional systems of autonomous ODEs. Hence, we consider $\mathbb{X}(t) = (X(t), Y(t)) \in \mathbb{R}^2$, which is the solution to the ODE

$$\frac{d\mathbb{X}(t)}{dt} = \varphi(\mathbb{X}(t))$$

where $\varphi \in C^2(\mathbb{R}^2)$ and $\mathbb{X}(0) = \mathbbm{x} = (x, y)$. We are interested in perturbing this deterministic system by the addition of noise. Our model for noise in this context will be the classical white noise. Thus, we consider the perturbed system $\mathbb{X}^\epsilon(t) =$

$(X^\epsilon(t), Y^\epsilon(t)) \in \mathbb{R}^2$, with $\mathbb{X}^\epsilon(0) = \mathbb{x} = (x, y)$, which is the solution to the stochastic differential equation (SDE)

$$\frac{d\mathbb{X}^\epsilon(t)}{dt} = \varphi(\mathbb{X}^\epsilon(t)) + \epsilon \, \frac{dW(t)}{dt}$$

where

$$\epsilon = \begin{pmatrix} \epsilon_x & 0 \\ 0 & \epsilon_y \end{pmatrix} \quad \text{with} \quad \epsilon_x \geqslant 0, \epsilon_y \geqslant 0$$

is a parameter which controls the noise strength and $|\epsilon| = \sqrt{\epsilon_x^2 + \epsilon_y^2}$. $W(t) = (W^x(t), W^y(t))$ is standard two-dimensional Brownian motion; i.e., $W^x(t)$ and $W^y(t)$ are independent, standard one-dimensional Brownian motions. Now $W(t)$ is in fact non-differentiable at every point; hence, $\frac{dW(t)}{dt}$ does not exist in the classical sense. To overcome this technical difficulty, it is possible to make rigorous sense of the integral equation

$$\mathbb{X}^\epsilon(t) = \mathbb{X}^\epsilon(0) + \int_0^t \varphi(\mathbb{X}^\epsilon(s))ds + \epsilon \, W(t)$$

from the theory of stochastic calculus [30]. As shorthand for the integral equation, we will use the notation

$$d\mathbb{X}_t^\epsilon = \varphi(\mathbb{X}_t^\epsilon)dt + \epsilon \, dW_t \tag{1.1}$$

to represent the SDE.

This thesis focuses on the phenomenon where there is a fundamental change in the behavior of the system between the case $|\epsilon| = 0$ and the case $|\epsilon| > 0$: when $|\epsilon| = 0$ the system is unstable and when $|\epsilon| > 0$ the system is stable. This phenomenon is referred to as *noise-induced stabilization*. For most systems which exhibit noise-induced stabilization, stabilization will not occur for all ϵ with $|\epsilon| > 0$, but rather for only certain ϵ in a particular range.

If $\epsilon_x > 0$ and $\epsilon_y > 0$, we say that the noise is *nondegenerate*; otherwise we say that the noise is *degenerate*. It is often easier to analyze the case of nondegenerate noise with noise added equally in both directions, i.e., $\epsilon_x = \epsilon_y$. However, for many systems noise is only needed in one particular direction in order to induce stability, and it is often of interest to identify the minimum amount of perturbation necessary for stabilization. Thus we present the perturbed system in more generality by allowing $\epsilon_x \neq \epsilon_y$. We show in Chapter 4 that the model problem exhibits noise-induced stabilization for any ϵ with $\epsilon_y > 0$.

Under certain properties of the function φ in (1.1), if the system $\mathbb{X}_t^\epsilon$ is stable, then $\mathbb{X}_t^{\tilde{\epsilon}}$ is also stable for any $\tilde{\epsilon} = \ell\epsilon$ with $\ell > 0$. We make this statement precise in the following proposition, which essentially states that the process $\mathbb{X}_t^\epsilon$ is invariant under a suitable rescaling of the noise, time, and space.

Proposition 1.1. *Suppose φ scales homogeneously in the sense that there exists $n \in \mathbb{R}$ so that*

$$\varphi(\ell^{-n}\mathbb{x}) = \ell^{n+2}\varphi(\mathbb{x}) \quad \textit{for all } \ell > 0 \textit{ and } \mathbb{x} \in \mathbb{R}^2 .$$

Then for all $\ell > 0$,

$$\ell^n \mathbb{X}_{\tilde{t}}^{\tilde{\epsilon}} \quad \textit{is equal in distribution to} \quad \mathbb{X}_t^\epsilon$$

where $\tilde{\epsilon} = \ell\epsilon$ and $\tilde{t} = \ell^{-2n-2}t$, provided that the two processes have the same initial condition.

Proof of Proposition 1.1. For simplicity of notation, let $\tilde{\mathbb{X}}_t^\epsilon = \ell^n \mathbb{X}_{\tilde{t}}^{\tilde{\epsilon}}$. Assume $\tilde{\mathbb{X}}_0^\epsilon = \mathbb{X}_0^\epsilon = \mathbb{x}$. Then

$$\begin{aligned}
\tilde{\mathbb{X}}_t^\epsilon &= \mathbb{x} + \ell^n \int_0^{\tilde{t}} \varphi(\mathbb{X}_s^{\tilde{\epsilon}})ds + \epsilon\ \ell^{n+1}\ W_{\tilde{t}} \\
&= \mathbb{x} + \ell^{-n-2} \int_0^t \varphi(\ell^{-n}\tilde{\mathbb{X}}_s^\epsilon)ds + \epsilon\ \ell^{n+1}\ W_{\tilde{t}} \\
&= \mathbb{x} + \int_0^t \varphi(\tilde{\mathbb{X}}_s^\epsilon)ds + \epsilon\ \ell^{n+1}\ W_{\tilde{t}} \ \text{ by the assumption on } \ \varphi\,.
\end{aligned}$$

Now by properties of Brownian motion, $\ell^{n+1}W_{\tilde{t}}$ is equal in distribution to W_t. Therefore, $\tilde{\mathbb{X}}_t^\epsilon$ is equal in distribution to $\mathbb{X}_t^\epsilon$, which completes the proof. □

In the model problem presented in Chapter 4, the φ under consideration is defined by

$$\varphi(\mathbb{x}) = \varphi(x, y) = (x^2 - y^2, 2xy)$$

Hence, $\varphi(\ell^{-n}\mathbb{x}) = \ell^{-2n}\varphi(\mathbb{x})$ for all $n \in \mathbb{R}$, $\ell > 0$, and $\mathbb{x} \in \mathbb{R}^2$. Thus, this φ satisfies the hypothesis of Proposition 1.1 with $n = -\frac{2}{3}$.

We now present in the following section the precise definition of stability used in this thesis.

1.2 Notion of Stability

As mentioned earlier, we are not concerned about the stability of particular fixed points, but rather the global stability of the system. We wish to choose a definition of stability which corresponds to boundedness in the deterministic setting. Thus, the deterministic systems which are unstable will be those which either blow up in finite time or wander off to infinity. A stochastic system will almost never be truly bounded; hence, our notion of stability used in this context is a relaxed version of stochastic boundedness. Any system which is stochastically bounded, the definition of which appears below, will be stable.

Definition 1.1. *A system $\mathbb{X}_t^\epsilon$ is **stochastically bounded** if and only if for all $\mathbb{X}_0^\epsilon = \mathbb{x} \in \mathbb{R}^2$ and for all $\delta > 0$, there exists $M < \infty$ such that*

$$\inf_{t \geqslant 0} \{\mathbb{P}_{\mathbb{x}} (|\mathbb{X}_t^\epsilon| \leqslant M)\} > 1 - \delta\,.$$

Note that the bound M may depend upon the initial condition. Also note that in the deterministic setting when $|\epsilon| = 0$, this definition does indeed reduce to the standard definition of boundedness.

The terminology for stochastic boundedness is unfortunately not consistent in the literature. In the classic text [22], Has'minskiĭ uses the term *bounded in proba-*

bility for the exact same notion. It is also important to note that the process $\mathbb{X}_t^\epsilon$ is stochastically bounded if and only if the corresponding measures $\nu_t(\cdot)$ defined by

$$\nu_t(A) = \mathbb{P}_{\mathbb{x}}\left(\mathbb{X}_t^\epsilon \in A\right),$$

where $A \subset \mathbb{R}^2$ is any Borel (i.e., measurable) set, are *tight* for each $\mathbb{x} \in \mathbb{R}^2$.

The Markov semigroup associated to the process $\mathbb{X}_t^\epsilon$ is defined by

$$(P_t\phi)(\mathbb{x}) = \mathbb{E}_{\mathbb{x}}[\phi(\mathbb{X}_t^\epsilon)] = \int_{\mathbb{R}^2} \phi(\mathbb{x}')P_t(\mathbb{x}, d\mathbb{x}')$$

for any $\phi \in C(\mathbb{R}^2)$, where $P_t(\mathbb{x}, \cdot)$ is the probability transition kernel; i.e,

$$P_t(\mathbb{x}, A) = \mathbb{P}_{\mathbb{x}}(\mathbb{X}_t^\epsilon \in A)$$

for any measurable set $A \subset \mathbb{R}^2$. Moreover, the action of P_t on a probability measure μ is defined by

$$(\mu P_t)(A) = \int_{\mathbb{R}^2} P_t(\mathbb{x}, A)\mu(d\mathbb{x})$$

for any measurable set $A \subset \mathbb{R}^2$. An invariant probability measure μ is any probability measure such that $\mu P_t = \mu$ for all t.

Now the assumption in (1.1) that $\varphi \in C^2(\mathbb{R}^2)$ implies that $\mathbb{X}_t^\epsilon$ is weak Feller continuous, i.e., the Markov semigroup P_t maps bounded continuous functions to continuous functions [30]. Hence, by the Krylov-Bogolyubov theroem [12], if $\mathbb{X}_t^\epsilon$ is stochastically bounded, then there exists at least one invariant probability measure μ for the system.

Now the condition of stochastic boundedness is slightly more stringent than the notion of stability we wish to use in this thesis. Instead of requiring the measures

$$\nu_t(A) = \mathbb{P}_{\mathbb{x}}\left(\mathbb{X}_t^\epsilon \in A\right)$$

to be tight for each $\mathbb{x} \in \mathbb{R}^2$, we will only require the Cesáro means

$$\bar{\nu}_t(A) = \frac{1}{t}\int_0^t \mathbb{P}_{\mathbb{x}}\left(\mathbb{X}^\epsilon_s \in A\right) ds$$

to be tight for each $\mathbb{x} \in \mathbb{R}^2$. This weaker requirement nevertheless still ensures the existence of an invariant probability measure. We remark that this requirement is stronger than the existence of an invariant probability measure, but only slightly; the existence of an invariant probability measure implies that *there exists* $\mathbb{x} \in \mathbb{R}^2$ such that the corresponding Cesáro means $\bar{\nu}_t(\cdot)$ are tight, but it does not have to hold *for all* $\mathbb{x} \in \mathbb{R}^2$ [22].

We now present our precise definition of stability.

Definition 1.2. *A system $\mathbb{X}^\epsilon_t$ is* ***stable*** *if and only if for all $\mathbb{X}^\epsilon_0 = \mathbb{x} \in \mathbb{R}^2$ and for all $\delta > 0$, there exists $M < \infty$ such that*

$$\inf_{t \geqslant 0}\left\{\frac{1}{t}\int_0^t \mathbb{P}_{\mathbb{x}}\left(|\mathbb{X}^\epsilon_s| \leqslant M\right) ds\right\} > 1 - \delta\,.$$

As in the definition of stochastic boundedness, the bound M may depend upon the initial condition. Also note that in the deterministic setting when $|\epsilon| = 0$, this definition again reduces to the standard definition of boundedness. Hence, the deterministic systems which are unstable are those which either blow up in finite time or wander off to infinity, as desired.

It can be difficult to verify that a system is stable directly from the definition given above. Hence, in the next section we present an overview of the theory of Lyapunov functions, which provides sufficient criteria for guaranteeing the stability of a system.

1.3 Lyapunov Functions

In deterministic settings, Lyapunov functions can be used to prove the stability of solutions to ODEs. In stochastic settings, Lyapunov functions serve an analogous role. In this section, we describe Lyapunov functions for two-dimensional systems since planar systems are the primary focus of this thesis; however, all of the definitions and results for Lyapunov functions presented in this section hold in $\mathbb{R}^d$ for any $d \geqslant 1$.

The well-known theorem below (see [22]) provides sufficient criteria to guarantee that the system $\mathbb{X}_t^\epsilon$ is stable.

Theorem 1.1. *Suppose there exists a C^2 Lyapunov function $V : \mathbb{R}^2 \to (0, \infty)$ satisfying*

1. *$V(\mathbb{x}) \to \infty$ as $|\mathbb{x}| \to \infty$*
2. *$(\mathcal{L}V)(\mathbb{x}) \to -\infty$ as $|\mathbb{x}| \to \infty$, where $\mathcal{L}$ is the generator corresponding to the system $\mathbb{X}_t^\epsilon$, i.e.*

$$\mathcal{L} = \varphi \cdot \nabla + \frac{\epsilon_x^2}{2}\partial_{xx} + \frac{\epsilon_y^2}{2}\partial_{yy} \, .$$

Then $\mathbb{X}_t^\epsilon$ is stable.

Before proving the above theorem, we first state the following lemma, which guarantees that the system does not blow up in finite time under weaker conditions than those listed in Theorem 1.1.

Lemma 1.1. *Suppose there exists a C^2 function $V : \mathbb{R}^2 \to (0, \infty)$ satisfying*

1. *$V(\mathbb{x}) \to \infty$ as $|\mathbb{x}| \to \infty$*
2. *There exist constants $m, b > 0$ such that for all $\mathbb{x} \in \mathbb{R}^2$*

$$(\mathcal{L}V)(\mathbb{x}) \leqslant mV(\mathbb{x}) + b \, .$$

Then

$$\mathbb{P}_{\mathbb{x}}\left(\lim_{n\to\infty}\tau_n<\infty\right)=0$$

where

$$\tau_n=\inf\{t\geqslant 0: |\mathbb{X}_t^\epsilon|\geqslant n\}\,.$$

Proof of Lemma 1.1. See [22]. □

Note that standard existence and uniqueness theorems for SDEs (e.g., [30]) usually assume that the function φ in (1.1) is globally Lipschitz. However, this assumption is very restrictive and the above lemma gives sufficient criteria for (1.1) to have a unique, continuous solution defined for all time whenever φ is merely locally Lipschitz (which is implied by our assumption that $\varphi \in C^2$).

We now use the above lemma to prove Theorem 1.1.

Proof of Theorem 1.1. Let $\mathbb{x}\in\mathbb{R}^2$ and $\delta>0$ be fixed. Then there exist constants $t_0, M_0>0$ so that

$$\mathbb{P}_{\mathbb{x}}\left(|\mathbb{X}_t^\epsilon|\leqslant M_0\right)>1-\delta\quad\text{for all } t\in[0,t_0]\,.$$

Dynkin's Formula [30] states that

$$\mathbb{E}_{\mathbb{x}}[V(\mathbb{X}_{t\wedge\tau_n}^\epsilon)]=V(\mathbb{x})+\mathbb{E}_{\mathbb{x}}\left[\int_0^{t\wedge\tau_n}(\mathcal{L}V)(\mathbb{X}_s^\epsilon)ds\right] \tag{1.2}$$

where τ_n is defined in Lemma 1.1. Now by assumption,

$$\mathbb{E}_{\mathbb{x}}[V(\mathbb{X}_{t\wedge\tau_n}^\epsilon)]>0\quad\text{for all } t,n\geqslant 0$$

which implies that

$$\mathbb{E}_{\mathbb{x}}\left[\int_0^{t\wedge\tau_n}(\mathcal{L}V)(\mathbb{X}_s^\epsilon)ds\right]>-V(\mathbb{x})\,.$$

Now for any $M > 0$,

$$(\mathcal{L}V)(\mathbb{X}_s^\epsilon) = (\mathcal{L}V)(\mathbb{X}_s^\epsilon) \cdot \mathbf{1}(\mathbb{X}_s^\epsilon \in \bar{B}_M) + (\mathcal{L}V)(\mathbb{X}_s^\epsilon) \cdot \mathbf{1}(\mathbb{X}_s^\epsilon \in \bar{B}_M^c)$$
$$\leqslant c + \sup_{\mathbb{x} \in \bar{B}_M^c} \{(\mathcal{L}V)(\mathbb{x})\} \, \mathbf{1}(\mathbb{X}_s^\epsilon \in \bar{B}_M^c)$$

where

$$c = \sup_{\mathbb{x} \in \mathbb{R}^2} \{(\mathcal{L}V)(\mathbb{x})\} \vee 0 < \infty$$

by the second assumption in the statement of the theorem. Hence,

$$- \sup_{\mathbb{x} \in \bar{B}_M^c} \{(\mathcal{L}V)(\mathbb{x})\} \, \mathbb{E}_{\mathbb{x}} \left[\int_0^{t \wedge \tau_n} \mathbf{1}(\mathbb{X}_s^\epsilon \in \bar{B}_M^c) ds \right] \leqslant V(\mathbb{x}) + c\, t \, .$$

Since the hypotheses of the theorem imply the hypotheses of Lemma 1.1,

$$\lim_{n \to \infty} t \wedge \tau_n = t \text{ for all } \ t \geqslant 0 \, .$$

Thus, taking the limit as $n \to \infty$, changing the order of integration, and then dividing through by t, we obtain that for all $t \geqslant t_0$,

$$- \sup_{\mathbb{x} \in \bar{B}_M^c} \{(\mathcal{L}V)(\mathbb{x})\} \, \frac{1}{t} \int_0^t \mathbb{P}_{\mathbb{x}}(\mathbb{X}_s^\epsilon \in \bar{B}_M^c) ds \leqslant \frac{V(\mathbb{x})}{t_0} + c \, .$$

Now by the second assumption in the statement of the theorem,

$$- \sup_{\mathbb{x} \in \bar{B}_M^c} \{(\mathcal{L}V)(\mathbb{x})\} \to \infty \ \text{ as } \ M \to \infty \, .$$

Hence, $M \geqslant M_0$ can be chosen so that

$$0 < \frac{1}{- \sup_{\mathbb{x} \in \bar{B}_M^c} \{(\mathcal{L}V)(\mathbb{x})\}} \left[\frac{V(\mathbb{x})}{t_0} + c \right] < \delta \, .$$

Note that this choice of M depends upon the initial condition $\mathbb{x}$ and t_0, but is independent of t. Therefore,

$$\frac{1}{t} \int_0^t \mathbb{P}_{\mathbb{x}}(|\mathbb{X}_s^\epsilon| \leqslant M) ds > 1 - \delta \ \text{ for all } \ t \geqslant 0$$

which completes the proof that $\mathbb{X}_t^\epsilon$ is stable. $\square$

Note that under slightly different conditions than those listed in Theorem 1.1, we can show that the system $\mathbb{X}_t^\epsilon$ is not only stable, but actually stochastically bounded.

Corollary 1.1. *Suppose there exists a C^2 function $V : \mathbb{R}^2 \to (0, \infty)$ satisfying*

1. $V(\mathbb{x}) \to \infty$ *as* $|\mathbb{x}| \to \infty$

2. $(\mathcal{L}V)(\mathbb{x}) \leqslant 0$ *for all* $\mathbb{x} \in \mathbb{R}^2$.

Then $\mathbb{X}_t^\epsilon$ is stochastically bounded (and hence stable).

Proof of Corollary 1.1. Let $\mathbb{x} \in \mathbb{R}^2$ and $\delta > 0$ be fixed. By (1.2),

$$\mathbb{E}_{\mathbb{x}}[V(\mathbb{X}_{t\wedge\tau_n}^\epsilon)] = V(\mathbb{x}) + \mathbb{E}_{\mathbb{x}}\left[\int_0^{t\wedge\tau_n} (\mathcal{L}V)(\mathbb{X}_s^\epsilon)ds\right] \leqslant V(\mathbb{x}),$$

where the last inequality results from the second assumption in the statement of the corollary. Since the assumptions of the corollary imply the assumptions of Lemma 1.1,

$$\lim_{n\to\infty} t \wedge \tau_n = t \text{ for all } \ t \geqslant 0\,.$$

Thus, taking the limit as $n \to \infty$, we obtain that $\mathbb{E}_{\mathbb{x}}[V(\mathbb{X}_t^\epsilon)] \leqslant V(\mathbb{x})$ for all $t \geqslant 0$. Now for any $M > 0$,

$$\begin{aligned}\mathbb{E}_{\mathbb{x}}[V(\mathbb{X}_t^\epsilon)] &= \int_{\mathbb{R}^2} V(\mathbb{x}')P_t(\mathbb{x}, d\mathbb{x}') \\ &\geqslant \int_{\bar{B}_M^c} V(\mathbb{x}')P_t(\mathbb{x}, d\mathbb{x}') \\ &\geqslant \left(\inf_{\mathbb{x}'\in\bar{B}_M^c} V(\mathbb{x}')\right) P_t(\mathbb{x}, \bar{B}_M^c)\,.\end{aligned}$$

Thus,

$$\mathbb{P}_{\mathbb{x}}(|X_t^\epsilon| > M) = P_t(\mathbb{x}, \bar{B}_M^c) \leqslant \frac{\mathbb{E}_{\mathbb{x}}[V(\mathbb{X}_t^\epsilon)]}{\inf_{\mathbb{x}'\in\bar{B}_M^c} V(\mathbb{x}')} \leqslant \frac{V(\mathbb{x})}{\inf_{\mathbb{x}'\in\bar{B}_M^c} V(\mathbb{x}')}\,.$$

Now by the first assumption in the statement of the corollary,

$$\inf_{\mathbb{x}' \in \bar{B}_M^c} V(\mathbb{x}') \to \infty \text{ as } M \to \infty\,.$$

Thus, M can be chosen large enough so that

$$\frac{V(\mathbb{x})}{\inf_{\mathbb{x}' \in \bar{B}_M^c} V(\mathbb{x}')} \leqslant \delta\,.$$

Note that this choice of M is independent of t, but depends upon the initial condition $\mathbb{x}$. This completes the proof that $\mathbb{X}_t^\epsilon$ is stochastically bounded. □

Any function satisfying the assumptions in Theorem 1.1 will be referred to simply as a (general) Lyapunov function. Stronger versions of the Lyapunov condition exist, which we define below, and these stronger versions imply specific rates of convergence to a unique invariant probability measure. These rates of convergence are described in Chapter 3.

Definition 1.3. *A C^2 function $V : \mathcal{R} \to (0, \infty)$, where $\mathcal{R} \subset \mathbb{R}^2$, satisfying*

1. *$V(\mathbb{x}) \to \infty$ as $|\mathbb{x}| \to \infty$ with $\mathbb{x} \in \mathcal{R}$*

2. *There exist constants $m, b, \gamma > 0$ such that for all $\mathbb{x} \in \mathcal{R}$,*

$$(\mathcal{L}V)(\mathbb{x}) \leqslant -mV^\gamma(\mathbb{x}) + b$$

is called a ***standard Lyapunov function*** *on $\mathcal{R}$ if $\gamma = 1$ and a* ***super Lyapunov function*** *on $\mathcal{R}$ if $\gamma > 1$. Moreover, if $\mathcal{R} = \mathbb{R}^2$ we say that V is a* ***global*** *Lyapunov function; otherwise we say that V is a* ***local*** *Lyapunov function.*

When $\gamma \in (0, 1)$, V may be referred to as a sub Lyapunov function. In [13], the authors prove properties of rates of convergence to stationarity for sub Lyapunov functions. The key distinction is that with a sub Lyapunov function, it is only possible to show that the rate of convergence to stationarity is subexponential, while with a standard or super Lyapunov function, we show in Chapter 3 that the rate of convergence is exponential. This thesis will be primarily concerned with general, standard, and super Lyapunov functions.

The above discussion illustrates that it is possible to show the stability of a system by constructing a suitable Lyapunov function, rather than proving stability directly from the definition. However, there is no standard methodology for showing the existence of a Lyapunov function, and constructing a Lyapunov function is usually quite ad hoc and often follows from "guess-and-check." In Chapter 2, we present a systematic algorithm for constructing a Lyapunov function for planar systems exhibiting noise-induced stabilization. In Chapter 4, we illustrate this algorithm by applying it to a model problem. We show that the Lyapunov function constructed is in fact a super Lyapunov function.

2

Systematic Algorithm for Lyapunov Construction

As mentioned in Chapter 1, there is no standard methodology for showing the existence of a Lyapunov function, and constructing a Lyapunov function can often be quite ad hoc and tedious. However, in this chapter, we present a systematic algorithm for constructing a Lyapunov function for planar systems exhibiting noise-induced stabilization. The general approach is to build local Lyapunov functions as solutions to associated partial differential equations, which are defined in regions delineated by different asymptotic behaviors of the system.

First, a "priming region" in the plane is identified where there is an obvious choice of a Lyapunov function; it is often characterized as a subset of phase space in which the deterministic flow is directed toward the origin. Next, by contrast, a "diffusive region" in the plane is identified in which the deterministic dynamics exhibit instability and for which noise is essential to the stabilization. The local Lyapunov function in the priming region is then "propagated" to the diffusive region, through a series of intermediate regions of the plane, until all possible routes to infinity are covered. Thus, a sequence of local Lyapunov functions is obtained, which is then mollified in order to produce a global Lyapunov function.

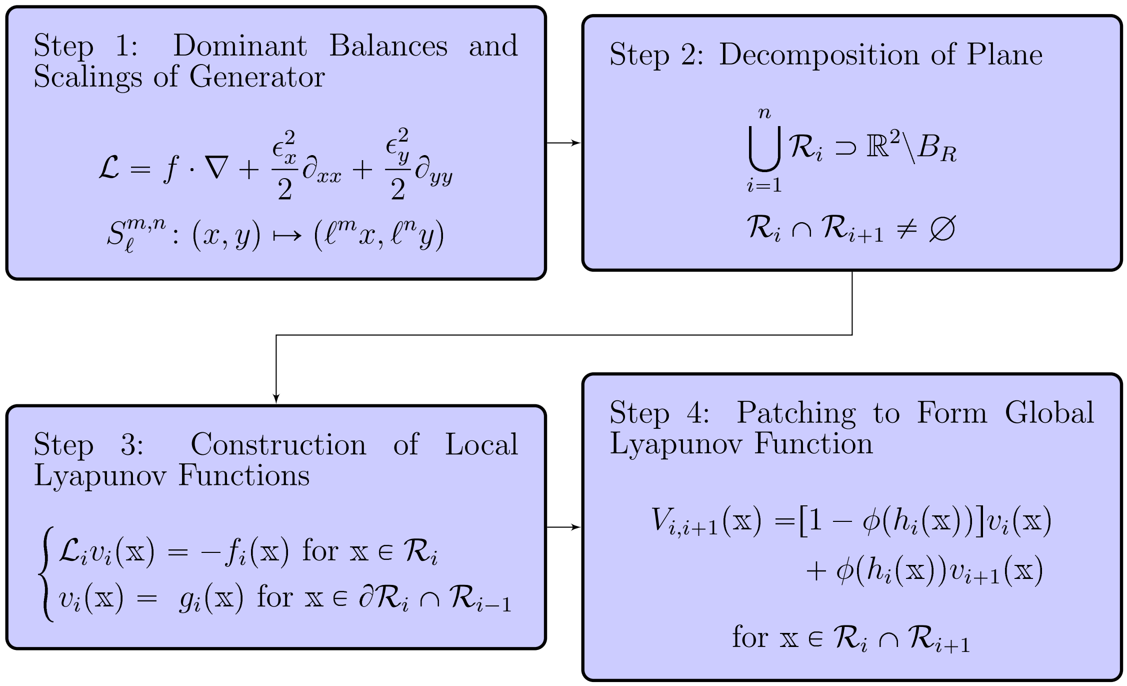

FIGURE 2.1: A flow chart of the algorithm for the construction of a global Lyapunov function.

Below, we outline in more detail each of the steps in the generalized algorithm to construct a Lyapunov function for the process $\mathbb{X}_t^\epsilon$ defined by (1.1). Figure 2.1 gives a summary of the algorithm. This algorithm is more clearly illustrated in Chapter 4, where we apply it to a model problem to show that the model system exhibits noise-induced stabilization. The algorithm presented below is for the construction of a general Lyapunov function. However, we show in Chapter 4 that the Lyapunov function constructed for the model problem is in fact super Lyapunov.

We emphasize that this algorithm may not work for all planar systems and that it may have to be slightly adapted for each specific problem. Nevertheless, the algorithm offers a systematic approach, and its underlying ideas, especially the use of scaling, may be applied to a wide variety of situations. The scaling techniques are the main tool utilized in the algorithm for constructing the Lyapunov function and verifying that it does indeed satisfy the Lyapunov property, and they are one of the most significant contributions of this thesis.

Step 1: Dominant Balances and Scalings of Generator

The first step of the algorithm is to analyze the various dominant balances and scalings of the generator. This analysis will be crucial for decomposing the plane in the next step into the regions in which the local Lyapunov functions will be constructed. While the concept of a dominant balance is clear heuristically, we will introduce some notation to make the concept a bit more precise. We consider a scaling of the form

$$S_\ell^{m,n} \colon (x, y) \mapsto (\ell^m x, \ell^n y) \tag{2.1}$$

for some choice of $m, n \in \mathbb{R}$. The dominant part of the operator $\mathcal{L}$ under the scaling $S_\ell^{m,n}$ consists of those terms in

$$(\mathcal{L} \circ S_\ell^{m,n})(x, y) = \ell^{-m}\varphi_1(\ell^m x, \ell^n y)\partial_x + \ell^{-n}\varphi_2(\ell^m x, \ell^n y)\partial_y + \ell^{-2m}\frac{\epsilon_x^2}{2}\partial_{xx} + \ell^{-2n}\frac{\epsilon_y^2}{2}\partial_{yy}$$

which have the largest exponent of ℓ, where the φ in (1.1) satisfies

$$\varphi(x, y) = (\varphi_1(x, y), \varphi_2(x, y)) \, .$$

For different choices of m and n, there will be different dominant operators. We define the curve $\mathcal{C}^{m,n}(x_0, y_0)$ by

$$\mathcal{C}^{m,n}(x_0, y_0) = \{(x, y) \in \mathbb{R}^2 : (x, y) = S_\ell^{m,n}(x_0, y_0) \text{ for some } \ell \geqslant 1\} \, . \tag{2.2}$$

Heuristically, the dominant part of $\mathcal{L}$ under the scaling $S_\ell^{m,n}$ will be the dominant operator in a region of the plane in which there exist routes to infinity along curves of the form $\mathcal{C}^{m,n}(x_0, y_0)$ for $x_0, y_0 \neq 0$. The curve $\mathcal{C}^{m,n}(x_0, y_0)$ is invariant under the scaling $S_\ell^{m,n}$ in the sense that for any $\ell \geqslant 1$,

$$(\mathcal{C}^{m,n} \circ S_\ell^{m,n})(x_0, y_0) = \mathcal{C}^{m,n}(\ell^m x_0, \ell^n y_0) \subset \mathcal{C}^{m,n}(x_0, y_0) \, .$$

We say that an operator M scales homogeneously under a scaling $S_\ell^{m,n}$ if

$$(M \circ S_\ell^{m,n})(x,y) = M(\ell^m x, \ell^n y) = \ell^\delta M(x,y) \quad \text{for some} \quad \delta \in \mathbb{R}\,.$$

Step 2: Decomposition of Plane

The next step of the algorithm is to decompose the plane into regions $\mathcal{R}_i$, for $i \in \{1, ..., n\}$, in which the local Lyapunov functions will be constructed in the third step. Each of the regions is chosen to be unbounded, and thus contains possible routes to infinity. The first region, $\mathcal{R}_1$, is called the "priming region" and is chosen to be a subset of the plane in which the deterministic dynamics flow inward. This ensures that a very natural Lyapunov function, such as the norm to some power, will satisfy the Lyapunov property in $\mathcal{R}_1$. The last region, $\mathcal{R}_n$, is chosen next and is called the "diffusive region." This region corresponds to a particular dominant balance of the generator $\mathcal{L}$ in which one or more of the diffusion terms is dominant. The diffusive region is a region of the plane in which there is instability in the deterministic setting. Hence, the construction of a local Lyapunov function in $\mathcal{R}_n$ is critical to the proof of noise-induced stabilization.

The intermediate regions $\mathcal{R}_i$, for $i \in \{2, ..., n-1\}$, are each chosen to correspond to different dominant balances of the generator. The precise number of intermediate regions depends upon the number of different dominant balances, and the number must be chosen so that the entire plane is covered, except perhaps for a bounded set about the origin. Thus, the regions satisfy

$$\bigcup_{i=1}^{n} \mathcal{R}_i \supset \mathbb{R}^2 \backslash B_R$$

for some $R \geqslant 0$.

We let $\mathcal{L}_i$ denote the dominant operator in region $\mathcal{R}_i$. In addition, we let $S_\ell^{(i)}$ denote the scaling under which $\mathcal{L}_i$ is dominant. Each i corresponds to a specific choice of m and n. We also let

$$\mathcal{C}^{(i)}(x_0, y_0) = \{(x, y) \in \mathbb{R}^2 : (x, y) = S_\ell^{(i)}(x_0, y_0) \text{ for some } \ell \geqslant 1\}.$$

The precise boundaries of the regions $\mathcal{R}_i$ are chosen to scale homogeneously under the scalings $S_\ell^{(i)}$, respectively. Therefore, the boundaries are chosen to be of the form $\mathcal{C}^{(i)}(x_0, y_0)$. In order to facilitate the mollification of the local Lyapunov functions, the regions are chosen so that the intersection between adjoining regions is both nonempty and unbounded. Thus, the regions satisfy

$$\mathcal{R}_i \cap \mathcal{R}_{i+1} \neq \varnothing$$

for $i \in \{1, ..., n-1\}$.

We note that the regions do not necessarily have to be connected subsets of the plane, which will be evident in the choice of regions in Chapter 4 for the model problem. In addition, there may be multiple "priming" and "diffusive" regions, in which case the algorithm can easily be modified by constructing intermediate regions between each of the priming and diffusive regions.

Step 3: Construction of Local Lyapunov Functions

The third step of the algorithm is to construct a local Lyapunov function in each of the regions $\mathcal{R}_i$. In the priming region $\mathcal{R}_1$, a natural Lyapunov function exists. Often this natural Lyapunov function, v_1, is the norm to some power, i.e., $|\mathbb{x}|^\delta$ for some $\delta > 0$. Verifying that this function has the Lyapunov property in $\mathcal{R}_1$ is usually straightforward and explicit. The priming Lyapunov function, v_1, is then propagated to the other regions through a series of Poisson equations. More specifically, for $i \in \{2, ..., n\}$, v_i is defined as the solution to a PDE of the form

$$\begin{cases} (\mathcal{L}_i v_i)(\mathbb{x}) & = -f_i(\mathbb{x}) \text{ for } \mathbb{x} \in \mathcal{R}_i \\ v_i(\mathbb{x}) & = \; g_i(\mathbb{x}) \text{ for } \mathbb{x} \in \partial\mathcal{R}_i \cap \mathcal{R}_{i-1} \end{cases}$$

where $\mathcal{L}_i$ is the dominant operator in $\mathcal{R}_i$. The boundary data, g_i, for the Poisson equation is given by the dominant behavior of v_{i-1} on the boundary between $\mathcal{R}_i$ and $\mathcal{R}_{i-1}$. The right-hand side, f_i, of the Poisson equation is chosen to be a positive definite function with unbounded growth. Moreover, f_i and g_i are chosen to scale homogeneously under the scaling which is dominant in $\mathcal{R}_i$.

This procedure is iterated to construct local Lyapunov functions as solutions to associated Poisson equations in each of the intermediate regions. Furthermore, a local Lyapunov function is constructed in the diffusive region by solving a Poisson equation as well, again with the boundary data determined by the dominant behavior of the local Lyapunov function in the adjoining intermediate region. An advantage of this approach, therefore, is its consistency: the same procedure is used to construct local Lyapunov functions in all but the priming region (where the local Lyapunov function is straightforward to deduce).

As the sequence of Poisson equations is solved, boundaries without boundary data may be encountered. While a priori this could be an issue, we will see that in our model problem, in all but the diffusive region, the deterministic dynamics are dominant. Hence the associated Poisson equations are governed by first-order operators requiring only one boundary condition. This is consistent with the idea of the priming Lyapunov function being propagated through a sequence of intermediate transport regions. Again, a priori this could lead to an incompatibility between two different boundaries of a given region, particularly if the relevant operator in a region is only first-order and cannot accept generic initial data on two boundaries. However, in our model problem and all of the other problems we have explored, sequences of compatible transport regions are separated from each other by diffusive regions. Since the associated differential operator in the diffusive region is second-order, the associated Poisson equation produces a smooth solution even with all of its boundary data specified.

In order to prove that v_i satisfies the Lyapunov condition in $\mathcal{R}_i$, it is necessary to show that

$$\begin{aligned}(\mathcal{L}v_i)(\mathbb{x}) &= (\mathcal{L}_i v_i)(\mathbb{x}) + ([\mathcal{L} - \mathcal{L}_i]v_i)(\mathbb{x}) \\ &= -f_i(\mathbb{x}) + ([\mathcal{L} - \mathcal{L}_i]v_i)(\mathbb{x}) \to -\infty \quad \text{as} \quad |\mathbb{x}| \to \infty \quad \text{with} \quad \mathbb{x} \in \mathcal{R}_i\,.\end{aligned}$$

This can be show through the scaling relations since $\mathcal{L}_i$ is the dominant operator in $\mathcal{R}_i$, and $\mathcal{L} - \mathcal{L}_i$ is subdominant.

Step 4: Patching to Form Global Lyapunov Function

The last step of the algorithm is to patch the local Lyapunov functions v_i defined in the regions $\mathcal{R}_i$, respectively, together in order to form one smooth Lyapunov function on the entire plane. Thus, the function $\tilde{V}$ is defined by

$$\tilde{V}(\mathbb{x}) = \begin{cases} v_i(\mathbb{x}) & \text{for } \mathbb{x} \in \mathcal{R}_i \cap \left(\bigcup_{j \neq i} \mathcal{R}_j\right)^c \text{ and } i \in \{1, ..., n\} \\ V_{i,i+1}(\mathbb{x}) & \text{for } \mathbb{x} \in \mathcal{R}_i \cap \mathcal{R}_{i+1} \text{ and } i \in \{1, ..., n-1\} \\ 0 & \text{otherwise.} \end{cases}$$

The functions $V_{i,i+1}$ defined in the overlap regions are specified to be smooth, convex combinations of the two local Lyapunov functions in the regions of intersection. More specifically,

$$V_{i,i+1}(\mathbb{x}) = [1 - \phi(h_i(\mathbb{x}))]v_i(\mathbb{x}) + \phi(h_i(\mathbb{x}))v_{i+1}(\mathbb{x})$$

where $\phi(t)$ is a smooth, increasing function which varies from 0 when $t = 0$ to 1 when $t = 1$, and is suitably normalized to integrate to one on the entire real line. In particular, we take

$$\phi(t) = \frac{1}{m}\int_{-\infty}^{t} \psi(s)ds \quad \text{with} \quad m = \int_{-\infty}^{\infty} \psi(s)ds$$

where

$$\psi(t) = \begin{cases} \exp\left(\frac{-1}{1-(2t-1)^2}\right) & \text{for } 0 < t < 1 \\ 0 & \text{otherwise.} \end{cases}$$

Moreover, h_i is defined to be a smooth increasing function that varies from 0 on the boundary of $\mathcal{R}_i$ to 1 on the boundary of $\mathcal{R}_{i+1}$. Thus, $V_{i,i+1} = v_i$ on the boundary of $\mathcal{R}_i$ and $V_{i,i+1} = v_{i+1}$ on the boundary of $\mathcal{R}_{i+1}$. It is not immediately obvious that $V_{i,i+1}$ will satisfy the Lyapunov condition in its region of definition, despite the fact that both v_i and v_{i+1} do. However, the scaling relations can be used to prove that $V_{i,i+1}$ is a Lyapunov function. The global Lyapunov function $V(\mathbb{x}) \in C^2(\mathbb{R}^2)$ is then chosen to satisfy

$$V(x,y) = \begin{cases} \tilde{V}(\mathbb{x}) & \text{for } \mathbb{x} \in B_R^c \\ \text{arbitrary positive and smooth} & \text{for } \mathbb{x} \in B_R \end{cases}$$

for some R so that $\bigcup_{i=1}^{n} \mathcal{R}_i \supset B_R^c$.

As mentioned at the beginning of the chapter, this algorithm is more clearly illustrated when applied to a concrete example. Thus, in Chapter 4, we apply the algorithm to a model problem in order to construct a global Lyapunov function to show noise-induced stabilization. Before doing so, we analyze in Chapter 3 properties of super Lyapunov functions since the Lyapunov function constructed in Chapter 4 for the model problem is in fact super Lyapunov.

3

Consequences of Super Lyapunov Structure

In this chapter, we show that the existence of a super Lyapunov function, as defined in Definition 1.3, implies a stronger rate of convergence to the invariant probability measure than the existence of a standard Lyapunov function. It is important to note that although all of the results in this chapter are presented for $\mathbb{R}^2$ in order to be consistent with the mathematical setting described in Section 1.1, all of the results actually hold in $\mathbb{R}^d$ for any $d \geqslant 1$. In order to quantify the rates of convergence, we first introduce a family of weighted total variation metrics, ρ_β.

3.1 Weighted Total Variation Metrics

For $\beta \geqslant 0$ and probability measures μ_1 and μ_2, we define

$$\rho_\beta(\mu_1, \mu_2) = \sup_{\|\phi\|_\beta \leqslant 1} \int_{\mathbb{R}^2} \phi(\mathbb{x})(\mu_1 - \mu_2)(d\mathbb{x})$$

where

$$\|\phi\|_\beta = \sup_{\mathbb{x}} \frac{|\phi(\mathbb{x})|}{1 + \beta V(\mathbb{x})}$$

and V is a Lyapunov function corresponding to the process $\mathbb{X}_t^\epsilon$. Note that $\rho_0(\mu_1, \mu_2) = \|\mu_1 - \mu_2\|_{TV}$, where $\|\cdot\|_{TV}$ is simply the standard total variation norm. It is clear from the definition of the $\|\cdot\|_\beta$ norms that they are monotonically decreasing in β. Hence, if $\beta_1 < \beta_2$,

$$\{\phi : \|\phi\|_{\beta_1} \leqslant 1\} \subset \{\phi : \|\phi\|_{\beta_2} \leqslant 1\}.$$

This in turn implies that the ρ_β metrics are monotonically increasing since the supremum is taken over larger sets as β increases; i.e., if $\beta_1 < \beta_2$,

$$\rho_{\beta_1}(\mu_1, \mu_2) \leqslant \rho_{\beta_2}(\mu_1, \mu_2).$$

Thus the total variation distance, which corresponds to $\beta = 0$, is the smallest of all the ρ_β metrics.

Now if we define ψ as

$$\psi(\mathbb{x}) = \frac{\phi(\mathbb{x})}{1 + \beta V(\mathbb{x})}$$

then $\|\phi\|_\beta = \|\psi\|_0$. Hence

$$\rho_\beta(\mu_1, \mu_2) = \sup_{\|\psi\|_0 \leqslant 1} \int_{\mathbb{R}^2} (1 + \beta V(\mathbb{x}))\psi(\mathbb{x})(\mu_1 - \mu_2)(d\mathbb{x})$$

which illustrates that the ρ_β metric is just a weighted total variation distance.

3.2 Convergence to Stationarity

It is well known that in the case of nondegenerate noise with $\epsilon_x > 0$ and $\epsilon_y > 0$, the existence of a standard Lyapunov function implies that there exist positive constants C and η so that

$$\|\mu_1 P_t - \mu_2 P_t\|_{TV} \leqslant Ce^{-\eta t}\rho_1(\mu_1, \mu_2) \tag{3.1}$$

for all $t \geqslant 0$ and any probability measures μ_1 and μ_2. If μ_1 is taken to be the initial probability measure, i.e., $\mu_1 = \delta_{\mathbb{x}}$, and μ_2 is taken to be the invariant probability measure, μ, then (3.1) quantifies the rate of convergence to stationarity. More specifically, (3.1) translates to

$$\|P_t(\mathbb{x},\ \cdot\ ) - \mu\|_{TV} \leqslant Ce^{-\eta t}\rho_1(\delta_{\mathbb{x}}, \mu) = C(1 + V(\mathbb{x}))e^{-\eta t}\,.$$

Note that in the above inequality, the upper bound on the right-hand side depends upon the initial condition $\mathbb{x}$.

We show that the existence of a super Lyapunov function implies a stronger rate of convergence to the invariant probability measure, which is encapsulated in the following theorem.

Theorem 3.1. *Suppose there exists a super Lyapunov function, V, for the process $\mathbb{X}_t^\epsilon$, with $\epsilon_x > 0$ and $\epsilon_y > 0$ and associated Markov semigroup P_t. Then there exist positive constants C and η so that*

$$\|\mu_1 P_t - \mu_2 P_t\|_{TV} \leqslant Ce^{-\eta t}\|\mu_1 - \mu_2\|_{TV} \tag{3.2}$$

for all $t \geqslant 0$ and any probability measures μ_1 and μ_2.

In this case, if we consider $\mu_1 = \delta_{\mathbb{x}}$ and $\mu_2 = \mu$, where μ is the invariant probability measure, then (3.2) translates to the statement

$$\|P_t(\mathbb{x},\ \cdot\ ) - \mu\|_{TV} \leqslant Ce^{-\eta t}\|\delta_{\mathbb{x}} - \mu\|_{TV} = Ce^{-\eta t}\,.$$

Note that the upper bound on the right-hand side is now independent of the initial condition $\mathbb{x}$. Thus, the existence of a super Lyapunov function implies a rate of convergence to the invariant probability measure that is not only exponential, but also uniform in initial condition.

As already mentioned in Section 1.3, the existence of an invariant probability measure μ follows quickly from the existence of a general Lyapunov function. The fact that there is only one invariant probability measure is immediate from (3.1) or (3.2). Before proving Theorem 3.1, we first prove a bound on the action of the semigroup, which results from the existence of a super Lyapunov function. This bound is essential to the proof of Theorem 3.1. We then state a minorization condition in Section 3.4, which is also used in the proof of Theorem 3.1. We show that this minorization condition easily holds in the case of nondegenerate noise, but we also give sufficient criteria for it to hold in the case of degenerate noise. Hence, the assumptions on Theorem 3.1 may be extended to allow for degenerate noise, as long as the minorization condition holds.

3.3 Bound on Action of Semigroup

We begin with a lemma which translates the bound on the generator (given in the definition of a super Lyapunov function) to a bound on the action of the semigroup. Despite its simplicity, it is nonetheless a key component to the more sophisticated consequences of the existence of a super Lyapunov function (as opposed to merely a standard Lyapunov function).

Lemma 3.1. *Suppose that* $V : \mathbb{R}^2 \to (0, \infty)$ *is a super Lyapunov function for the process* $\mathbb{X}_t^\epsilon$*, i.e.,*

$$(\mathcal{L}V)(\mathbb{x}) < -mV^\gamma(\mathbb{x}) + b \quad \textit{for all} \quad \mathbb{x} \in \mathbb{R}^2$$

where $\mathcal{L}$ *is the generator corresponding to* $\mathbb{X}_t^\epsilon$*,* $m, b > 0$ *and* $\gamma > 1$*. Then for every* $t > 0$*, there exists a positive constant* K_t*, such that* K_t *is a continuous, monotone decreasing function of* t *on* $(0, \infty)$ *with*

$$\lim_{t \to \infty} K_t = \left(\frac{2b}{m} \right)^{\frac{1}{\gamma}} .$$

Moreover, if P_t *is the Markov semigroup associated to* $\mathbb{X}_t^\epsilon$*, then*

$$(P_tV)(\mathbb{x}) \leqslant K_t \quad \textit{for all} \quad \mathbb{x} \in \mathbb{R}^2 \ \textit{and} \ t \geqslant 0 \, .$$

Proof of Lemma 3.1. Let $V_t = V(\mathbb{X}_t^\epsilon)$. By Dynkin's formula (see [30]),

$$
\begin{aligned}
(P_tV)(\mathbb{x}) = \mathbb{E}_{\mathbb{x}}[V_t] &= V(\mathbb{x}) + \mathbb{E}_{\mathbb{x}}\left[\int_0^t \mathcal{L}V_s ds\right] \\
&\leqslant V(\mathbb{x}) - m\int_0^t \mathbb{E}_{\mathbb{x}}[V_s^\gamma]ds + bt \\
&\leqslant V(\mathbb{x}) - m\int_0^t (\mathbb{E}_{\mathbb{x}}[V_s])^\gamma ds + bt \quad \text{by convexity.}
\end{aligned}
$$

For simplicity of notation, let $\phi_{\mathbb{x}}(t) = (P_tV)(\mathbb{x}) = \mathbb{E}_{\mathbb{x}}[V_t]$. Then $\phi_{\mathbb{x}}(t)$ satisfies the following differential inequality:

$$
\phi'_{\mathbb{x}}(t) \leqslant -m[\phi_{\mathbb{x}}(t)]^\gamma + b \leqslant -\frac{m}{2}[\phi_{\mathbb{x}}(t)]^\gamma \quad \text{if} \quad \phi_{\mathbb{x}}(t) \geqslant R
$$

where

$$
R = \left(\frac{2b}{m}\right)^{\frac{1}{\gamma}}.
$$

Since $\phi'_{\mathbb{x}}(t) < 0$ if $\phi_{\mathbb{x}}(t) \geqslant R$, this implies that once $\phi_{\mathbb{x}}(t) \leqslant R$, $\phi_{\mathbb{x}}(t)$ remains less than or equal to R for all times afterward. Thus, $\phi_{\mathbb{x}}(t) \leqslant R$ for all $t \geqslant \tau$, where τ is the stopping time defined by

$$
\tau = \inf\{t > 0 : \phi_{\mathbb{x}}(t) \leqslant R\}.
$$

Now suppose $\psi_{\mathbb{x}}(t)$ satisfies the following differential equation:

$$
\begin{cases}
\psi'_{\mathbb{x}}(t) = -\dfrac{m}{2}[\psi_{\mathbb{x}}(t)]^\gamma \quad \text{for all } t \in [0,\tau] \\
\psi_{\mathbb{x}}(0) = \phi_{\mathbb{x}}(0) = V(\mathbb{x}).
\end{cases}
$$

Then by the comparison proposition in Section 3.6, $\phi_{\mathbb{x}}(t) \leqslant \psi_{\mathbb{x}}(t)$ for all $t \in [0,\tau]$. Now the differential equation for $\psi_{\mathbb{x}}(t)$ can be solved explicitly to obtain that for all $t \in [0,\tau]$:

$$
\psi_{\mathbb{x}}(t) = \left(\frac{m(\gamma-1)t}{2} + V(\mathbb{x})^{-(\gamma-1)}\right)^{-\frac{1}{\gamma-1}} \leqslant \left(\frac{m(\gamma-1)t}{2}\right)^{-\frac{1}{\gamma-1}}.
$$

Defining the constants K_t as follows

$$K_t = \max\left\{\left(\frac{2b}{m}\right)^{\frac{1}{\gamma}}, \left(\frac{m(\gamma-1)t}{2}\right)^{-\frac{1}{\gamma-1}}\right\}$$

we conclude that $\phi_{\mathbb{x}}(t) \leqslant K_t$ for all $t > 0$, which completes the proof. □

3.4 Minorization Condition

We now state a minorization condition, which in addition to the above bound on the action of the Markov semigroup, is essential to the proof of Theorem 3.1. We show in Section 3.4.1 that this condition is easily satisfied in the case of nondegenerate noise, i.e., $\epsilon_x > 0$ and $\epsilon_y > 0$. In Section 3.4.2, we state and prove sufficient criteria for the minorization condition to hold in the case of degenerate noise. The minorization condition is as follows: we seek a probability measure ν and positive constants α, R, and T so that

$$\inf_{|\mathbb{x}| \leqslant R} P_T(\mathbb{x}, \,\cdot\,) \geqslant \alpha\nu(\,\cdot\,) \tag{3.3}$$

and $R > 2K_T$, where K_T is the constant from Lemma 3.1. This condition is a localized version of the classical Doeblin condition and central to the theory of Harris chains [19, 21, 29]. While the Lyapunov condition ensures the existence of an invariant probability measure and guarantees sufficiently rapid returns to the "center of phase space" to produce geometric mixing, the minorization condition ensures the existence of probabilistic mixing.

3.4.1 Minorization with Nondegenerate Noise

In the case of nondegenerate noise with $\epsilon_x > 0$ and $\epsilon_y > 0$, the system is uniformly elliptic, and the minorization condition is relatively straightforward. Since the diffusion associated with $\mathbb{X}_t^\epsilon$ has a constant, positive definite diffusion matrix, classical results guarantee the existence of a density function $p_t(\mathbb{x}, \mathbb{x}') : (0, \infty) \times \mathbb{R}^2 \times \mathbb{R}^2 \to (0, \infty)$ such that

$$P_t(\mathbb{x}, A) = \int_A p_t(\mathbb{x}, \mathbb{x}') d\mathbb{x}' \tag{3.4}$$

for all measurable sets $A \subset \mathbb{R}^2$. Moreover, $p_t(\mathbb{x}, \mathbb{x}')$ is jointly continuous in t and $(\mathbb{x}, \mathbb{x}')$.

Now since for each $t > 0$, $p_t(\mathbb{x}, \mathbb{x}')$ is an everywhere positive, continuous function of $(\mathbb{x}, \mathbb{x}')$, it is elementary that for any $R > 0$, there exists a positive constant α', depending upon both R and t, so that

$$\inf_{|\mathbb{x}|, |\mathbb{x}'| \leqslant R} p_t(\mathbb{x}, \mathbb{x}') \geqslant \alpha' \,.$$

Then for any measurable set $A \subset \mathbb{R}^2$

$$P_t(\mathbb{x}, A) = \int_A p_t(\mathbb{x}, \mathbb{x}') d\mathbb{x}' \geqslant \alpha' \; \lambda(A \cap B_R(0)) = \alpha' \; \lambda(B_R(0)) \; \nu(A)$$

where λ is Lebesgue measure and ν is the probability measure defined by

$$\nu(A) = \frac{\lambda(A \cap B_R(0))}{\lambda(B_R(0))} \,.$$

Thus the minorization condition (3.3) holds for any $T > 0$ and $R > 2K_T$ with ν defined as above and

$$\alpha = \alpha' \; \lambda(B_R(0)) \,.$$

3.4.2 Minorization with Degenerate Noise

When $|\epsilon| > 0$, but either $\epsilon_x = 0$ or $\epsilon_y = 0$, the situation is more delicate since the system is no longer elliptic. However, if the generator of $\mathbb{X}_t^\epsilon$ is hypoelliptic, then there still exists a continuous density function $p_t(\mathbb{x}, \mathbb{x}')$ so that (3.4) holds. Hypoellipticity can be verified using Hormander's classic condition [5]. The main difference between this setting and the nondegenerate noise setting is that it is no longer necessarily true that $p_t(\mathbb{x}, \mathbb{x}')$ is positive for all $t > 0$ and $\mathbb{x}, \mathbb{x}' \in \mathbb{R}^2$.

We now state and prove a lemma which gives sufficient criteria for the minorization condition to hold in the case of degenerate noise. The criteria essentially consist of continuity and a relaxed positivity assumption for the density function.

Lemma 3.2. *Let $p_t(\mathbb{x}, \mathbb{x}')$ be the density function corresponding to the process $\mathbb{X}_t^\epsilon$, with $|\epsilon| > 0$, where $p_t \colon \mathbb{R}^2 \times \mathbb{R}^2 \to [0, \infty)$ is a continuous function for each $t > 0$. In addition, we assume that there exists $\mathbb{x}^* \in \mathbb{R}^2$ such that for all $R > 0$, there exists $T_R > 0$ such that*

$$\inf_{\mathbb{x} \in \bar{B}_R} p_t(\mathbb{x}, \mathbb{x}^*) > 0$$

for all $t \geqslant T_R$. Under this assumption, for all $R > 0$ and $t \geqslant T_R$, there exist $\alpha \in (0, 1)$ and a probability measure ν such that

$$\inf_{\mathbb{x} \in \bar{B}_R} P_t(\mathbb{x}, \,\cdot\,) \geqslant \alpha\ \nu(\,\cdot\,)\,.$$

Proof of Lemma 3.2. Let $R > 0$ and $t \geqslant T_R$ be fixed. By the assumption, for all $\mathbb{x} \in \bar{B}_R$, $p_t(\mathbb{x}, \mathbb{x}^*) > 0$. By continuity of p_t, for all $\mathbb{x} \in \bar{B}_R$ there exists $\delta_{\mathbb{x}} > 0$ such that for all $\hat{\mathbb{x}} \in B_{\delta_{\mathbb{x}}}(\mathbb{x})$ and for all $\tilde{\mathbb{x}} \in B_{\delta_{\mathbb{x}}}(\mathbb{x}^*)$, $p_t(\hat{\mathbb{x}}, \tilde{\mathbb{x}}) > 0$. Now

$$\bigcup_{\mathbb{x} \in \bar{B}_R} [B_{\delta_{\mathbb{x}}}(\mathbb{x}) \times B_{\delta_{\mathbb{x}}}(\mathbb{x}^*)]$$

is an open cover of the set $\{(\mathbb{x}, \mathbb{x}^*) : \mathbb{x} \in \bar{B}_R\}$. Since this set is compact, there must exist a finite subcover. Hence, there exist $\mathbb{x}_1, ..., \mathbb{x}_n$ such that

$$\bigcup_{i=1}^{n} \left[B_{\delta_{\mathbb{x}_i}}(\mathbb{x}_i) \times B_{\delta_{\mathbb{x}_i}}(\mathbb{x}^*)\right] \supset \{(\mathbb{x}, \mathbb{x}^*) : \mathbb{x} \in \bar{B}_R\}\,.$$

Let $\delta = \min\{\delta_{\mathbb{x}_1}, ..., \delta_{\mathbb{x}_n}\}/2$. Then

$$\bigcup_{i=1}^{n} \left[B_{\delta_{\mathbb{x}_i}}(\mathbb{x}_i) \times B_{\delta_{\mathbb{x}_i}}(\mathbb{x}^*)\right] \supset \{(\mathbb{x}, \tilde{\mathbb{x}}) : \mathbb{x} \in \bar{B}_R \text{ and } \tilde{\mathbb{x}} \in \bar{B}_\delta(\mathbb{x}^*)\}\,.$$

Hence for all $\mathbb{x} \in \bar{B}_R$ and for all $\tilde{\mathbb{x}} \in \bar{B}_\delta(\mathbb{x}^*)$, $p_t(\mathbb{x}, \tilde{\mathbb{x}}) > 0$. Since $\bar{B}_R \times \bar{B}_\delta(\mathbb{x}^*)$ is a compact set, p_t must achieve its minimum on this set. Thus, there exists $\eta > 0$ such that $p_t(\mathbb{x}, \tilde{\mathbb{x}}) \geqslant \eta$ for all $\mathbb{x} \in \bar{B}_R$ and for all $\tilde{\mathbb{x}} \in \bar{B}_\delta(\mathbb{x}^*)$. Now

$$\begin{aligned}\inf_{\mathbb{x} \in \bar{B}_R} P_t(\mathbb{x}, A) = \inf_{\mathbb{x} \in \bar{B}_R} \int_A p_t(\mathbb{x}, \tilde{\mathbb{x}}) d\tilde{\mathbb{x}} &\geqslant \inf_{\mathbb{x} \in \bar{B}_R} \int_{A \cap \bar{B}_\delta(\mathbb{x}^*)} p_t(\mathbb{x}, \tilde{\mathbb{x}}) d\tilde{\mathbb{x}} \\ &\geqslant \lambda\left[A \cap \bar{B}_\delta(\mathbb{x}^*)\right] \inf_{\mathbb{x} \in \bar{B}_R, \tilde{\mathbb{x}} \in \bar{B}_\delta(\mathbb{x}^*)} p_t(\mathbb{x}, \tilde{\mathbb{x}}) \\ &\geqslant \lambda\left[A \cap \bar{B}_\delta(\mathbb{x}^*)\right] \eta = \left(\lambda\left[\bar{B}_\delta(\mathbb{x}^*)\right] \eta\right) \cdot \left(\frac{\lambda\left[A \cap \bar{B}_\delta(\mathbb{x}^*)\right]}{\lambda\left[\bar{B}_\delta(\mathbb{x}^*)\right]}\right)\end{aligned}$$

where $A \subset \mathbb{R}^2$ is any measurable set and λ is Lebesgue measure on $\mathbb{R}^2$. Choose

$$\alpha = \min\left\{\frac{1}{2}, \lambda\left[\bar{B}_\delta(\mathbb{x}^*)\right] \eta\right\} \quad \text{and} \quad \nu(A) = \frac{\lambda\left[A \cap \bar{B}_\delta(\mathbb{x}^*)\right]}{\lambda\left[\bar{B}_\delta(\mathbb{x}^*)\right]}.$$

Then $\alpha \in (0, 1)$, ν is a valid probability measure, and

$$\inf_{\mathbb{x} \in \bar{B}_R} P_t(\mathbb{x}, \cdot\,) \geqslant \alpha\, \nu(\,\cdot\,).$$

□

Since the assumption in the statement of the above lemma trivially holds in the case of nondegenerate noise, as demonstrated in Section 3.4.1, we may use the lemma more generally to guarantee the minorization condition holds in either the case of nondegenerate noise or the case of degenerate noise with the additional assumption.

In Appendix B, we prove through the use of control theory that the model problem defined in Chapter 4 satisfies the assumption of the above lemma for any ϵ with $\epsilon_y > 0$. Thus, the model problem satisfies the minorization condition with degenerate noise in the x-direction.

3.5 Proof of Convergence Theorem

We now turn to proving Theorem 3.1, which relies on the two lemmas above and a result from [19]. Thus, we prove that the existence of a super Lyapunov function, plus a minorization condition, implies that there is a uniform, exponential rate of convergence to the invariant probability measure.

Proof of Theorem 3.1. Choose $R > 2\left(\frac{2b}{m}\right)^{\frac{1}{\gamma}}$. Then choose $T \geqslant T_R$ such that $R > 2K_T$, where K_T is the constant from Lemma 3.1 and T_R is the constant from Lemma 3.2. Note that it is possible to choose such a T since K_t is a monotone decreasing function of t and R was chose to be greater than

$$2\lim_{t\to\infty} K_t\,.$$

Then by Lemma 3.2, there exist $\alpha \in (0,1)$ and a probability measure ν such that

$$\inf_{\mathbb{x}\in\bar{B}_R} P_T(\mathbb{x},\ \cdot\ ) \geqslant \alpha\ \nu(\ \cdot\ )\,.$$

By Theorem 1.3 in [19], whose assumptions hold with the choice of R and T above, there exist constants $\bar{\alpha} \in (0,1)$ and $\bar{\beta} > 0$ such that

$$\rho_{\bar{\beta}}(\mu_1 P_T, \mu_2 P_T) \leqslant \bar{\alpha}\rho_{\bar{\beta}}(\mu_1, \mu_2) \tag{3.5}$$

for any probability measures μ_1 and μ_2. Now $\|\phi\|_{\bar{\beta}} \leqslant 1$ implies $|\phi(\mathbb{x})| \leqslant 1 + \bar{\beta}V(\mathbb{x})$. This in turn implies that

$$(P_T\phi)(\mathbb{x}) \leqslant 1 + \bar{\beta}(P_T V)(\mathbb{x}) \leqslant 1 + \bar{\beta}K_T$$

by Lemma 3.1. Thus $\|\phi\|_{\bar{\beta}} \leqslant 1$ implies that

$$\left\|\frac{(P_T\phi)(\mathbb{x})}{1+\bar{\beta}K_T}\right\|_0 \leqslant 1\,.$$

Therefore we have that

$$\begin{aligned}\rho_{\bar{\beta}}(\mu_1 P_T, \mu_2 P_T) &= (1+\bar{\beta}K_T) \sup_{\|\phi\|_{\bar{\beta}} \leqslant 1} \int \frac{(P_T\phi)(\mathbb{x})}{1+\bar{\beta}K_T}(\mu_1 - \mu_2)(d\mathbb{x}) \\ &\leqslant (1+\bar{\beta}K_T) \sup_{\|\psi\|_0 \leqslant 1} \int \psi(\mathbb{x})(\mu_1-\mu_2)(d\mathbb{x}) \\ &= (1+\bar{\beta}K_T)\rho_0(\mu_1,\mu_2)\,. \end{aligned} \tag{3.6}$$

For $n \in \mathbb{N}$, we will upper bound $\rho_{\bar{\beta}}(\mu_1 P_{nT}, \mu_2 P_{nT})$ by applying the inequality in (3.5) $n-1$ times and then using the inequality in (3.6) for the last step. Hence we have that

$$\begin{aligned}\rho_{\bar{\beta}}(\mu_1 P_{nT}, \mu_2 P_{nT}) &\leqslant \bar{\alpha}^{n-1}\rho_{\bar{\beta}}(\mu_1 P_T, \mu_2 P_T) \\ &\leqslant \bar{\alpha}^n \left(\frac{1+\bar{\beta}K_T}{\bar{\alpha}}\right)\rho_0(\mu_1,\mu_2)\,. \end{aligned} \tag{3.7}$$

We would now like to extend the result in (3.7) to hold not just at non-negative integer multiples of T, but for all $t \geqslant 0$. Note that any $t \geqslant 0$ may be written in the form $t = nT + T'$ where $n = \lfloor \frac{t}{T} \rfloor$ and $T' \in [0, T)$. Thus for all $t \geqslant 0$,

$$\begin{aligned}\rho_{\bar{\beta}}(\mu_1 P_t, \mu_2 P_t) &= \rho_{\bar{\beta}}(\mu_1 P_{T'} P_{nT}, \mu_2 P_{T'} P_{nT}) \\ &\leqslant \bar{\alpha}^n \left(\frac{1+\bar{\beta}K_T}{\bar{\alpha}}\right)\rho_0(\mu_1 P_{T'}, \mu_2 P_{T'}) \text{ by (3.7)} \\ &\leqslant \bar{\alpha}^n \left(\frac{1+\bar{\beta}K_T}{\bar{\alpha}}\right)\rho_0(\mu_1,\mu_2) \\ &\leqslant \bar{\alpha}^{\frac{t}{T}-1} \left(\frac{1+\bar{\beta}K_T}{\bar{\alpha}}\right)\rho_0(\mu_1,\mu_2) \\ &= \left(\frac{1+\bar{\beta}K_T}{\bar{\alpha}^2}\right) e^{\frac{\ln(\bar{\alpha})}{T}t}\rho_0(\mu_1,\mu_2)\,. \end{aligned}$$

Since $\|\mu_1 P_t - \mu_2 P_t\|_{TV} = \rho_0(\mu_1 P_t, \mu_2 P_t) \leqslant \rho_{\bar{\beta}}(\mu_1 P_t, \mu_2 P_t)$, we have that for all $t \geqslant 0$,

$$\|\mu_1 P_t - \mu_2 P_t\|_{TV} \leqslant Ce^{-\eta t}\|\mu_1 - \mu_2\|_{TV}$$

where

$$C = \left(\frac{1 + \bar{\beta} K_T}{\bar{\alpha}^2}\right) \quad \text{and} \quad \eta = -\frac{\ln(\bar{\alpha})}{T}.$$

□

3.6 Comparison Proposition

In this section we state and prove a proposition comparing the solutions to two differential equations. This proposition is used in the proof of Lemma 3.1.

Proposition 3.1. *Suppose $f \in C(\mathbb{R})$ is a non-increasing function and that $\phi(t)$ and $\psi(t)$ are C^1 functions on $\mathbb{R}$ satisfying the following differential inequalities:*

$$\begin{cases} \phi'(t) & \leqslant f(\phi(t)) \\ \psi'(t) & = f(\psi(t)) \\ \phi(0) & = \psi(0). \end{cases}$$

Then $\phi(t) \leqslant \psi(t)$ for all $t \geqslant 0$.

Proof of Proposition 3.1. For all $0 \leqslant r \leqslant t$, we have that

$$\phi(t) \leqslant \phi(r) + \int_r^t f(\phi(s))ds \quad \text{and} \quad \psi(t) = \psi(r) + \int_r^t f(\psi(s))ds$$

which implies

$$\psi(t) - \phi(t) \geqslant (\psi(r) - \phi(r)) + \int_r^t (f(\psi(s)) - f(\phi(s)))ds.$$

Let $T_1 = \inf\{t > 0 : \psi(t) - \phi(t) < 0\}$. Suppose for contradiction that $T_1 < \infty$. Then by continuity, $\psi(T_1) - \phi(T_1) = 0$ and there exists $T_2 \in (T_1, \infty)$ such that for all $t \in (T_1, T_2)$, $\psi(t) - \phi(t) < 0$. Then for all $t \in (T_1, T_2)$,

$$\begin{aligned} \psi(t) - \phi(t) &\geqslant (\psi(T_1) - \phi(T_1)) + \int_{T_1}^t (f(\psi(s)) - f(\phi(s)))ds \\ &= \int_{T_1}^t (f(\psi(s)) - f(\phi(s)))ds. \end{aligned} \tag{3.8}$$

Now since f is non-increasing, $\psi(t) < \phi(t)$ implies that $f(\psi(t)) \geqslant f(\phi(t))$. Hence this result combined with (3.8) implies that for all $t \in (T_1, T_2)$, $\psi(t) - \phi(t) \geqslant 0$. This is a contradiction. Hence T_1 must be infinite and $\phi(t) \leqslant \psi(t)$ for all $t \geqslant 0$. $\square$

4

A Model Problem

In this chapter we apply our systematic algorithm for the construction of a Lyapunov function to a model problem in order to prove that it exhibits the noise-induced stabilization phenomenon. We first present the deterministic system and demonstrate that it is indeed unstable by showing that it blows up in finite time for certain initial conditions. Furthermore, by careful analysis of the deterministic dynamics, we are able to gain much intuition into why the addition of noise creates a stabilizing effect. The key idea is that the instability is isolated on the positive x-axis. For all initial conditions off the positive x-axis, the deterministic dynamics flow in towards the origin. Hence, the addition of noise allows the process to escape from a narrow region about the axis and flow in towards the origin, avoiding finite time blow-up.

The model problem considered here is essentially the same example presented in [8] and [16]. In these works, the authors construct a Lyapunov function in order to prove noise-induced stabilization. However, as mentioned earlier, the construction is quite complicated and ad hoc. In this chapter, we present a simpler and more systematic construction. Moreover, the Lyapunov function constructed is actually a super Lyapunov function, allowing us to prove a stronger form of convergence, i.e., that the rate of convergence to the invariant probability measure is not only exponential, but also uniformly bounded in initial condition.

4.1 Deterministic System

We consider the deterministic system $\mathbb{X}_t = (X_t, Y_t) \in \mathbb{R}^2$ which is the solution to the ODE

$$\begin{aligned} dX_t &= (X_t^2 - Y_t^2)dt \\ dY_t &= 2X_tY_tdt\,. \end{aligned} \tag{4.1}$$

This two-dimensional system of ODEs originates from the complex-valued ODE

$$dZ_t = Z_t^2 dt$$

where $Z_t = X_t + iY_t \in \mathbb{C}$. In the complex form, it is straightforward to find the explicit solution

$$Z_t = \frac{Z_0}{1 - Z_0 t}\,.$$

Separating the real and imaginary components leads us to the following explicit solution to the two-dimensional system of ODEs:

$$\begin{aligned} X_t &= \frac{X_0 - (X_0^2 + Y_0^2)t}{(1 - X_0t)^2 + (Y_0t)^2} \\ Y_t &= \frac{Y_0}{(1 - X_0t)^2 + (Y_0t)^2}\,. \end{aligned}$$

From the explicit solution, we observe that X_t is an even function of Y_0 and Y_t is an odd function of Y_0. Thus, the deterministic process is symmetric about the x-axis. In addition, there is symmetry about the y-axis in the sense that (X_{-t}, Y_{-t}) starting from $(-X_0, Y_0)$ is equal to $(-X_t, Y_t)$ starting from (X_0, Y_0).

When $Y_0 \neq 0$, we obtain the following relationship between X_t and Y_t:

$$X_t^2 + \left(Y_t - \frac{X_0^2 + Y_0^2}{2Y_0}\right)^2 = \left(\frac{X_0^2 + Y_0^2}{2Y_0}\right)^2\,.$$

Thus, the deterministic system follows circular trajectories with

$$\text{radius} = \frac{X_0^2 + Y_0^2}{2|Y_0|} \quad \text{and} \quad \text{center} = \left(0, \frac{X_0^2 + Y_0^2}{2Y_0}\right).$$

Furthermore, when $Y_0 \neq 0$, it is clear from the explicit equations that

$$\lim_{t\to\infty} X_t = 0 \quad \text{and} \quad \lim_{t\to\infty} Y_t = 0\,.$$

Now when $Y_0 = 0$, $Y_t \equiv 0$ and the equation for X_t simplifies to

$$X_t = \frac{X_0}{1 - X_0 t}\,.$$

When $X_0 = 0$, X_t is also identically zero; thus, the origin is an equilibrium point. When $X_0 < 0$,

$$\lim_{t\to\infty} X_t = 0\,.$$

Now the instability results from the case $X_0 > 0$; more specifically

$$\lim_{t\to\frac{1}{X_0}} X_t = +\infty\,.$$

Thus the system blows up in finite time, where $t = \frac{1}{X_0}$ is the blow-up time. The trajectories of the system are shown in Figure 4.1.

While the origin is not reached in finite time by any trajectory with initial condition $(X_0, Y_0) \neq (0,0)$, for all choices of initial conditions not on the positive x-axis, the following proposition shows that the time to return to a fixed ball of radius R about the origin is uniformly bounded.

Proposition 4.1. *For all initial conditions (X_0, Y_0) not on the positive x-axis, the time T at which the process (X_t, Y_t) defined by* (4.1) *returns to a fixed ball of radius R about the origin is less than or equal to $\frac{2}{R}$.*

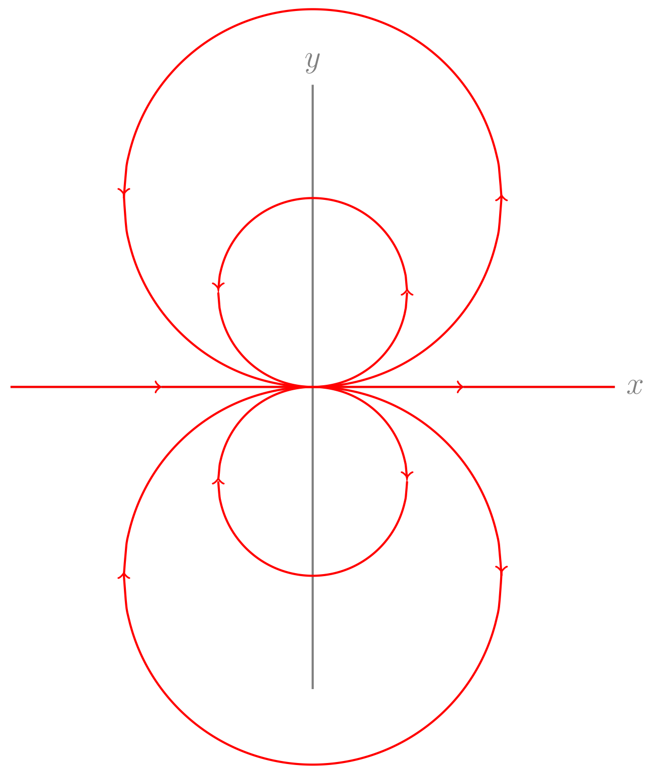

FIGURE 4.1: A number of representative orbits of the deterministic dynamics governed by (4.1).

Proof of Proposition 4.1. We assume $X_0^2+Y_0^2 > R^2$ since otherwise $T = 0$. If $Y_0 = 0$, then $X_T = R$ implies $T = \frac{1}{R} - \frac{1}{|X_0|} \leqslant \frac{1}{R}$. We now consider the case where $Y_0 \neq 0$. Now for any X_0, it follows by the symmetry of the process about the y-axis that the time to reach the ball of radius R is less than or equal to twice the time to reach the ball of radius R starting from the point $(0, y_*)$ with $y_* = \frac{X_0^2+Y_0^2}{Y_0}$, since this point lies on the deterministic circular trajectory going through (X_0, Y_0). Note that $X_0^2 + Y_0^2 > R^2$ implies $|y_*| > R$. Starting from $(0, y^*)$, the equations for X_t and Y_t simplify to

$$X_t = \frac{-y_*^2 t}{1 + (y_* t)^2} \quad \text{and} \quad Y_t = \frac{y_*}{1 + (y_* t)^2} \,.$$

Thus, we can solve for the return time by setting

$$X_T^2 + Y_T^2 = \frac{y_*^2}{1 + (y_* T)^2} = R^2$$

which implies

$$T = \sqrt{\frac{1}{R^2} - \frac{1}{y_*^2}} \leqslant \frac{1}{R}\,.$$

Hence, for all initial conditions, $T \leqslant \frac{2}{R}$. $\square$

This proposition highlights the underlying reason why noise is able to induce stabilization in the system: the instability is isolated in a particular region of phase space, and in all other regions of phase space, the deterministic dynamics flow in towards the origin. Furthermore, in Appendix B, we employ this uniform bound to prove a positivity condition on the transition density for the perturbed system.

In Appendix A, the model problem is analyzed in polar coordinates. The polar form helps illuminate many of the key properties of the dynamics since the deterministic trajectories are circular. In particular, we are able to calculate an asymptotic expansion of the invariant probability density in polar coordinates in order to analyze its decay rate. However, to parallel the Lyapunov construction outlined in Chapter 2, we present our main results in Cartesian coordinates. We now progress our analysis to the perturbed system.

4.2 Perturbed System

We consider the perturbed system $\mathbb{X}_t^\epsilon = (X_t^\epsilon, Y_t^\epsilon)$ given by the following SDE:

$$\begin{aligned} dX_t^\epsilon &= ((X_t^\epsilon)^2 - (Y_t^\epsilon)^2)dt + \epsilon_x \, dW_t^x \\ dY_t^\epsilon &= 2X_t^\epsilon Y_t^\epsilon dt + \epsilon_y \, dW_t^y \,. \end{aligned} \tag{4.2}$$

Unlike in the deterministic setting, it is not possible to find an explicit solution for the perturbed system.

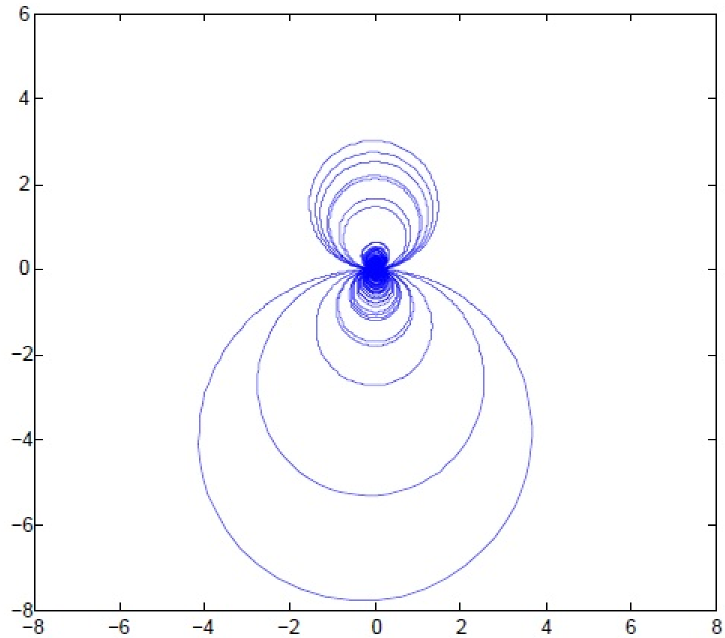

FIGURE 4.2: A simulation of the perturbed system defined by (4.2), with $\epsilon_x, \epsilon_y > 0$ and initial condition on the positive x-axis.

In the above section, we showed that when $\epsilon_x = \epsilon_y = 0$, this system blows up in finite time for certain initial conditions, i.e., those located on the positive x-axis. We also previously described the intuition for why the addition of noise is able to stabilize this system. We simulated the process with $\epsilon_x, \epsilon_y > 0$ and initial condition on the positive x-axis in Matlab using a basic Forward Euler-Maruyama method, and the result is shown in Figure 4.2.

The simulation helps confirm our intuition that the system does indeed exhibit noise-induced stabilization. Rather than flowing towards infinity in finite time along the positive x-axis, the noise kicks the process away from the axis and then it flows back in towards the origin along the deterministic circular trajectories. In the deterministic setting, the process is not able to reach the origin in finite time. However, the noise allows the process to flow through the origin and out along the positive x-axis, where the motion is repeated, forming a quasi-periodic behavior. We see from the figure that the process spends a significant amount of time around the origin, which is expected since the origin is a fixed point of the deterministic system. Also as expected, we observe that the process spends about half its time above the x-axis and about half its time below the x-axis, since as the process flows through the origin, the noise has a one-half probability of kicking the process above the axis and a one-half probability of kicking the process below the axis.

While it is clear from the simulation that the model problem exhibits noise-induced stabilization, in order to rigorously prove that the perturbed system is stable, we construct a Lyapunov function, as described in Section 1.3. The Lyapunov function constructed will have the super Lyapunov property, as defined in Definition 1.3. This is encapsulated in the following theorem.

Theorem 4.1. *There exists a super Lyapunov function V for the process $\mathbb{X}_t^\epsilon$ defined by* (4.2)*; i.e., there exists a C^2 function $V : \mathbb{R}^2 \to (0, \infty)$ such that*

$$V(\mathbb{x}) \to \infty \text{ as } |\mathbb{x}| \to \infty$$

and there exist constants $m, b > 0$ and $\gamma > 1$ such that for all $\mathbb{x} \in \mathbb{R}^2$,

$$(\mathcal{L}V)(\mathbb{x}) \leqslant -mV^\gamma(\mathbb{x}) + b$$

where $\mathcal{L}$ is the generator for $\mathbb{X}_t^\epsilon$ defined by

$$\mathcal{L}(x, y) = (x^2 - y^2)\partial_x + 2xy\partial_y + \sigma_x\partial_{xx} + \sigma_y\partial_{yy}$$

with

$$\sigma_x = \frac{\epsilon_x^2}{2} \text{ and } \sigma_y = \frac{\epsilon_y^2}{2}.$$

We will use the systematic algorithm described in Chapter 2 to construct the Lyapunov function V with the desired properties. The general methodology will be to construct a series of local Lyapunov functions in different regions of the plane, based upon different asymptotic behaviors of the deterministic dynamics. The regions are determined through the calculations in Sections 4.3 and 4.4, while the local Lyapunov functions are constructed in Section 4.5. We will then patch these local Lyapunov functions together to form one smooth global Lyapunov function in Section 4.6. We remark that the notation in the following construction varies slightly from that used in Chapter 2.

4.3 Dominant Balances and Scalings of Generator

We now begin to apply the algorithm laid out in Chapter 2 to construct a Lyapunov function for the model problem. The first step is to consider the various possible dominant balances between the terms in the generator, and to analyze their different scalings in phase space. In order to prove that the addition of noise arrests the blow-up on the positive x-axis sufficiently to produce an invariant probability measure, we need to understand the behavior of the dynamics at infinity. There are many different routes to infinity and we now consider the various possible dominant balances associated with different routes.

The dominant part of the operator $\mathcal{L}$ under the scaling $S_\ell^{m,n}$, defined by (2.1), consists of those terms in

$$(\mathcal{L} \circ S_\ell^{m,n})(x, y) = \ell^m(x^2\partial_x + 2xy\partial_y) - \ell^{2n-m}y^2\partial_x + \ell^{-2m}\sigma_x\partial_{xx} + \ell^{-2n}\sigma_y\partial_{yy}$$

which have the largest exponent of ℓ. We introduce two scaling transformations

$$S_\ell^{(1)} = S_\ell^{1,-\frac{1}{2}} : (x, y) \mapsto (\ell x, \ell^{-\frac{1}{2}}y) \quad \text{and} \quad S_\ell^{(2)} = S_\ell^{1,1} : (x, y) \mapsto (\ell x, \ell y)$$

corresponding to two specific sets of choices for m and n. We define the operator dominant under scaling $S_\ell^{(1)}$ by

$$A = x^2\partial_x + 2xy\partial_y + \sigma_y\partial_{yy} \tag{4.3}$$

and the operator dominant under scaling $S_\ell^{(2)}$ by

$$T = (x^2 - y^2)\partial_x + 2xy\partial_y\,.$$

T is simply the operator corresponding to the deterministic system; hence the notation T indicates that it is a transport operator. We define the curves $\mathcal{C}^{(1)}(x_0, y_0)$ and $\mathcal{C}^{(2)}(x_0, y_0)$ by

$$\mathcal{C}^{(i)}(x_0, y_0) = \{(x, y) \in \mathbb{R}^2 : (x, y) = S_\ell^{(i)}(x_0, y_0) \text{ for some } \ell \geqslant 1\}\,.$$

Thus, $\mathcal{C}^{(1)}(x_0, y_0) = \mathcal{C}^{1,-\frac{1}{2}}(x_0, y_0)$ and $\mathcal{C}^{(2)}(x_0, y_0) = \mathcal{C}^{1,1}(x_0, y_0)$, where $\mathcal{C}^{m,n}(x_0, y_0)$ is defined by (2.2). To make these curves more clear, we observe that

$$\mathcal{C}^{(1)}(x_0, y_0) = \left\{(x, y) \in \mathbb{R}^2 : \operatorname{sgn}(x_0)x > |x_0| \text{ and } y = \operatorname{sgn}(y_0)\sqrt{\frac{x_0}{y_0^2 x}}\right\}$$

$$\mathcal{C}^{(2)}(x_0, y_0) = \left\{(x, y) \in \mathbb{R}^2 : \operatorname{sgn}(x_0)x > |x_0| \text{ and } y = \frac{y_0}{x_0}x\right\}\,.$$

When $m = 1$, there is a transition in the dominant terms in the generator $\mathcal{L}$ as the value of n varies. If $n > -\frac{1}{2}$, then the dominant part is contained in the operator T. If $n \leqslant -\frac{1}{2}$, then the dominant part is contained in the operator A. We note that when $n < -\frac{1}{2}$, the dominant term in $\mathcal{L}$ is just $\sigma_y\partial_{yy}$; but again, this term is contained in the operator A. We also note that the term $\sigma_x\partial_{xx}$ is never dominant, which is one of the underlying reasons why the system exhibits noise-induced stabilization for any $\sigma_y > 0$, regardless of whether $\sigma_x > 0$ or $\sigma_x = 0$. The noise in the y-direction is key to stabilization, while the noise in the x-direction is insignificant.

From the above observations, we see that the operator T will be dominant in a region of the plane in which there do not exist any routes to infinity along paths of the form $\mathcal{C}^{1,n}(x_0, y_0)$ with $n \leqslant -\frac{1}{2}$ and $x_0, y_0 \neq 0$. Likewise, the operator A will be dominant in a region of the plane in which there do not exist any routes to infinity along paths of the form $\mathcal{C}^{1,n}(x_0, y_0)$ with $n > -\frac{1}{2}$ and $x_0, y_0 \neq 0$. These facts will be crucial to our decisions in Section 4.4 for how to decompose the plane into different regions in which we construct the various local Lyapunov functions. But before we do this, we investigate the homogeneous scalings of the two operators A and T.

Observe that the operator A scales homogeneously under the scaling $S_\ell^{(1)}$, while the operator T scales homogeneously under the scaling $S_\ell^{(2)}$. In fact, the dominant operator corresponding to scaling $S_\ell^{m,n}$ will always scale homogeneously under that scaling. We would like the operator T to scale homogeneously under the scaling $S_\ell^{(1)}$ as well; however, this does not hold for all of the terms in T. We remedy this by introducing a non-negative parameter λ and defining the family of operators

$$T_\lambda = (x^2 - \lambda y^2)\partial_x + 2xy\partial_y \,. \tag{4.4}$$

Note that $T = T_1$. We extend the definition of the scaling operators by

$$S_\ell^{(1)} \colon (x, y, \lambda) \mapsto (\ell x, \ell^{-\frac{1}{2}} y, \ell^3 \lambda) \quad \text{and} \quad S_\ell^{(2)} \colon (x, y, \lambda) \mapsto (\ell x, \ell y, \lambda) \,.$$

Now T_λ scales homogeneously under both scaling map $S_\ell^{(1)}$ and $S_\ell^{(2)}$, while A continues to scale homogeneously under $S_\ell^{(1)}$. We will use the family of operators T_λ to help construct a local Lyapunov function in the transport region in Section 4.5.2. This gambit of introducing an extra parameter to produce a homogeneous scaling was also used in a similar way in [11].

We extend the definition of the curves $\mathcal{C}^{(i)}(x_0, y_0)$ by

$$\mathcal{C}^{(i)}(x_0, y_0, \lambda_0) = \{(x, y, \lambda) \in \mathbb{R}^2 \times [0, \infty) : (x, y, \lambda) = S_\ell^{(i)}(x_0, y_0, \lambda_0) \text{ for some } \ell \geqslant 1\} \,.$$

Given a function $\phi : \mathbb{R}^2 \times [0, \infty) \to \mathbb{R}$, we say that ϕ scales homogeneously under the scaling $S_\ell^{(i)}$ if $\phi \circ S_\ell^{(i)} = \ell^\delta \phi$ for some δ. In this case, we say that ϕ scales like ℓ^δ under the i-th scaling. We write this compactly as $\phi \overset{i}{\sim} \ell^\delta$.

Proposition 4.2. *If* $\phi \overset{1}{\sim} \ell^{\delta}$*, then* $\partial_x \phi \overset{1}{\sim} \ell^{\delta-1}$ *and* $\partial_y \phi \overset{1}{\sim} \ell^{\delta+\frac{1}{2}}$*. Similarly, if* $\phi \overset{2}{\sim} \ell^{\delta}$*, then* $\partial_x \phi \overset{2}{\sim} \ell^{\delta-1}$ *and* $\partial_y \phi \overset{2}{\sim} \ell^{\delta-1}$*. In all cases, if one side is infinite, then so is the other.*

Proof of Proposition 4.2. We only show one case; all others follow similarly. If $\phi \overset{1}{\sim} \ell^{\delta}$, then $\phi(\ell x, \ell^{-\frac{1}{2}} y, \ell^3 \lambda) = \ell^{\delta} \phi(x, y, \lambda)$. Differentiating in x, we obtain

$$\ell(\partial_x \phi)(\ell x, \ell^{-\frac{1}{2}} y, \ell^3 \lambda) = \ell^{\delta} (\partial_x \phi)(x, y, \lambda) \,.$$

Dividing through by ℓ, we conclude that $\partial_x \phi \overset{1}{\sim} \ell^{\delta-1}$. □

In the next section, we decompose the plane into regions where the different dominant balances hold. These regions are defined by boundary curves which are well-behaved under one or both of the scalings.

4.4 Decomposition of Plane

Based on the discussion in the previous section, we will divide the plane into three regions $\mathcal{R}_i(\alpha)$, where α is a positive parameter that we specify later. As described in Chapter 2, we call these regions the "priming," "transport," and "diffusive" regions, respectively. The regions are defined so that there is a non-empty overlap between adjacent regions and so that their union covers the entire plane, except a bounded region about the origin. In addition, each of the regions will be symmetric about the x-axis, since the deterministic dynamics are symmetric about the x-axis. This results in the transport region being the union of two disconnected subsets.

Our priming region, $\mathcal{R}_1(\alpha)$, is a subset of the left-half plane where there exists a very natural Lyapunov function. This is because in this region, the deterministic drift is directed toward the origin. On the other hand, the diffusive region, $\mathcal{R}_3(\alpha)$, is a funnel-like region around the positive x-axis where there is finite time blow-up in the deterministic setting. Demonstrating the existence of a local Lyapunov function

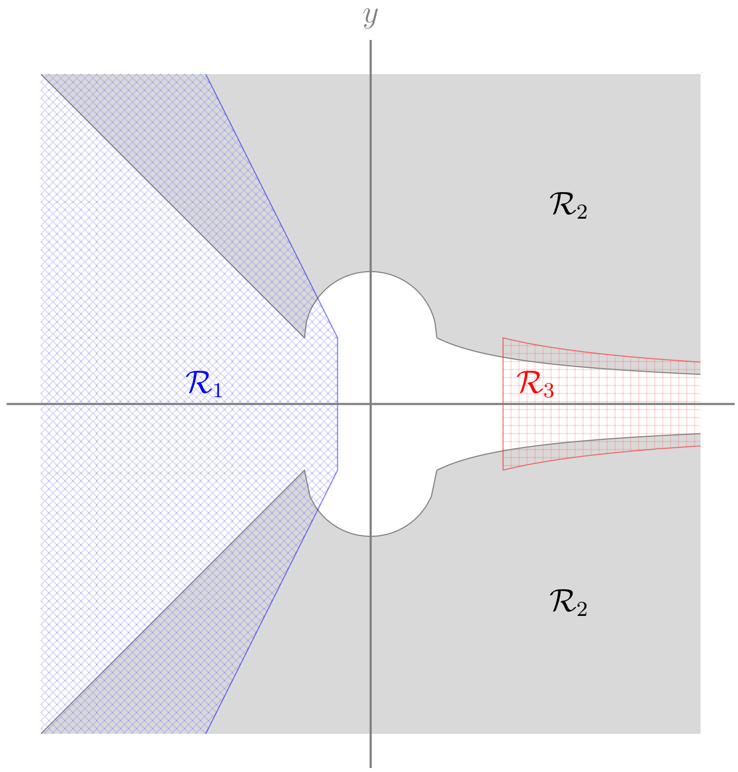

FIGURE 4.3: The different regions in which local Lyapunov functions are constructed: $\mathcal{R}_1$ is the priming region, $\mathcal{R}_2$ is the transport region, and $\mathcal{R}_3$ is the diffusive region.

in the diffusive region is a key piece in proving noise-induced stabilization in the model problem. The transport region $\mathcal{R}_2(\alpha)$ is governed primarily by deterministic transport from the diffusive region to the priming region. We now describe the placement of these various regions, which are indicated pictorially in Figure 4.3.

We will see in Section 4.5.1 that the natural Lyapunov function $(x^2+y^2)^{\frac{\delta}{2}}$ with $\delta > 0$ is super Lyapunov in a subset of the left-half plane bounded by the vertical line $\{(x,y) : x = -x_0, |y| \leqslant y_0\}$ and the two curves $\mathcal{C}^{(2)}(-x_0, \pm y_0)$, for any $x_0, y_0 > 0$. Note that the curves intersect the vertical line at the points $(x_0, \pm y_0)$. For now we choose $(x_0, y_0) = (\frac{\alpha}{2}, 1)$ to define the boundary curves for the priming region $\mathcal{R}_2(\alpha)$, where α will be specified later as mentioned above.

The boundaries for the diffusive region $\mathcal{R}_3(\alpha)$ are chosen to scale nicely under the scaling $S_\ell^{(1)}$, since A will be the dominant operator in this region and A is the operator dominant under the scaling $S_\ell^{(1)}$. Hence, we choose the boundaries of the region to be the vertical line $\{(x,y) : x = 2\alpha, |y| \leqslant 1\}$ and the two curves $\mathcal{C}^{(1)}(2\alpha, \pm 1)$. The curves intersect the vertical line at the points $(2\alpha, \pm 1)$.

The transport region $\mathcal{R}_2(\alpha)$ will connect the priming and diffusive regions. We will define it to be the union of two subregions, $\mathcal{R}_{21}(\alpha)$ and $\mathcal{R}_{22}(\alpha)$, where the first subregion overlaps with $\mathcal{R}_1(\alpha)$ and the second subregion overlaps with $\mathcal{R}_3(\alpha)$. The precise boundaries of the region $\mathcal{R}_{21}(\alpha)$ where it overlaps with $\mathcal{R}_1(\alpha)$ are chosen to scale nicely under the scaling $S_\ell^{(2)}$, since T will be the dominant operator in this region and T is the operator dominant under the scaling $S_\ell^{(2)}$. Hence, we choose the boundaries to be $\mathcal{C}^{(2)}(-\alpha, \pm 1)$. The boundaries of the region $\mathcal{R}_{22}(\alpha)$ where it overlaps with $\mathcal{R}_3(\alpha)$ are chosen to have a similar form to the boundaries of the region $\mathcal{R}_3(\alpha)$. Hence, the boundaries are chosen to be $\mathcal{C}^{(1)}(\alpha, \pm 1)$.

The boundaries between the two subregions $\mathcal{R}_{21}(\alpha)$ and $\mathcal{R}_{22}(\alpha)$ are chosen to be $\mathcal{C}^{(1)}(\alpha, \pm 1)$. This choice ensures that the operator T is the dominant operator in the entire subregion $\mathcal{R}_{21}(\alpha)$. Note that the subregion $\mathcal{R}_{22}(\alpha)$ contains routes to infinity along curves of the form $\mathcal{C}^{1,-\frac{1}{2}}(x_0, y_0)$. Thus, the operator T is not in fact dominant in this entire subregion; nevertheless, it is necessary to have such a region in order to overlap with the diffusive region, and we will show in Section 4.5.2 that it is still possible to construct a local super Lyapunov function in this region using the operator T. The inner boundary of the subregion $\mathcal{R}_{21}(\alpha)$ is chosen to be a circular arc which connects the curves $\mathcal{C}^{(2)}(-\alpha, \pm 1)$ and $\mathcal{C}^{(2)}(\alpha, \pm 1)$, i.e., it is chosen to be

$$\left\{(x,y) : -\alpha \leqslant x \leqslant \alpha \text{ and } x^2 + y^2 = \alpha^2 + 1\right\} .$$

Note that the curves intersect this circular arc at the points $(\pm\alpha, \pm 1)$.

In summary, for each $\alpha > 0$, we have defined three regions

$$\text{Priming Region: } \mathcal{R}_1(\alpha) = \left\{(x, y) : x \leqslant -\frac{\alpha}{2} \text{ and } |y| \leqslant \frac{2}{\alpha}|x|\right\}$$

$$\text{Transport Region: } \mathcal{R}_2(\alpha) = \mathcal{R}_{21}(\alpha) \cup \mathcal{R}_{22}(\alpha)$$

$$\text{Diffusive Region: } \mathcal{R}_3(\alpha) = \left\{(x, y) : x \geqslant 2\alpha \text{ and } |y| \leqslant \sqrt{\frac{2\alpha}{x}}\right\}$$

where

$$\mathcal{R}_{21}(\alpha) = \left\{(x, y) : x^2 + y^2 \geqslant \alpha^2 + 1 \text{ and } |y| \geqslant \frac{|x|}{\alpha}\right\}$$

$$\mathcal{R}_{22}(\alpha) = \left\{(x, y) : x \geqslant \alpha \text{ and } \sqrt{\frac{\alpha}{x}} \leqslant |y| \leqslant \frac{x}{\alpha}\right\}.$$

Notice that $\mathcal{R}_1(\alpha) \cap \mathcal{R}_2(\alpha)$ and $\mathcal{R}_2(\alpha) \cap \mathcal{R}_3(\alpha)$ are nonempty and that $\mathbb{R}^2 \backslash (\mathcal{R}_1(\alpha) \cup \mathcal{R}_2(\alpha) \cup \mathcal{R}_3(\alpha))$ is a bounded set.

4.5 Construction of Local Lyapunov Functions

In this section, we construct local super Lyapunov functions in each of the three regions $\mathcal{R}_i(\alpha)$ defined above. Hence, we construct a C^2 function $v_i : \mathcal{R}_i(\alpha) \to (0, \infty)$ such that

$$v_i(\mathbb{x}) \to \infty \text{ as } |\mathbb{x}| \to \infty \text{ with } \mathbb{x} \in \mathcal{R}_i(\alpha)$$

and there exist constants $m_i > 0$, $R_i \geqslant 0$ and $\gamma_i > 1$ such that for all $\mathbb{x} \in \mathcal{R}_i(\alpha)$ with $|\mathbb{x}| \geqslant R_i$,

$$(\mathcal{L}v_i)(\mathbb{x}) \leqslant -m_i v_i^{\gamma_i}(\mathbb{x})$$

for $i = 1, 2, 3$. Then in Section 4.6, we smoothly patch the three local super Lyapunov functions together to form one global super Lyapunov function on the entire plane.

4.5.1 The Priming Region

When constructing a priming Lyapunov function, it is natural to consider the norm to some power. In the model problem, we expect the norm to some power to be a Lyapunov function in the left-half plane since the drift vector field points at least partially towards the origin, as illustrated in Figure 4.1. This was our rationale for constructing $\mathcal{R}_1(\alpha)$ as specified in the above section. In the proposition below, we state and prove that the norm to some power is in fact a super Lyapunov function in this region.

Proposition 4.3. *For any $\alpha > 0$ and $\delta \in (0,2)$, if $(x,y) \in \mathcal{R}_1(\alpha)$ with $|(x,y)| \geqslant R_1$, then $v_1(x,y) = (x^2+y^2)^{\frac{\delta}{2}}$ satisfies*

$$(\mathcal{L}v_1)(x,y) \leqslant -m_1 \; v_1^{\gamma_1}(x,y)$$

where $m_1 = \frac{\alpha\delta}{2\sqrt{\alpha^2+4}} > 0$, $\gamma_1 = \frac{\delta+1}{\delta} > 1$, and $R_1 = \left[2\left(\sigma_x+\sigma_y\right)\frac{\sqrt{\alpha^2+4}}{\alpha}\right]^{\frac{1}{3}}$.

Proof of Proposition 4.3. We first note that based upon the definition of $\mathcal{R}_1(\alpha)$, for any $(x,y) \in \mathcal{R}_1(\alpha)$,

$$\frac{x}{\sqrt{x^2+y^2}} \leqslant \frac{x}{\sqrt{x^2+\frac{4}{\alpha^2}x^2}} = \frac{-\alpha}{\sqrt{\alpha^2+4}}\,.$$

Applying the generator to v_1 we obtain that

$$\begin{aligned}\mathcal{L}v_1(x,y) =& \delta x(x^2+y^2)^{\frac{\delta}{2}} + \delta(\delta-2)(x^2+y^2)^{\frac{\delta}{2}-2}(\sigma_x x^2+\sigma_y y^2)\\ &+ \delta(\sigma_x+\sigma_y)(x^2+y^2)^{\frac{\delta}{2}-1}\\ \leqslant& -\frac{\alpha\delta}{\sqrt{\alpha^2+4}}(x^2+y^2)^{\frac{\delta+1}{2}} + \delta(\sigma_x+\sigma_y)(x^2+y^2)^{\frac{\delta}{2}-1}\\ =& -\frac{\alpha\delta}{\sqrt{\alpha^2+4}}(x^2+y^2)^{\frac{\delta+1}{2}}\left[1-\frac{\sqrt{\alpha^2+4}}{\alpha}\left(\frac{\sigma_x+\sigma_y}{(x^2+y^2)^{\frac{3}{2}}}\right)\right].\end{aligned}$$

R_1 is chosen so that for all $(x, y) \in \mathcal{R}_1(\alpha)$ with $|(x, y)| \geqslant R_1$, the term in the brackets is greater than or equal to $\frac{1}{2}$. This completes the proof with the values of m_1 and γ_1 as specified in the statement of the proposition. □

Note that neither σ_x nor σ_y must be strictly positive in order for v_1 to be a super Lyapunov function in the priming region. We will see that the only restriction on the noise coefficients occurs in the diffusive region.

We now propagate the priming Lyapunov function v_1 which is defined in $\mathcal{R}_1(\alpha)$ to the other regions by solving a succession of Poisson equations. Throughout most of our construction, δ and α will remain free parameters which we specify later to ensure a number of necessary estimates. We now turn to the transport region $\mathcal{R}_2(\alpha)$.

4.5.2 The Transport Region

For any $\delta > 0$ and $\alpha > 0$, we define v_2 as the solution to the following Poisson equation:

$$\begin{cases} (Tv_2)(x, y) = -\left(\dfrac{x^2 + y^2}{|y|}\right)^{\delta+1} & \text{on } \mathcal{R}_2(\alpha) \\ \\ v_2(x, y) = v_1(x, y) & \text{on } \mathcal{C}^{(2)}(-\alpha, \pm 1)\,. \end{cases} \tag{4.5}$$

Recall that $\mathcal{C}^{(2)}(-\alpha, \pm 1)$ is the boundary of $\mathcal{R}_2(\alpha)$ where it overlaps with the priming region $\mathcal{R}_1(\alpha)$.

The scaling properties of the solution to the above Poisson equation are one of the main tools we use to show that v_2 is a local Lyapunov function. This is because, with one exception, points at infinity in $\mathcal{R}_2(\alpha)$ can be scaled back to points in the same region by the scaling $S_\ell^{(2)}$, which is the scaling transformation under which the associated differential operator T is homogeneous. As we discuss below, the exception is the subregion of $\mathcal{R}_2(\alpha)$ which lies near the boundary of $\mathcal{R}_3(\alpha)$.

Care must be taken when scaling the points in the subregion of $\mathcal{R}_2(\alpha)$ which lie close to the boundary of $\mathcal{R}_3(\alpha)$. The points in this region naturally scale with $S_\ell^{(1)}$ while the operator T which is associated to $\mathcal{R}_2(\alpha)$ scales homogeneously under $S_\ell^{(2)}$. This issue was also addressed in Section 4.3 where we introduced the parameter λ to generate a family of operators T_λ which scale homogeneously with $S_\ell^{(1)}$.

With this in mind, it is natural to introduce the function $v_2(x, y, \lambda)$ which, for a given $\lambda \in (0, 1]$, solves the following family of auxiliary Poisson equations in $\mathcal{R}_2(\alpha, \lambda)$:

$$\begin{cases} (T_\lambda v_2)(x, y, \lambda) = -h(x, y, \lambda) & \text{on } \mathcal{R}_2(\alpha\sqrt{\lambda}) \\ v_2(x, y, \lambda) = f(x, y, \lambda) & \text{on } \mathcal{C}^{(2)}(-\alpha\sqrt{\lambda}, \pm 1) \end{cases} \tag{4.6}$$

where we define

$$h(x, y, \lambda) = \left(\frac{x^2 + \lambda y^2}{|y|}\right)^{\delta+1} \quad \text{and} \quad f(x, y, \lambda) = \lambda^{\frac{\delta+1}{2}}(x^2 + \lambda y^2)^{\frac{\delta}{2}}.$$

For simplicity of notation, we write

$$h(x, y) = h(x, y, 1) \quad \text{and} \quad f(x, y) = f(x, y, 1).$$

Notice that $h \overset{1}{\sim} \ell^{\hat{\delta}+1}$, $f \overset{1}{\sim} \ell^{\hat{\delta}}$, $h \overset{2}{\sim} \ell^{\delta+1}$, and $f \overset{2}{\sim} \ell^{\delta}$ where $\hat{\delta} = \frac{5\delta+3}{2}$. Also observe that $v_2(x, y, 1)$ coincides with the $v_2(x, y)$ defined by (4.5).

We employ the method of characteristics to produce the following unique, explicit solution to (4.6):

$$v_2(x, y, \lambda) = \left(\frac{x^2 + \lambda y^2}{|y|}\right)^{\delta} \left[\frac{x}{|y|} + \alpha\sqrt{\lambda} + \sqrt{\lambda}\left(\frac{1}{\alpha^2 + 1}\right)^{\frac{\delta}{2}}\right]. \tag{4.7}$$

It is clear from this explicit solution that $v_2 \overset{1}{\sim} \ell^{\hat{\delta}}$, $v_2 \overset{2}{\sim} \ell^{\delta}$, and $(x, y, \lambda) \mapsto v_2(x, y, \lambda)$ is continuous on $U \times [0, 1]$ where $U \subset \mathbb{R}^2$ is any closed set not containing any point on the x-axis.

As $\lambda \to 0$, $v_2(x, y, \lambda)$ given in (4.7) converges to $x^{2\delta+1}|y|^{-(\delta+1)}$. This was expected since formally taking $\lambda \to 0$ in (4.5) produces the equation

$$\begin{cases} (x^2\partial_x + 2xy\partial_y)v_2(x, y, 0) = -\dfrac{x^{2(\delta+1)}}{|y|^{\delta+1}} & \text{on } \mathcal{R}_2(0) \\ \\ v_2(x, y, 0) = 0 & \text{on } \mathcal{C}^{(2)}(0, \pm 1)\,. \end{cases} \tag{4.8}$$

Note that the limiting region $\mathcal{R}_2(0) = \mathcal{R}_{21}(0) \cup \mathcal{R}_{22}(0)$, where

$$\mathcal{R}_{21}(0) = \mathcal{C}^{(2)}(0, \pm 1) = \{(x, y) : x = 0, y \geqslant 1\}$$

$$\mathcal{R}_{22}(0) = \{(x, y) : x \geqslant 0, y \neq 0\}\,.$$

There is not a unique solution to the simplified Poisson equation given by (4.8), but it is easily seen that any solution is of the form

$$v_2(x, y, 0) = \frac{x^{2\delta+1}}{|y|^{\delta+1}} + \phi\left(\frac{x^2}{|y|}\right)$$

for any $\phi(t) \in C^1(\mathbb{R})$ with $\phi(0) = 0$ since $x^2|y|^{-1}$ is in the kernel of the operator $(x^2\partial_x + 2xy\partial_y)$ and satisfies the boundary requirement. Thus $x^{2\delta+1}|y|^{-(\delta+1)}$ is the solution corresponding to $\phi(t) \equiv 0$.

In a more complex setting than the model problem where the local Lyapunov function in the transport region cannot be solved explicitly, an analogous simplified Poisson equation corresponding to a limiting value of a parameter may be easy to solve. We will see in Section 4.6 that the most delicate parts of the patching only require precise information about the limit of v_2 as $\lambda \to 0$, not the fully explicit formula. This suggests that the analysis may be feasible even when (4.5) is not solvable.

We now turn to proving the super Lyapunov property for v_2. This is encapsulated in the following proposition.

Proposition 4.4. *For any $\delta > 0$ and $\alpha > 2\sigma_y(\delta+1)(\delta+2)$, there exist constants $m_2 > 0$, $R_2 \geqslant 0$, and $\gamma_2 > 1$ so that if $(x,y) \in \mathcal{R}_2(\alpha)$ with $|(x,y)| \geqslant R_2$, then v_2 satisfies*

$$(\mathcal{L}v_2)(x,y) \leqslant -m_2 \; v_2^{\gamma_2}(x,y)\,.$$

This proposition will be proven in two parts, where we consider the subregions $\mathcal{R}_{21}(\alpha)$ and $\mathcal{R}_{21}(\alpha)$ separately. In Lemmas 4.1 and 4.2, we prove that v_2 is a super Lyapunov function on each of these subregions. The restriction on the size of α in the proposition originates from Lemma 4.2. We start with $\mathcal{R}_{21}(\alpha)$ where the analysis is more straightforward.

Lemma 4.1. *For any $\delta > 0$ and $\alpha > 0$, there exist constants $m_{21} > 0$, $R_{21} \geqslant 0$, and $\gamma_{21} > 1$ so that if $(x,y) \in \mathcal{R}_{21}(\alpha)$ with $|(x,y)| \geqslant R_{21}$, then v_2 satisfies*

$$(\mathcal{L}v_2)(x,y) \leqslant -m_{21} \; v_2^{\gamma_{21}}(x,y)\,.$$

Proof of Lemma 4.1. Applying the generator to v_2, we obtain that

$$\begin{aligned}(\mathcal{L}v_2)(x,y) &= (Tv_2)(x,y) + \sigma_x\partial_{xx}v_2(x,y) + \sigma_y\partial_{yy}v_2(x,y)\\ &= -h(x,y) + \sigma_x\partial_{xx}v_2(x,y) + \sigma_y\partial_{yy}v_2(x,y)\,.\end{aligned}$$

Note that any point in $\mathcal{R}_{21}(\alpha)$ can be connected by a radial line to a point on the circular arc of radius $r = \sqrt{\alpha^2+1}$, which is the arc that forms the inner boundary of this subregion. It follows from this that any $(x,y) \in \mathcal{R}_{21}(\alpha)$ can be written in the form $(x,y) = S_\ell^{(2)}(a,b)$ where $\ell = \frac{|(x,y)|}{r}$ and (a,b) is a point on the circle of radius r centered at the origin. Therefore, we have the following homogeneous scaling properties:

$$\begin{aligned} v_2(x,y) &= \ell^\delta v_2(a,b) & (\partial_{xx}v_2)(x,y) &= \ell^{\delta-2}(\partial_{xx}v_2)(a,b)\\ h(x,y) &= \ell^{\delta+1}h(a,b) & (\partial_{yy}v_2)(x,y) &= \ell^{\delta-2}(\partial_{yy}v_2)(a,b)\,.\end{aligned}$$

These scaling relations lead us to choose $\gamma_{21} = \frac{\delta+1}{\delta}$, which is the ratio of the exponents of ℓ in Tv_2 and v_2. With this choice of γ_{21}, if $(x, y) = S_\ell^{(2)}(a, b)$, we have that

$$\left[\frac{h - \sigma_x|\partial_{xx}v_2| - \sigma_y|\partial_{yy}v_2|}{v_2^{\gamma_{21}}}\right](x, y) = \left[\frac{h - \ell^{-3}\sigma_x|\partial_{xx}v_2| - \ell^{-3}\sigma_y|\partial_{yy}v_2|}{v_2^{\gamma_{21}}}\right](a, b) .$$

We define

$$\rho = \inf_{(a,b)\in\partial B_r\cap\mathcal{R}_{21}} \frac{h(a, b)}{v_2^{\gamma_{21}}(a, b)} \quad \text{and} \quad M = \sup_{(a,b)\in\partial B_r\cap\mathcal{R}_{21}} \left[\frac{\sigma_x|\partial_{xx}v_2| + \sigma_y|\partial_{yy}v_2|}{v_2^{\gamma_{21}}}\right](a, b) .$$

Because h and v_2 are C^2 in (x, y) and strictly positive and $\partial B_r \cap \mathcal{R}_{21}$ is a compact set, we conclude that $\rho > 0$ and $M < \infty$. From this we see that

$$\begin{aligned}
(\mathcal{L}v_2)(x, y) &= -h(x, y) + \sigma_x\partial_{xx}v_2(x, y) + \sigma_y\partial_{yy}v_2(x, y) \\
&= -v_2^{\gamma_{21}}(x, y)\left[\frac{h - \ell^{-3}\sigma_x\partial_{xx}v_2 - \ell^{-3}\sigma_y\partial_{yy}v_2}{v_2^{\gamma_{21}}}\right](a, b) \\
&\leqslant -v_2^{\gamma_{21}}(x, y)(\rho - \ell^{-3}M) .
\end{aligned}$$

We set $m_{21} = \frac{\rho}{2}$ and we choose R_{21} sufficiently large in order to ensure that

$$\rho - \ell^{-3}M = \rho - \frac{Mr^3}{|(x, y)|^3} \geqslant \rho - \frac{Mr^3}{(R_{21})^3} \geqslant m_{21} .$$

Hence, we choose

$$R_{21} = \left(\frac{2Mr^3}{\rho}\right)^{\frac{1}{3}} .$$

□

Note that this proof only relies on the explicit solution for v_2 in order to verify its scaling and smoothness properties. In a more complex problem where an explicit solution to the Poisson equation in the transport region is not available, it is still possible to prove the super Lyapunov property of v_2 as long as its scaling and smoothness properties can be verified.

We now turn to the subregion $\mathcal{R}_{22}(\alpha)$. We prove that v_2 is super Lyapunov on $\mathcal{R}_{22}(\alpha)$ through the following lemma.

Lemma 4.2. *For any $\delta > 0$ and $\alpha > 2\sigma_y(\delta+1)(\delta+2)$, there exist constants $m_{22} > 0$, $R_{22} \geqslant 0$, and $\gamma_{22} > 1$ so that if $(x, y) \in \mathcal{R}_{22}(\alpha)$ with $|(x, y)| \geqslant R_{22}$, then v_2 satisfies*

$$(\mathcal{L}v_2)(x, y) \leqslant -m_{22}\ v_2^{\gamma_{22}}(x, y)\,.$$

Proof of Lemma 4.2. Applying the generator to v_2, we obtain that

$$\begin{aligned}(\mathcal{L}v_2)(x, y) &= (Tv_2)(x, y) + \sigma_x \partial_{xx} v_2(x, y) + \sigma_y \partial_{yy} v_2(x, y)\\ &= -h(x, y) + \sigma_x \partial_{xx} v_2(x, y) + \sigma_y \partial_{yy} v_2(x, y)\,.\end{aligned}$$

We again use the scaling properties to bound the right-hand side in the above equation. Every point $(x, y) \in \mathcal{R}_{22}(\alpha)$ can be mapped back to a point (a, b) on the curve

$$\mathcal{C}^{(2)}(\alpha, \pm 1) = \left\{(a, b) : a \geqslant \alpha, |b| = \frac{a}{\alpha}\right\},$$

which forms the boundary between the subregions $\mathcal{R}_{21}(\alpha)$ and $\mathcal{R}_{22}(\alpha)$, such that $(x, y) = S_\ell^{(1)}(a, b)$, where

$$\ell = \left(\frac{x}{\alpha|y|}\right)^{\frac{2}{3}},\ a = \alpha^{\frac{2}{3}}(xy^2)^{\frac{1}{3}}, \ \text{ and } \ b = \operatorname{sgn}(y)\alpha^{-\frac{1}{3}}(xy^2)^{\frac{1}{3}}\,.$$

Hence we obtain the following scaling relations:

$$\begin{aligned} v_2(x, y, 1) &= \ell^{\hat\delta} v_2(a, b, \ell^{-3}) \qquad & (\partial_{xx} v_2)(x, y, 1) &= \ell^{\hat\delta-2}(\partial_{xx} v_2)(a, b, \ell^{-3})\\ h(x, y, 1) &= \ell^{\hat\delta+1} h(a, b, \ell^{-3}) \qquad & (\partial_{yy} v_2)(x, y, 1) &= \ell^{\hat\delta+1}(\partial_{yy} v_2)(a, b, \ell^{-3})\,.\end{aligned}$$

These scaling relations lead us to choose

$$\gamma_{22} = \frac{\hat\delta + 1}{\hat\delta} = \frac{5\delta + 5}{5\delta + 3}$$

which is the ratio of the exponents of ℓ in Tv_2 and v_2. We now use the scaling transformation $S^{(2)}_{|b|}$ to map (a,b) back to $\left(\alpha, \frac{b}{|b|}\right)$. Then since v_2 and h are even functions of b, we obtain

$$v_2(a,b,\ell^{-3}) = |b|^{\delta} v_2(\alpha,1,\ell^{-3}) \qquad (\partial_{xx}v_2)(a,b,\ell^{-3}) = |b|^{\delta-2}(\partial_{xx}v_2)(\alpha,1,\ell^{-3})$$
$$h(a,b,\ell^{-3}) = |b|^{\delta+1} h(\alpha,1,\ell^{-3}) \qquad (\partial_{yy}v_2)(a,b,\ell^{-3}) = |b|^{\delta-2}(\partial_{yy}v_2)(\alpha,1,\ell^{-3})\,.$$

Combining the two sets of scaling estimates and setting $\bar{\gamma} = \delta(1-\gamma_{22})+1$, we get that

$$\begin{aligned}
&\left[\frac{h - \sigma_x\partial_{xx}v_2 - \sigma_y\partial_{yy}v_2}{v_2^{\gamma_{22}}}\right](x,y,1) \\
&\quad = |b|^{\bar{\gamma}}\left[\frac{h - (|b|\ell)^{-3}\sigma_x\partial_{xx}v_2 - |b|^{-3}\sigma_y\partial_{yy}v_2}{v_2^{\gamma_{22}}}\right](\alpha,1,\ell^{-3}) \\
&\quad = b^{\bar{\gamma}}\left[\frac{h}{v_2^{\gamma_{22}}}\left(1 - (|b|\ell)^{-3}\sigma_x\frac{\partial_{xx}v_2}{h} - |b|^{-3}\sigma_y\frac{\partial_{yy}v_2}{h}\right)\right](\alpha,1,\ell^{-3}) \\
&\quad \geqslant \left[\frac{h}{v_2^{\gamma_{22}}}\left(1 - \ell^{-3}\sigma_x\frac{\partial_{xx}v_2}{h} - \sigma_y\frac{\partial_{yy}v_2}{h}\right)\right](\alpha,1,\ell^{-3})
\end{aligned}$$

since $|b| \geqslant 1$. We define

$$\rho(\lambda^*) = \inf_{\lambda\in[0,\lambda^*]}\frac{h(\alpha,1,\lambda)}{v_2^{\gamma_{22}}(\alpha,1,\lambda)} \quad\text{and}\quad M_1(\lambda^*) = \sup_{\lambda\in[0,\lambda^*]}\frac{|\partial_{xx}v_2(\alpha,1,\lambda)|}{h(\alpha,1,\lambda)}$$
$$\text{and}\quad M_2(\lambda^*) = \sup_{\lambda\in[0,\lambda^*]}\frac{|\partial_{yy}v_2(\alpha,1,\lambda)|}{h(\alpha,1,\lambda)}$$

where $\lambda^* = (\ell^*)^{-3}$ with ℓ^* defined by

$$\ell_* = \inf\{\ell = (x/\alpha y)^{\frac{2}{3}} : (x,y)\in\mathcal{R}_{22}(\alpha) \text{ and } |(x,y)| \geqslant R_{22}\}\,.$$

We choose R_{22} momentarily. From the explicit equations for v_2 and h, it is clear that h, v_2, $\partial_{xx}v_2$, and $\partial_{yy}v_2$ are all continuous functions of λ for $\lambda \geqslant 0$ and are nonvanishing with $x = \alpha$ and $y = 1$. Hence, we can conclude that $\rho(\lambda^*) > 0$ and

$M_1(\lambda^*), M_2(\lambda^*) < \infty$. We set $m_{22} = \frac{\rho(\lambda^*)}{2}$ and choose R_{22} so that

$$1 - \lambda^* \sigma_x M_1(\lambda^*) - \sigma_y M_2(\lambda^*) \geqslant \frac{1}{2}\,.$$

In order to show that such an R_{22} is feasible, we need to understand the behavior of $M_1(\lambda^*)$ and $M_2(\lambda^*)$ as $\lambda^* \to 0$. By direct calculation from the explicit formula for v_2, we see that

$$M_1(0) = \frac{2\delta(2\delta+1)}{\alpha^3} \quad \text{and} \quad M_2(0) = \frac{(\delta+1)(\delta+2)}{\alpha}\,.$$

Since $M_1(\lambda^*)$ and $M_2(\lambda^*)$ are continuous functions of λ^* and since $\lambda^* \to 0$ as $R_{22} \to \infty$, it is possible to choose R_{22} large enough to make

$$1 - \lambda^* \sigma_x M_1(\lambda^*) - \sigma_y M_2(\lambda^*)$$

arbitrarily close to

$$1 - \sigma_y M_2(0) = 1 - \sigma_y \frac{(\delta+1)(\delta+2)}{\alpha}\,.$$

However, this will only be greater than or equal to $\frac{1}{2}$ as long as $\frac{\sigma_y(\delta+1)(\delta+2)}{\alpha} \leqslant \frac{1}{2}$. This in turn implies that

$$\alpha \geqslant 2\sigma_y(\delta+1)(\delta+2)\,.$$

This completes the proof that v_2 is super Lyapunov in $\mathcal{R}_2(\alpha)^{(2)}$, but we have obtained our first restriction on the value of α. □

Again we emphasize that the above proof only invokes the explicit solution for v_2 in order to verify its scaling and smoothness properties, as well as its limiting behavior as the parameter λ approaches zero. In a more complex setting, an explicit solution is unnecessary as long as these properties can be verified. We also note that the dependence upon σ_y in the restriction on α is unsurprising in light of the qualitative behavior of the dynamics. α may be viewed as a parameter which determines how closely the transport region approaches the positive x-axis. In the case of no noise in the y-direction, transport is the governing behavior all the way up to the axis. However, when there is noise added in the y-direction, the noise is dominant close to the axis, and hence, the region where transport dominates is restricted.

Choosing $m_2 = \min\{m_{21}, m_{22}\}$, $\gamma_2 = \min\{\gamma_{21}, \gamma_{22}\}$, and $R_2 = \max\{R_{21}, R_{22}\}$ completes the proof that v_2 is a local super Lyapunov function in the entire region $\mathcal{R}_2(\alpha)$.

4.5.3 The Diffusive Region

For $\delta > 0$ and $\alpha > 0$, we define v_3 by the following Poisson equation:

$$\begin{cases} (Av_3)(x,y) = -c_1 x^{\hat{\delta}+1} & \text{on } \mathcal{R}_3(\alpha) \\ v_3(x,y) = c_2 x^{\hat{\delta}} & \text{on } \mathcal{C}^{(1)}(2\alpha, \pm 1) \end{cases} \tag{4.9}$$

where $c_1, c_2 > 0$ are constants which will be chosen later. We will show that the values of c_1 and c_2 do not affect the local super Lyapunov property of v_3, but rather are chosen in order to facilitate the patching of the local super Lyapunov functions into one global super Lyapunov function in Section 4.6. Recall that

$$\mathcal{C}^{(1)}(2\alpha, \pm 1) = \left\{ x \geqslant 2\alpha, |y| = \sqrt{\frac{2\alpha}{x}} \right\}$$

is the boundary of $\mathcal{R}_3(\alpha)$ where it overlaps with $\mathcal{R}_2(\alpha)$ and $\hat{\delta} = \frac{5\delta+3}{2}$.

As before for the transport region, we have chosen a right-hand side for the Poisson equation which is negative definite, scales homogeneously under the appropriate scaling, namely $S_\ell^{(1)}$, and has unbounded growth in the region. We use a constant multiple of $x^{\hat{\delta}}$ as the boundary condition rather than the function v_2 from the neighboring region because we want a function which scales homogeneously under $S_\ell^{(1)}$. However, $x^{\hat{\delta}}$ is in fact the asymptotic behavior (up to a constant multiple) of $v_2(x,y)$ as $|(x,y)| \to \infty$ on the specified boundary.

The dynamics associated to the operator A, which is dominant in $\mathcal{R}_3(\alpha)$, can be understood as having one diffusive direction and one deterministic direction which is uncoupled from the diffusion and acts as the "clock" of the diffusion. To see this, observe that A is the operator associated to the system of SDEs given by

$$\begin{aligned} d\hat{X}_t &= \hat{X}_t^2 dt \\ d\hat{Y}_t &= 2\hat{X}_t\hat{Y}_t dt + \epsilon_y \, dW_t^y \, . \end{aligned} \tag{4.10}$$

Now let $(\hat{X}_0, \hat{Y}_0) = (x, y)$ lie in the interior of $\mathcal{R}_3(\alpha)$ and define the stopping time

$$\hat{\tau} = \inf\{t > 0 : (\hat{X}_t, \hat{Y}_t) \in \mathcal{C}^{(1)}(2\alpha, \pm 1)\} \, .$$

Then $v_3(x, y)$, which was defined in (4.9), can be represented probabilistically as

$$\begin{aligned} v_3(x, y) &= c_2 \mathbb{E}_{(x,y)} \left[\hat{X}_{\hat{\tau}}^{\hat{\delta}} \right] + c_1 \mathbb{E}_{(x,y)} \left[\int_0^{\hat{\tau}} \hat{X}_s^{\hat{\delta}+1} ds \right] \\ &= \left(\tfrac{c_1}{\hat{\delta}} + c_2 \right) \mathbb{E}_{(x,y)} \left[\hat{X}_{\hat{\tau}}^{\hat{\delta}} \right] - \tfrac{c_1}{\hat{\delta}} x^{\hat{\delta}} \end{aligned} \tag{4.11}$$

since $\hat{X}_t = \frac{x}{1-xt}$. This probabilistic representation is only valid provided, first, that the expectation is finite and, second, that the right-hand side of equation (4.11) depends in a C^2 fashion on $(x, y) \in \mathcal{R}_3(\alpha)$. Both of these facts will follow from Lemma 4.3, which we use in Proposition 4.5 to show that v_3 is a super Lyapunov function.

Since $\hat{X}_t$ is deterministic, the representation of v_3 given in (4.11) amounts to a deterministic function of $\hat{\tau}$. To better understand the properties of $\hat{\tau}$, we introduce the time change

$$T(t) = \int_0^t \hat{X}_s ds = -\ln|1 - xt|$$

and the process $Z_{T(t)} = \hat{X}_t^{\frac{1}{2}} \hat{Y}_t$. Due to the scaling of the boundary of $\mathcal{R}_3(\alpha)$, if we define the stopping time

$$\tau = \inf\{T > 0 : |Z_T| \geqslant \sqrt{2\alpha}\}$$

then

$$\hat{\tau} = \frac{1}{x}(1 - e^{-\tau}) \quad \text{and} \quad \hat{X}_t = xe^{T(t)}\,.$$

In addition, Z_T satisfies the SDE

$$dZ_T = \frac{5}{2} Z_T dT + \sqrt{2\sigma_y}\, dW_T\,, \qquad Z_0 = x^{\frac{1}{2}} y\,. \tag{4.12}$$

The time change enables us to rewrite the stochastic representation formula for v_3 given by (4.11) as

$$v_3(x, y) = x^{\hat{\delta}} \left[\left(\tfrac{c_1}{\hat{\delta}} + c_2 \right) \mathbb{E}_{(x,y)} \left[e^{\hat{\delta}\tau} \right] - \tfrac{c_1}{\hat{\delta}} \right]. \tag{4.13}$$

As a consequence of this formula, we observe that $v_3 \overset{1}{\sim} \ell^{\hat{\delta}}$. To prove this scaling, it suffices to show that

$$\mathbb{E}_{(x,y)}\big[e^{\hat{\delta}\tau}\big] \overset{1}{\sim} \ell^0\,.$$

This is clear, since the only dependence of τ upon x and y results from Z_0, and $Z_0 = x^{\frac{1}{2}} y = (\ell x)^{\frac{1}{2}} (y \ell^{-\frac{1}{2}})$ is invariant under $S_\ell^{(1)}$.

Using the time change and the process Z_T, we are able to prove the following lemma, which combined with classical results (see, for example, [5]), justifies the stochastic representation for v_3 given by (4.13).

Lemma 4.3. *For any $\sigma_y > 0$, $\hat{\delta} < \frac{5}{2}$ and $(x, y) \in \mathcal{R}_3(\alpha)$, $\mathbb{E}_{(x,y)}\big[e^{\hat{\delta}\tau}\big] < \infty$ and the map $(x, y) \mapsto \mathbb{E}_{(x,y)}\big[e^{\hat{\delta}\tau}\big]$ is C^2.*

Proof of Lemma 4.3. To show the finiteness of the expectation, we observe that

$$\begin{aligned}
\mathbb{P}_{(x,y)}(e^{\hat{\delta}\tau} > s) &= \mathbb{P}_{(x,y)}\left(\sup_{0\leqslant T\leqslant \frac{\ln s}{\hat{\delta}}} |Z_T| < \sqrt{2\alpha}\right) \leqslant \mathbb{P}_{(x,y)}\left(|Z_{\frac{1}{\hat{\delta}}\ln s}| < \sqrt{2\alpha}\right)\\
&\leqslant \mathbb{P}\left(\left|\sqrt{x}\ y\ s^{\frac{5}{2\hat{\delta}}} + \sqrt{2\sigma_y}\ s^{\frac{5}{2\hat{\delta}}}\int_0^{\frac{1}{\hat{\delta}}\ln s} e^{-\frac{5}{2}r} dW_r\right| < \sqrt{2\alpha}\right)\\
&\leqslant \mathbb{P}\left(\left|\sqrt{2\sigma_y}\ s^{\frac{5}{2\hat{\delta}}}\int_0^{\frac{1}{\hat{\delta}}\ln s} e^{-\frac{5}{2}r} dW_r\right| < \sqrt{2\alpha}\right)\\
&= \mathbb{P}\left(|Z| < \sqrt{\frac{5\alpha}{\sigma_y(s^{\frac{5}{\hat{\delta}}}-1)}}\right) \quad \text{where} \quad Z \sim N(0,1)\\
&\leqslant \left(\frac{10\alpha}{\sigma_y\pi(s^{5/\hat{\delta}}-1)}\right)^{\frac{1}{2}}.
\end{aligned}$$

Hence for $\hat{\delta} < \frac{5}{2}$, this upper bound decays sufficiently rapidly in order to guarantee that $\mathbb{E}_{(x,y)}\big[e^{\hat{\delta}\tau}\big]$ is finite. The continuity properties now follow from the continuity properties of τ. Specifically, $\mathbb{E}_{(x,y)}\big[e^{\hat{\delta}\tau}\big] = g(\sqrt{x}\,y)$ where $g(z)$ solves the following ODE:

$$\begin{cases} \sigma_y g''(z) + \frac{5}{2}zg'(z) + \hat{\delta}g(z) = 0 \quad \text{for } g \in (-\sqrt{2\alpha}, \sqrt{2\alpha}) \\ g(\sqrt{2\alpha}) = g(-\sqrt{2\alpha}) = 1\,. \end{cases} \tag{4.14}$$

Since by standard results on the regularity of ODEs, $g(z) \in C^2([-\sqrt{2\alpha}, \sqrt{2\alpha}])$, we conclude that $\mathbb{E}_{(x,y)}\big[e^{\hat{\delta}\tau}\big] = g(\sqrt{x}\ y) \in C^2(\mathcal{R}_3(\alpha))$ as desired. □

We remark that this lemma imposes a further restriction on the size of the parameter δ, which previously was only required to be in $(0, 2)$ from the priming region; namely, that $\hat{\delta} < \frac{5}{2}$, which forces $\delta \in (0, \frac{2}{5})$. In addition, this lemma produces the first restriction on the noise coefficients; namely, that $\sigma_y > 0$. Hence, the noise in the y-direction is crucial to the existence of a local Lyapunov function in the diffusive region and to the overall stabilization. On the other hand, there is no restriction on the value of σ_x in the diffusive region, or in any of the other regions, implying that the noise in the x-direction is unnecessary for stabilization to occur.

Using Lemma 4.3, we now turn to proving that v_3 is a super Lyapunov function on $\mathcal{R}_3(\alpha)$.

Proposition 4.5. *For any $\delta \in (0, \frac{2}{5})$ and $\alpha > 0$, there exist constants $m_3 > 0$ and $R_3 \geqslant 0$ so that if $(x, y) \in \mathcal{R}_3(\alpha)$ with $|(x, y)| \geqslant R_3$, then v_3 satisfies*

$$(\mathcal{L}v_3)(x, y) \leqslant -m_3 v_3^{\gamma_3}(x, y)$$

where $\gamma_3 = \gamma_{22} = \frac{5\delta+5}{5\delta+3} > 1$.

Proof of Proposition 4.5. Applying the generator to v_3, we obtain

$$\begin{aligned}(\mathcal{L}v_3)(x, y) &= (Av_3)(x, y) - y^2 \partial_x v_3(x, y) + \sigma_x \partial_{xx} v_3(x, y) \\ &= -c_1 x^{\hat{\delta}+1} - y^2 \partial_x v_3(x, y) + \sigma_x \partial_{xx} v_3(x, y)\,.\end{aligned}$$

The scaling relations are again used to prove the super Lyapunov property. The scaling $S_\ell^{(1)}$ is used since it is the transformation under which the operator A is dominant and scales homogeneously. Every point $(x, y) \in \mathcal{R}_3(\alpha)$ can be mapped back to a point $(2\alpha, b)$, where $(x, y) = S_\ell^{(1)}(2\alpha, b)$, $\ell = \frac{x}{2\alpha}$, and $b = \sqrt{\ell}y \in [-1, 1]$. Therefore v_3 satisfies the following scaling relations:

$$\begin{aligned} v_3(x, y) &= \ell^{\hat{\delta}} v_3(2\alpha, b) & \qquad (\partial_x v_3)(x, y) &= \ell^{\hat{\delta}-1} (\partial_x v_3)(2\alpha, b) \\ x^{\hat{\delta}+1} &= \ell^{\hat{\delta}+1} (2\alpha)^{\hat{\delta}+1} & \qquad (\partial_{xx} v_3)(x, y) &= \ell^{\hat{\delta}-2} (\partial_{xx} v_3)(2\alpha, b)\,. \end{aligned}$$

These scaling relations lead us to choose $\gamma_3 = \frac{\hat{\delta}+1}{\hat{\delta}} = \frac{5\delta+5}{5\delta+3}$, which is the ratio of the exponents of ℓ in Av_3 and v_3. With this choice of γ_3, it follows that

$$\begin{aligned}&\frac{c_1 x^{\hat{\delta}+1} - y^2 \partial_x v_3(x, y) - \sigma_x \partial_{xx} v_3(x, y)}{v_3^{\gamma_3}(x, y)} \\ &\qquad\qquad = \frac{c_1 (2\alpha)^{\hat{\delta}+1} - \ell^{-3} (b^2 \partial_x v_3(2\alpha, b) + \sigma_x \partial_{xx} v_3(2\alpha, b))}{v_3^{\gamma_3}(2\alpha, b)}\,.\end{aligned}$$

Constants ρ and M are defined by

$$\rho = \inf_{b \in [-1,1]} \frac{c_1 (2\alpha)^{\hat{\delta}+1}}{v_3^{\gamma_3}(2\alpha, b)}$$

$$M = \sup_{b \in [-1,1]} \frac{b^2 |\partial_x v_3(2\alpha, b)| + \sigma_x |\partial_{xx} v_3(2\alpha, b)|}{v_3^{\gamma_3}(2\alpha, b)} \,.$$

Since v_3 is C^2 and nonvanishing on $\mathcal{R}_3(\alpha)$, $\rho > 0$ and $M < \infty$ (the infimum and supremum are over a compact set). From the definition of these constants, it follows that

$$(\mathcal{L} v_3)(x, y) \leqslant -v_3^{\gamma_3}(x, y)(\rho - \ell^{-3} M) \,.$$

Setting $m_3 = \frac{\rho}{2}$, it only remains to show that R_3 can be chosen so that $\rho - \ell^{-3} M \geqslant \frac{\rho}{2}$ for all $(x, y) \in \mathcal{R}_3(\alpha)$ with $|(x, y)| \geqslant R_3$. To accomplish this, we define

$$\ell_* = \inf \left\{ \ell = \frac{x}{2\alpha} : (x, y) \in \mathcal{R}_3(\alpha) \text{ with } |(x, y)| \geqslant R_3 \right\}$$

which approaches infinity as $R_3 \to \infty$. Thus, R_3 can be chosen large enough in order to ensure that $\ell \geqslant \left(\frac{2M}{\rho}\right)^{\frac{1}{3}}$. □

4.6 Construction of Global Super Lyapunov Function

We now patch together the three local Lyapunov functions that are defined in distinct regions of the plane in order to produce one smooth, global Lyapunov function defined on the entire plane for the model problem.

4.6.1 Definition of Global Super Lyapunov Function

To facilitate the construction of the global super Lyapunov function, we use the standard mollifier $\phi(t)$ described in Chapter 2; i.e, $\phi(t)$ is a smooth, increasing function which varies from zero to one and is suitable normalized to integrate to one on the entire real line. Next, we define the functions $h_1(x, y)$ and $h_2(x, y)$ as follows:

$$h_1(x, y) = 2 + \frac{\alpha |y|}{x} \quad \text{and} \quad h_2(x, y) = 2 - \frac{x y^2}{\alpha} \,.$$

The function $h_1(x,y) = 0$ on one boundary of the wedge-shaped region $\mathcal{R}_1(\alpha) \cap \mathcal{R}_2(\alpha)$ and $h_1(x,y) = 1$ on the other boundary of this region; moreover, h_1 varies smoothly between 0 and 1 in the interior. Similarly, $h_2(x,y) = 0$ on one boundary of the funnel-like region $\mathcal{R}_2(\alpha) \cap \mathcal{R}_3(\alpha)$ and $h_2(x,y) = 1$ on the other boundary, again varying smoothly between 0 and 1 in the interior.

Thus, outside of a fixed ball, we define our global Lyapunov function V to agree with the local Lyapunov functions in subregions of their domains of definition and to be a smooth, convex combination of the two local Lyapunov functions in the regions of intersection. In particular, let $\tilde{V}(x,y)$ be given by

$$\tilde{V}(x,y) = \begin{cases} v_1(x,y) & \text{for } (x,y) \in \mathcal{R}_1(\alpha) \cap \mathcal{R}_2(\alpha)^c \\ V_{1,2}(x,y) & \text{for } (x,y) \in \mathcal{R}_1(\alpha) \cap \mathcal{R}_2(\alpha) \\ v_2(x,y) & \text{for } (x,y) \in \mathcal{R}_2(\alpha) \cap \mathcal{R}_1(\alpha)^c \cap \mathcal{R}_3(\alpha)^c \\ V_{2,3}(x,y) & \text{for } (x,y) \in \mathcal{R}_2(\alpha) \cap \mathcal{R}_3(\alpha) \\ v_3(x,y) & \text{for } (x,y) \in \mathcal{R}_3(\alpha) \cap \mathcal{R}_2(\alpha)^c \\ 0 & \text{otherwise} \end{cases}$$

where

$$V_{1,2}(x,y) = [1 - \phi(h_1(x,y))]v_2(x,y) + \phi(h_1(x,y))v_1(x,y)$$

$$V_{2,3}(x,y) = [1 - \phi(h_2(x,y))]v_2(x,y) + \phi(h_2(x,y))v_3(x,y)\,.$$

We then choose $V(x,y) \in C^2(\mathbb{R}^2)$ to satisfy

$$V(x,y) = \begin{cases} \tilde{V}(x,y) & \text{for } x^2 + y^2 > R^2 \\ \text{arbitrary positive and smooth} & \text{for } x^2 + y^2 \leqslant R^2 \end{cases} \tag{4.15}$$

where R will be specified below.

At the start of the Lyapunov construction in Section 4.5.1, we fix a choice of $\delta \in (0,2)$ when defining v_1. This choice is then propagated through our construction and is explicitly present in the definition of v_2 and v_3. During the analysis of v_3, we noted in Lemma 4.3 that we must further restrict $\delta \in (0, \frac{2}{5})$. Except for this one restriction, the choice of δ is free. In addition, all three of the local Lyapunov functions and their regions of definition are dependent upon the parameter α. In Lemma 4.2, we noted that α must be chosen sufficiently large in order to guarantee that v_2 is super Lyapunov in the subregion of $\mathcal{R}_2$ that is near region $\mathcal{R}_3$. We will demonstrate below that further restrictions on α are required in order to show the existence of a global super Lyapunov function. Nevertheless, a choice of α can be fixed to ensure that V, defined by (4.15), is super Lyapunov.

4.6.2 Super Lyapunov Property in Overlap Regions

The main missing ingredient in showing that V is a global super Lyapunov function is the verification that V is a local Lyapunov function on the overlap regions. In Propositions 4.6 and 4.7, we show that $V_{1,2}$ and $V_{2,3}$, respectively, are local Lyapunov functions on their regions of definition. Then in Section 4.6.3, we put everything together to prove Theorem 4.1, which states that there exists a global super Lyapunov function for the model problem.

Before stating and proving the two propositions, we first define a few necessary constants. Let m_i, R_i, and γ_i for $i = 1, 2, 3$ be the constants from Propositions 4.3, 4.4, and 4.5. Then define $m_* = \min\{m_1, m_2, m_3\}$, $\gamma_* = \min\{\gamma_1, \gamma_2, \gamma_3\} = \frac{5\delta+5}{5\delta+3}$ and $R_* \geqslant \max\{R_1, R_2, R_3\}$ so that for $i = 1, 2, 3$, $v_i(x, y) > 1$ for all $(x, y) \in \mathcal{R}_i(\alpha)$ with $|(x, y)| \geqslant R_*$.

Proposition 4.6. *For any $\delta \in (0, \frac{2}{5})$, $V_{1,2}(x, y)$ is a local super Lyapunov function on $\mathcal{R}_1(\alpha) \cap \mathcal{R}_2(\alpha)$.*

Proof of Proposition 4.6. If $r > R_*$, we have that for all $(x,y) \in \mathcal{R}_1(\alpha) \cap \mathcal{R}_2(\alpha) \cap B_r^c$

$$\begin{aligned}
(\mathcal{L}V_{1,2})(x,y) = &\ (1-\phi(h_1(x,y)))(\mathcal{L}v_2)(x,y) + \phi(h_1(x,y))(\mathcal{L}v_1)(x,y) + E_1(x,y) \\
&\leqslant -(1-\phi(h_1(x,y)))m_2\ v_2^{\gamma_1}(x,y) - \phi(h_1(x,y))m_1\ v_1^{\gamma_1}(x,y) + E_1(x,y) \\
&\leqslant -m_*[(1-\phi(h_1(x,y)))v_2^{\gamma_1}(x,y) + \phi(h_1(x,y))v_1^{\gamma_1}(x,y)] + E_1(x,y) \\
&\leqslant -m_*[V_{1,2}(x,y)]^{\gamma_1} + E_1(x,y) \text{ by convexity} \\
&= -m_*[V_{1,2}(x,y)]^{\gamma_1}\left[1 - \frac{E_1(x,y)}{m_*[V_{1,2}(x,y)]^{\gamma_1}}\right] \\
&\leqslant -m_*(1-M_1)[V_{1,2}(x,y)]^{\gamma_1}
\end{aligned}$$

where M_1 and $E_1(x,y)$ are defined as

$$M_1 = \sup_{(x,y)\in\mathcal{R}_1(\alpha)\cap\mathcal{R}_2(\alpha)\cap B_r^c}\left[\frac{E_1(x,y)}{m_*[V_{1,2}(x,y)]^{\gamma_1}}\right]$$

and

$$\begin{aligned}
E_1(x,y) = &\ \mathcal{L}[\phi(h_1(x,y))](v_1(x,y) - v_2(x,y)) \\
&+ 2\sigma_x\frac{\partial}{\partial x}[\phi(h_1(x,y))]\frac{\partial}{\partial x}[v_1(x,y) - v_2(x,y)] \\
&+ 2\sigma_y\frac{\partial}{\partial y}[\phi(h_1(x,y))]\frac{\partial}{\partial y}[v_1(x,y) - v_2(x,y)]\,.
\end{aligned}$$

If we can choose r sufficiently large so that $M_1 < 1$, then $V_{1,2}(x,y)$ will be a local super Lyapunov function on $\mathcal{R}_1(\alpha) \cap \mathcal{R}_2(\alpha)$. To show $M_1 < 1$, we use the scaling properties of v_1 and v_2 to map back to a circular arc of radius $r = \sqrt{\alpha^2+4}$. Let

$$\ell = \frac{\sqrt{x^2+y^2}}{r}, \quad a = \frac{x}{\ell}, \quad \text{and} \quad b = \frac{y}{\ell}\,.$$

Then

$$(x,y) \in \mathcal{R}_1(\alpha) \cap \mathcal{R}_2(\alpha) \cap B_r^c \quad \Longrightarrow \quad (x,y) = S_\ell^{(2)}(a,b) \quad \text{with} \quad \ell \geqslant 1\,.$$

Note that $h_1(x,y) = h_1(a,b)$, so

$$V_{1,2}(x,y) = \ell^\delta V_{1,2}(a,b) \quad \text{and} \quad [V_{1,2}(x,y)]^{\gamma_1} = \ell^{\delta+1}[V_{1,2}(a,b)]^{\gamma_1}\,.$$

As a consequence of these scaling relations, we get that for all $(x,y) \in \mathcal{R}_1(\alpha) \cap \mathcal{R}_2(\alpha) \cap B_r^c$,

$$
\begin{aligned}
E_1(x,y) &= \alpha|y|\left(1+\frac{y^2}{x^2}\right)\phi'(h_1(x,y))[v_1-v_2](x,y) \\
&\quad + \sigma_x\left[\frac{2\alpha|y|}{x^3}\phi'(h_1(x,y)) - \frac{\alpha|y|}{x^2}\phi''(h_1(x,y))\right][v_1-v_2](x,y) \\
&\quad + \frac{\sigma_y \mathrm{sgn}(y)\alpha}{x}\phi''(h_1(x,y))[v_1-v_2](x,y) \\
&\quad - \frac{2\sigma_x\alpha|y|}{x^2}\phi'(h_1(x,y))\frac{\partial}{\partial x}[v_1-v_2](x,y) \\
&\quad + \frac{2\sigma_y \mathrm{sgn}(y)\alpha}{x}\phi'(h_1(x,y))\frac{\partial}{\partial y}[v_1-v_2](x,y) \\
&= \ell^{\delta+1}\alpha\phi'(h_1(a,b))|b|\left(1+\frac{b^2}{a^2}\right)[v_1(a,b)-v_2(a,b)] \\
&\quad + \ell^{\delta-1}\alpha\phi''(h_1(a,b))\left(\frac{-\sigma_x|b|}{a^2}+\frac{\sigma_y\mathrm{sgn}(b)}{a}\right)[v_1(a,b)-v_2(a,b)] \\
&\quad + \ell^{\delta-2}\alpha\phi'(h_1(a,b))\frac{2\sigma_x|b|}{a^3}[v_1(a,b)-v_2(a,b)] \\
&\quad + \ell^{\delta-2}\alpha\phi'(h_1(a,b))\frac{-2\sigma_x|b|}{a^2}\frac{\partial}{\partial x}[v_1(a,b)-v_2(a,b)]) \\
&\quad + \ell^{\delta-2}\alpha\phi'(h_1(a,b))\frac{2\sigma_y\mathrm{sgn}(b)}{a}\frac{\partial}{\partial y}[v_1(a,b)-v_2(a,b)]\,.
\end{aligned}
$$

Hence we have that

$$
M_1 \leqslant \sup_{\substack{(a,b)\in\mathcal{R}_1(\alpha)\cap\mathcal{R}_2(\alpha) \\ |(x,y)|\geqslant r}}\left[\frac{e_{1,1}(a,b)}{m_*[V_{1,2}(a,b)]^{\gamma_1}} + \frac{e_{1,2}(a,b)}{\ell^2 m_*[V_{1,2}(a,b)]^{\gamma_1}}\right] \tag{4.16}
$$

where

$$
\begin{aligned}
e_{1,1}(a,b) &= \alpha\phi'(h_1(a,b))|b|\Big(1+\frac{b^2}{a^2}\Big)[v_1(a,b)-v_2(a,b)] \\
e_{1,2}(a,b) &= \alpha\phi''(h_1(a,b))\Big(\frac{-\sigma_x|b|}{a^2}+\frac{\sigma_y \mathrm{sgn}(b)}{a}\Big)[v_1(a,b)-v_2(a,b)] \\
&\quad + \alpha\phi'(h_1(a,b))\frac{2\sigma_x|b|}{a^3}[v_1(a,b)-v_2(a,b)] \\
&\quad + \alpha\phi'(h_1(a,b))\frac{-2\sigma_x|b|}{a^2}\frac{\partial}{\partial x}[v_1(a,b)-v_2(a,b)] \\
&\quad + \alpha\phi'(h_1(a,b))\frac{2\sigma_y \mathrm{sgn}(b)}{a}\frac{\partial}{\partial y}[v_1(a,b)-v_2(a,b)]\,.
\end{aligned}
$$

By explicit computation with v_1 and v_2, we can show that $e_{1,1}(a,b)$ is always negative for (a,b) in the desired region. The second term of the sum in (4.16), the upper bound for M_1, can then be made arbitrarily small by choosing ℓ large enough; this corresponds to choosing r sufficiently large. This establishes that $M_1 < 1$, which completes the proof of the proposition. □

We now turn to proving that $V_{2,3}$ is super Lyapunov.

Proposition 4.7. *For any $\delta \in (0, \frac{2}{5})$, $V_{2,3}(x,y)$ is a local super Lyapunov function on $\mathcal{R}_2(\alpha) \cap \mathcal{R}_3(\alpha)$.*

Proof of Proposition 4.7. If $r > R_*$, then for all $(x,y) \in \mathcal{R}_2(\alpha) \cap \mathcal{R}_3(\alpha) \cap B_r^c$

$$
\begin{aligned}
(\mathcal{L}V_{2,3})(x,y) &= (1-\phi(h_2(x,y)))(\mathcal{L}v_2)(x,y) + \phi(h_2(x,y))(\mathcal{L}v_3)(x,y) + E_2(x,y) \\
&\leqslant -(1-\phi(h_2(x,y)))m_2\, v_2^{\gamma_3}(x,y) - \phi(h_2(x,y))m_3\, v_3^{\gamma_3}(x,y) + E_2(x,y) \\
&\leqslant -m_*[(1-\phi(h_2(x,y)))v_2^{\gamma_3}(x,y) + \phi(h_2(x,y))v_3^{\gamma_3}(x,y)] + E_2(x,y) \\
&\leqslant -m_*[V_{2,3}(x,y)]^{\gamma_3} + E_2(x,y) \quad \text{by convexity} \\
&= -m_*[V_{2,3}(x,y)]^{\gamma_3}\left[1 - \frac{E_2(x,y)}{m_*[V_{2,3}(x,y)]^{\gamma_3}}\right] \\
&\leqslant -m_*(1-M_2)[V_{2,3}(x,y)]^{\gamma_3}
\end{aligned}
$$

where

$$M_2 = \sup_{(x,y)\in\mathcal{R}_2(\alpha)\cap\mathcal{R}_3(\alpha)\cap B_r^c} \left[\frac{E_2(x,y)}{m_*[V_{2,3}(x,y)]^{\gamma_3}}\right]$$

and

$$\begin{aligned} E_2(x,y) &= \mathcal{L}[\phi(h_2(x,y))](v_3(x,y) - v_2(x,y)) \\ &\quad + 2\sigma_x \frac{\partial}{\partial x}[\phi(h_2(x,y))]\frac{\partial}{\partial x}[v_3(x,y) - v_2(x,y)] \\ &\quad + 2\sigma_y \frac{\partial}{\partial y}[\phi(h_2(x,y))]\frac{\partial}{\partial y}[v_3(x,y) - v_2(x,y)]\,. \end{aligned}$$

If $M_2 < 1$, then $V_{2,3}(x,y)$ will be a super Lyapunov function on $\mathcal{R}_2(\alpha) \cap \mathcal{R}_3(\alpha)$. To show $M_2 < 1$, we use the scaling properties of v_2 and v_3 to map back to a vertical line. Let

$$\ell = \frac{x}{2\alpha}, \qquad a = 2\alpha, \quad \text{and} \quad b = y\sqrt{\ell}\,.$$

Then

$$(x,y) \in \mathcal{R}_2(\alpha) \cap \mathcal{R}_3(\alpha) \cap B_r^c \quad \Longrightarrow \quad (x,y,1) = S_\ell^{(1)}(a,b,\ell^{-3})$$

where $|b| \in \left[\frac{1}{\sqrt{2}}, 1\right]$ and $\ell \geqslant 1$. Note that $h_2(x,y) = h_2(a,b)$, so $V_{2,3}(x,y) = V_{2,3}(x,y,1)$ satisfies

$$V_{2,3}(x,y,1) = \ell^{\hat{\delta}} V_{2,3}(a,b,\ell^{-3}) \quad \text{and} \quad [V_{2,3}(x,y,1)]^{\gamma_3} = \ell^{\hat{\delta}+1}[V_{2,3}(a,b,\ell^{-3})]^{\gamma_3}$$

where

$$V_{2,3}(x,y,\lambda) = [1 - \phi(h_2(x,y))]v_2(x,y,\lambda) + \phi(h_2(x,y))v_3(x,y)\,.$$

Now we have that for all $(x, y) \in \mathcal{R}_2(\alpha) \cap \mathcal{R}_3(\alpha) \cap B_r^c$,

$$\begin{aligned}
E_2(x,y) &= \frac{1}{\alpha}\left[-x(5xy^2+2\sigma_y)+y^4\right]\phi'(h_2(x,y))(v_3(x,y)-v_2(x,y)) \\
&\quad - \frac{1}{\alpha}\left[\sigma_x y^2 + 2\sigma_y xy\right]\phi''(h_2(x,y))(v_3(x,y)-v_2(x,y)) \\
&\quad - \frac{2\sigma_x y^2}{\alpha}\phi'(h_2(x,y))\frac{\partial}{\partial x}[v_3(x,y)-v_2(x,y)] \\
&\quad - \frac{4\sigma_y xy}{\alpha}\phi'(h_2(x,y))\frac{\partial}{\partial y}[v_3(x,y)-v_2(x,y)] \\
&= \ell^{\hat{\delta}+1}\frac{\phi'(h_2(a,b))}{\alpha}(-5a^2b^2-2a\sigma_y)[v_3(a,b)-v_2(a,b,\ell^{-3})] \\
&\quad + \ell^{\hat{\delta}+1}\frac{\phi'(h_2(a,b))}{\alpha}(-4ab\sigma_y)\frac{\partial}{\partial y}[v_3(a,b)-v_2(a,b,\ell^{-3})] \\
&\quad + \ell^{\hat{\delta}+\frac{1}{2}}\frac{\phi''(h_2(a,b))}{\alpha}(-2ab\sigma_y)[v_3(a,b)-v_2(a,b,\ell^{-3})] \\
&\quad + \ell^{\hat{\delta}-1}\frac{\phi''(h_2(a,b))}{\alpha}(-b^2\sigma_x)[v_3(a,b)-v_2(a,b,\ell^{-3})] \\
&\quad + \ell^{\hat{\delta}-2}\frac{\phi'(h_2(a,b))}{\alpha}b^4[v_3(a,b)-v_2(a,b,\ell^{-3})] \\
&\quad + \ell^{\hat{\delta}-2}\frac{\phi'(h_2(a,b))}{\alpha}(-2b^2\sigma_x)\frac{\partial}{\partial x}[v_3(a,b)-v_2(a,b,\ell^{-3})]\,.
\end{aligned}$$

Define $N(\lambda^*)$ as follows:

$$N(\lambda^*) = \sup_{\substack{|b|\in[\frac{1}{\sqrt{2}},1] \\ \lambda\in(0,\lambda^*]}}\left[\frac{e_{2,1}(a,b,\lambda)}{m_*[V_{2,3}(a,b,\lambda)]^{\gamma_3}} + \frac{e_{2,2}(a,b,\lambda)}{\sqrt{\ell}m_*[V_{2,3}(a,b,\lambda)]^{\gamma_3}}\right] \tag{4.17}$$

where

$$
\begin{aligned}
e_{2,1}(a,b,\lambda) = {} & \frac{\phi'(h_2(a,b))}{\alpha}(-5a^2b^2 - 2a\sigma_y)[v_3(a,b) - v_2(a,b,\lambda)] \\
& + \frac{\phi'(h_2(a,b))}{\alpha}(-4ab\sigma_y)\frac{\partial}{\partial y}[v_3(a,b) - v_2(a,b,\lambda)]\,, \\
e_{2,2}(a,b,\lambda) = {} & \frac{\phi''(h_2(a,b))}{\alpha}(-2ab\sigma_y)[v_3(a,b) - v_2(a,b,\lambda)] \\
& + \frac{\phi''(h_2(a,b))}{\alpha}(-b^2\sigma_x)[v_3(a,b) - v_2(a,b,\lambda)] \\
& + \frac{\phi'(h_2(a,b))}{\alpha}b^4[v_3(a,b) - v_2(a,b,\lambda)] \\
& + \frac{\phi'(h_2(a,b))}{\alpha}(-2b^2\sigma_x)\frac{\partial}{\partial x}[v_3(a,b) - v_2(a,b,\lambda)]\,.
\end{aligned}
$$

Note that for any $\lambda^* > 0$, we can choose r sufficiently large to force M_2 (which, we recall, depends on r) to be less than $N(\lambda^*)$. Ultimately, we will choose λ^* sufficiently small so that $N(\lambda^*) < 1$. The second term of the sum in (4.17) can be made arbitrarily small by increasing the size of ℓ; again, increasing the size of ℓ corresponds to increasing r. We now address the first term of the sum in (4.17). From Lemma 4.4 which is stated and proven below, we see that we can choose the parameters to make this term negative. Combining all of these observations, we have demonstrated that $M_2 < 1$, which completes the proof of the proposition. □

Lemma 4.4. *There exist positive constants c_1 and c_2 in the definition of the Poisson equation for $v_3(x,y)$, and positive α and λ^* such that for all $\lambda \in [0,\lambda^*]$, the following inequalities hold for $a = 2\alpha$ and $|b| \in \left[\frac{1}{\sqrt{2}}, 1\right]$:*

$$v_3(a,b) - v_2(a,b,\lambda) > 0 \tag{4.18}$$

$$b\Big[\frac{\partial v_3}{\partial y}(a,b) - \frac{\partial v_2}{\partial y}(a,b,\lambda)\Big] > 0\,. \tag{4.19}$$

Proof of Lemma 4.4. Recall that, from (4.13), $v_3(x,y)$ can be represented as

$$v_3(x,y) = x^{\hat{\delta}}\Big[\big(\tfrac{c_1}{\hat{\delta}} + c_2\big)\mathbb{E}_{(x,y)}\big[e^{\hat{\delta}\tau}\big] - \tfrac{c_1}{\hat{\delta}}\Big] \tag{4.20}$$

where $\tau = \inf\{t > 0 : |Z_t| \notin [-\sqrt{2\alpha}, \sqrt{2\alpha}]\}$ and Z_t is the process given in (4.12). Note that the expectation in (4.20) can be written as the solution to a second-order ODE, namely:

$$\mathbb{E}_{(x,y)}\big[e^{\hat{\delta}\tau}\big] = g_\epsilon(\sqrt{\epsilon x}\ y)$$

where $g_\epsilon(z)$ solves the following boundary value problem with $\epsilon = \frac{1}{2\alpha}$:

$$\begin{cases} \epsilon\sigma_y g_\epsilon''(z) + \frac{5}{2}zg_\epsilon'(z) + \hat{\delta}g_\epsilon(z) = 0 \quad \text{ for } z \in (-1,1) \\ g_\epsilon(-1) = g_\epsilon(1) = 1\,. \end{cases} \tag{4.21}$$

Define $g_0(z)$ to be the solution to the limiting ODE in (4.21) when $\epsilon = 0$ and note that $g_0(z)$ can be computed exactly for initial conditions $z \neq 0$:

$$g_0(z) = \frac{1}{|z|^{\delta + \frac{3}{5}}}\,. \tag{4.22}$$

Now, let $v_3^0(x,y)$ be defined as

$$v_3^0(x,y) = x^{\hat{\delta}}\Big[\big(\tfrac{c_1}{\hat{\delta}} + c_2\big)g_0(\sqrt{\epsilon x}\ y) - \tfrac{c_1}{\hat{\delta}}\Big]\,. \tag{4.23}$$

We address the first difference between v_3 and v_2 in (4.18) as follows. We write

$$v_3(a,b) - v_2(a,b,\lambda) = v_3(a,b) - v_3^0(a,b) \tag{4.24}$$

$$+ v_3^0(a,b) - v_2(a,b,0) \tag{4.25}$$

$$+ v_2(a,b,0) - v_2(a,b,\lambda)\,. \tag{4.26}$$

To show that this difference is positive, we will show that $v_3^0(a,b) - v_2(a,b,0) > 0$ and that the other two differences are small in comparison. Similarly, for the difference between the y-derivatives of v_3 and v_2 in (4.19), we write

$$b\Big[\frac{\partial v_3}{\partial y}(a,b) - \frac{\partial v_2}{\partial y}(a,b,\lambda)\Big] = b\Big[\frac{\partial v_3}{\partial y}(a,b) - \frac{\partial v_3^0}{\partial y}(a,b)\Big] \tag{4.27}$$

$$+ b\Big[\frac{\partial v_3^0}{\partial y}(a,b) - \frac{\partial v_2}{\partial y}(a,b,0)\Big] \tag{4.28}$$

$$+ b\Big[\frac{\partial v_2}{\partial y}(a,b,0) - \frac{\partial v_2}{\partial y}(a,b,\lambda)\Big] \tag{4.29}$$

and again, we will show that $b\big[\frac{\partial v_3^0}{\partial y}(a,b) - \frac{\partial v_2}{\partial y}(a,b,0)\big] > 0$ and the other two differences are small in comparison. Specifically, we demonstrate that there exist positive constants c_1 and c_2 in the Poisson equation for v_3 such that the differences in (4.25) and (4.28) are positive; and then, that there exists α sufficiently large such that the differences on the righthand sides of (4.24) and (4.27) are comparatively small; and lastly, that there exists a λ^* such that (4.26) and (4.29) are comparatively small for all $\lambda \in [0, \lambda^*]$. For the first of these claims, note that

$$v_3^0(a,b) - v_2(a,b,0) = \frac{(2\alpha)^{2\delta+1}}{|b|^{\delta+1}} q(b) \tag{4.30}$$

$$b\Big[\frac{\partial v_3^0}{\partial y}(a,b) - \frac{\partial v_2}{\partial y}(a,b,0)\Big] = \frac{(2\alpha)^{2\delta+1}}{|b|^{\delta+1}} \tilde{q}(b) \tag{4.31}$$

where q and $\tilde{q}$ are given by

$$q(b) = 2^{\frac{1}{2}\delta+\frac{1}{2}}\Big[\big(\tfrac{\tilde{c_1}}{\hat{\delta}} + \tilde{c_2}\big)|b|^{\frac{2}{5}} - \frac{\tilde{c_1}}{\hat{\delta}}|b|^{\delta+1}\Big] - 1$$

$$\tilde{q}(b) = -\big(\delta + \tfrac{3}{5}\big)2^{\frac{1}{2}\delta+\frac{1}{2}}\big(\tfrac{\tilde{c_1}}{\hat{\delta}} + \tilde{c_2}\big)|b|^{\frac{2}{5}} + \delta + 1$$

and $c_1 = \frac{\tilde{c_1}}{\alpha^{\frac{1}{2}\delta+\frac{1}{2}}}$ and $c_2 = \frac{\tilde{c_2}}{\alpha^{\frac{1}{2}\delta+\frac{1}{2}}}$. We note that c_1 and c_2 are chosen to scale with α so that v_2 and v_3 have identical scaling in α. Moreover, as we demonstrate below, $\tilde{c_1}$ and $\tilde{c_2}$ can be chosen independently of α. It is clear that $\tilde{q}$ is a monotone decreasing function of $|b|$; hence it is minimized at the right endpoint of the interval for $|b|$, that is, $|b| = 1$. Thus if we can show $\tilde{q}(1) > 0$, then it follows that

$$b\Big[\frac{\partial v_3^0}{\partial y}(a,b) - \frac{\partial v_2}{\partial y}(a,b,0)\Big] > 0 \tag{4.32}$$

for all $|b| \in [2^{-\frac{1}{2}}, 1]$. If we can also show $q(2^{-\frac{1}{2}}) > 0$, then from (4.30), we conclude that

$$v_3^0(a, 2^{-\frac{1}{2}}) - v_2(a, 2^{-\frac{1}{2}}, 0) > 0\,.$$

Combining this with (4.32) gives the desired positivity of (4.25) on the whole interval $|b| \in [2^{-\frac{1}{2}}, 1]$. Hence, we need only verify that there exist positive values of $\tilde{c_1}$ and $\tilde{c_2}$ such that

$$q(2^{-\frac{1}{2}}) = 2^{\frac{1}{2}\delta+\frac{1}{2}}\Big[\big(\tfrac{\tilde{c_1}}{\hat{\delta}} + \tilde{c_2}\big)2^{-\frac{1}{5}} - \tfrac{\tilde{c_1}}{\hat{\delta}}2^{-(\frac{1}{2}\delta+\frac{1}{2})}\Big] - 1 > 0 \tag{4.33}$$

$$\tilde{q}(1) = -\big(\delta + \tfrac{3}{5}\big)2^{\frac{1}{2}\delta+\frac{1}{2}}\big(\tfrac{\tilde{c_1}}{\hat{\delta}} + \tilde{c_2}\big) + \delta + 1 > 0\,. \tag{4.34}$$

The verification of this is elementary and we omit the details.

We remark that $\tilde{c_1}$ and $\tilde{c_2}$ can be chosen independently of α, since the above inequalities have no dependence on α. Thus, choosing positive $\tilde{c_1}$ and $\tilde{c_2}$ such that (4.33) and (4.34) are both satisfied, we obtain that for all $|b| \in [2^{-\frac{1}{2}}, 1]$,

$$v_3^0(a,b) - v_2(a,b,0) \geqslant \alpha^{2\delta+1}2^{\frac{5}{2}\delta+\frac{3}{2}}q(2^{-\frac{1}{2}}) > 0$$

$$b\Big[\frac{\partial v_3^0}{\partial y}(a,b) - \frac{\partial v_2}{\partial y}(a,b,0)\Big] \geqslant \alpha^{2\delta+1}2^{2\delta+1}\tilde{q}(1) > 0\,.$$

We turn our attention to making the differences

$$v_3(a,b) - v_3^0(a,b) \quad \text{and} \quad b\Big[\frac{\partial v_3}{\partial y}(a,b) - \frac{\partial v_3^0}{\partial y}(a,b)\Big]$$

comparatively small. Note that

$$v_3(a,b) - v_3^0(a,b) = \alpha^{2\delta+1}2^{\frac{5}{2}\delta+\frac{3}{2}}\big(\tfrac{\tilde{c_1}}{\hat{\delta}} + \tilde{c_2}\big)[g_\epsilon(b) - g_0(b)]$$

$$b\Big[\frac{\partial v_3}{\partial y}(a,b) - \frac{\partial v_3^0}{\partial y}(a,b)\Big] = \alpha^{2\delta+1}2^{\frac{5}{2}\delta+\frac{3}{2}}b\big(\tfrac{\tilde{c_1}}{\hat{\delta}} + \tilde{c_2}\big)[g_\epsilon'(b) - g_0'(b)]\,.$$

To be precise, we will show that

$$|v_3(a,b) - v_3^0(a,b)| < \tfrac{1}{3}\big[\alpha^{2\delta+1}2^{\frac{5}{2}\delta+\frac{3}{2}}q(2^{-\frac{1}{2}})\big]$$

$$\left|b\Big[\frac{\partial v_3}{\partial y}(a,b) - \frac{\partial v_3^0}{\partial y}(a,b)\Big]\right| < \tfrac{1}{3}\left[\alpha^{2\delta+1}2^{2\delta+1}\tilde{q}(1)\right] .$$

This is equivalent to establishing that

$$\left|\alpha^{2\delta+1}2^{\frac{5}{2}\delta+\frac{3}{2}}\big(\tfrac{\tilde{c_1}}{\hat{\delta}} + \tilde{c_2}\big)\big[g_\epsilon(b) - g_0(b)\big]\right| < \tfrac{1}{3}\big[\alpha^{2\delta+1}2^{\frac{5}{2}\delta+\frac{3}{2}}q(2^{-\frac{1}{2}})\big] \tag{4.35}$$

$$\left|\alpha^{2\delta+1}2^{\frac{5}{2}\delta+\frac{3}{2}}b\big(\tfrac{\tilde{c_1}}{\hat{\delta}} + \tilde{c_2}\big)\big[g_\epsilon'(b) - g_0'(b)\big]\right| < \tfrac{1}{3}\left[\alpha^{2\delta+1}2^{2\delta+1}\tilde{q}(1)\right] . \tag{4.36}$$

Observe that the same powers of α appear on both sides of each of the above inequalities. Therefore, to prove (4.35) and (4.36), it suffices to show that $g_\epsilon(b)$ and $g_\epsilon'(b)$ converge uniformly to $g_0(b)$ and $g_0'(b)$, respectively, for $|b| \in [2^{-\frac{1}{2}}, 1]$ as $\epsilon = \frac{1}{2\alpha} \to 0$. Both of these uniform convergences follow from classical results; see, for example, [1].

Since $\epsilon = \frac{1}{2\alpha}$, we can choose α sufficiently large to guarantee that both (4.35) and (4.36) hold and that v_2 remains a local super Lyapunov function on $\mathcal{R}_2(\alpha)$ (recall that in Proposition 4.4, a lower bound on the size of α was imposed). Finally, by choosing λ^* sufficiently small, the differences in (4.26) and (4.29) can be made small for all $\lambda \in [0, \lambda*]$. This is an immediate consequence of the fact that $v_2(a,b,\lambda)$ is a C^2 function of $\lambda \in [0,1]$. □

4.6.3 Proof of Global Super Lyapunov Property

Having established the super Lyapunov property in the overlap regions, we return to the proof of the existence of a global super Lyapunov function for the model problem, which is the essence of Theorem 4.1.

Proof of Theorem 4.1. Choose V as defined by (4.15). The local super Lyapunov condition has already been proven in the three regions $\mathcal{R}_i(\alpha)$ through Propositions

4.3, 4.4, and 4.5, and in the overlap regions through Propositions 4.6 and 4.7. All that remains is to make a global choice of constants. The constant r from Proposition 4.7 was chosen to be valid in all regions; thus, we set $R = r$. It is sufficient to choose

$$m = \min \{m_*(1 - M_1), m_*(1 - M_2)\} < m_*$$

$$b = \sup \left\{|(\mathcal{L}V)(x,y)| : x^2 + y^2 \leqslant R^2\right\}$$

$$\gamma = \min\{\gamma_1, \gamma_2, \gamma_3\} = \frac{5\delta + 5}{5\delta + 3}.$$

These choices guarantee that for all $(x, y) \in \mathbb{R}^2$

$$(\mathcal{L}V)(x,y) \leqslant -m\,[V(x,y)]^\gamma + b\,.$$

□

5

Conclusions and Future Research

In this chapter, we summarize the contributions of this thesis and describe various avenues for possible future research.

5.1 Summary of Results

This thesis has developed a generalized Lyapunov construction for proving stabilization by noise. Chapter 1 described the mathematical setting under consideration and the notion of stability used throughout this thesis. It also gave an overview of the theory of Lyapunov functions. In Chapter 2, we described the systematic algorithm for the construction of a global Lyapunov function. The main idea was to build local Lyapunov functions as solutions to associated Poisson equations, and then patch them together to form one smooth global Lyapunov function on the plane. One of the crucial components to the construction of the local Lyapunov functions was the use of homogenous scaling transformations. It is our thought that these scaling techniques could also be applicable to a much broader setting.

Chapter 3 proved key properties of a particular type of Lyapunov function, which we referred to as a super Lyapunov function. We showed that the existence of a super Lyapunov function guarantees that the associated system converges to a unique invariant probability measure at a uniform, exponential rate. The results for super Lyapunov functions can easily be extended to the case of general Markov chains and are not just limited to the phenomenon of noise-induced stabilization. In Chapter 4, we applied the systematic algorithm outlined in Chapter 2 to construct a super Lyapunov function for a model problem. This model problem blows up in finite time in the deterministic setting, but is stabilized by by the addition of white noise in the y-direction. The model problems helped to more clearly illustrate the details of the Lyapunov construction algorithm. It is our hope that the algorithm could be applied to a wide array of problems in order to greatly reduce the amount of time and difficulty often necessitated in order to construct a Lyapunov function to prove stabilization.

In the following sections, we describe several possible areas for further research.

5.2 Additional Examples

First of all, we are interested in gathering additional examples of planar systems which exhibit noise-induced stabilization to which we can apply the algorithm outlined in Chapter 2 for the systematic construction of a Lyapunov function. The model problem considered in Chapter 4 corresponded to the complex-valued SDE

$$dZ_t^\epsilon = (Z_t^\epsilon)^2 dt + \epsilon_x dW_t^x + \epsilon_y dW_t^y,$$

while in [23], Herzog proved a complete characterization of the conditions required on the noise coefficients in order for the more general system

$$dZ_t^\epsilon = (Z_t^\epsilon)^n dt + \epsilon_x dW_t^x + \epsilon_y dW_t^y,$$

with $n \geqslant 2$ to be stable. When $n > 2$, the system has multiple diffusive regions, and we wish to apply our systematic algorithm to this more general system in order to fully test the algorithm's capabilities.

Another particular system which we would like to analyze is similar to that considered in [10], given by

$$\begin{aligned} dX_t^\epsilon &= (-X_t^\epsilon + (X_t^\epsilon)^2 Y_t^\epsilon - (X_t^\epsilon)^3)dt + \epsilon_x \, dW_t^x \\ dY_t^\epsilon &= (aY_t^\epsilon - (X_t^\epsilon)^3)dt + \epsilon_y \, dW_t^y \end{aligned} \tag{5.1}$$

where a is the deterministic bifurcation parameter. When $a > 0$, the deterministic system, (X_t, Y_t), corresponding to $\epsilon_x = \epsilon_y = 0$ is unstable. This is evidenced by the fact that when $X_0 = 0$,

$$X_t \equiv 0 \text{ and } Y_t = Y_0 e^{at} .$$

This in turn implies that when $X_0 = 0$,

$$\lim_{t \to \infty} |Y_t| = \infty .$$

In fact, for all initial conditions other than the three equilibrium points, the process approaches infinity. Since the system wanders off to infinity for certain initial conditions, it is unstable according to the definition of stability given in Section 1.2.

Figure 5.1 shows an example deterministic trajectory with $a = 1$. Note that the system is symmetric about the origin and contains three unstable equilibrium points: the origin and two other points out of which spiraling trajectories emanate. If the process starts in the right half plane, it wanders off to infinity in the negative y-direction; similarly, if the process starts in the left half plane, it wanders off to infinity in the positive y-direction.

The instability of this system differs in many respects from the instability of the model problem analyzed in Chapter 4. First of all, the model problem blows up in finite time for certain initial conditions, while this system only blows up as $t \to \infty$. Secondly, the instability in the model problem is concentrated solely on the positive x-axis, and for all other initial conditions, the process flows in towards the origin. The addition of noise is able to stabilize the model problem by kicking the process

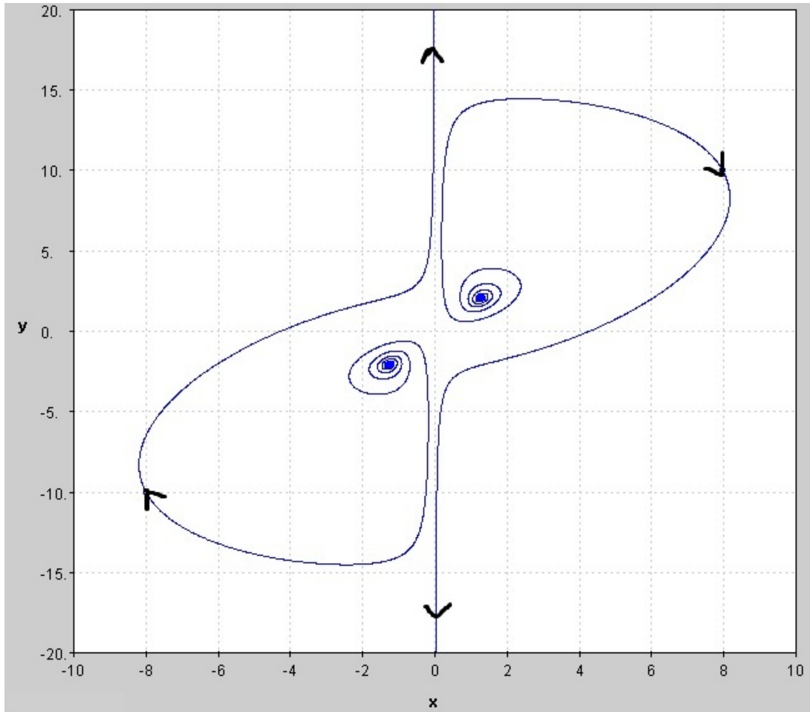

FIGURE 5.1: Example deterministic trajectory for the system given by (5.1) with $a = 1$.

away from the axis and directing the flow inward. However, in the system defined by (5.1), the noise stabilizes the system by connecting the two unstable trajectories in opposite halves of the plane to form a quasi-periodic orbit. A simulation of the system is shown in Figure 5.2.

While it is clear from the simulation that the process defined by (5.1) exhibits noise-induced stabilization, the systematic algorithm outlined in Chapter 2 could be utilized in order to rigorously prove that the system does indeed exhibit the stabilization phenomenon. Moreover, the algorithm could be applied to a wide variety of other systems in order to construct a global Lyapunov function to prove stabilization by noise.

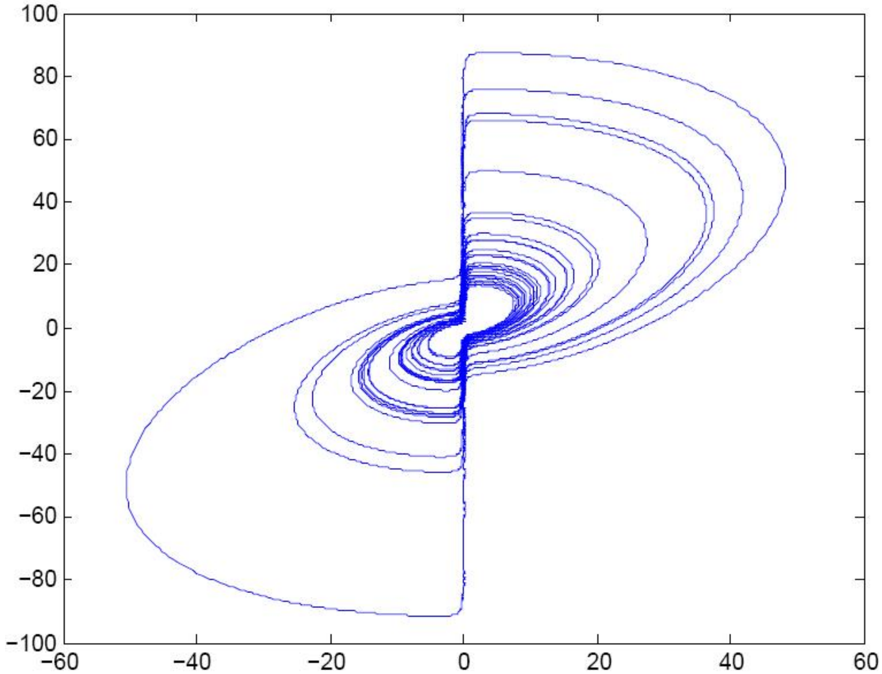

FIGURE 5.2: A simulation of the system given by (5.1) with $a = 1$.

5.3 Sufficient Criteria for Stabilization

It is often not clear a priori whether or not an unstable deterministic system will be able to be stabilized by the introduction of noise into the system. Hence, we would like to utilize the generalized algorithm outlined in Chapter 2 in order to produce at least sufficient criteria for guaranteeing when this phenomenon will occur. Due to the complexity of the stabilization phenomenon, we are doubtful that a complete characterization of the necessary and sufficient criteria for stabilization could be identified. Nevertheless, our hope is that the scaling techniques described in our algorithm could help elucidate properties of the generator that would ensure stabilization.

Below, we describe a particular example of a system that does *not* exhibit noise-induced stabilization, but which appears to on a superficial level, as a cautionary

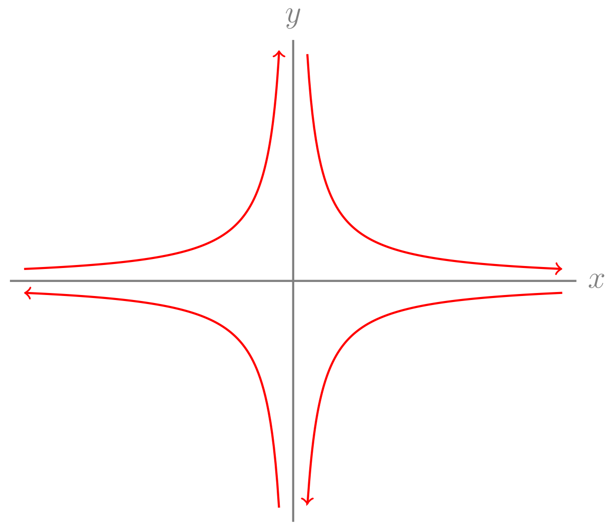

FIGURE 5.3: Example deterministic trajectories for the system given by (5.2).

tale of the delicacy of this issue. Consider the system

$$\begin{aligned} dX_t^\epsilon &= (X_t^\epsilon)^2 Y_t^\epsilon dt + \epsilon_x \, dW_t^x \\ dY_t^\epsilon &= -X_t^\epsilon (Y_t^\epsilon)^2 dt + \epsilon_y \, dW_t^y \end{aligned} \tag{5.2}$$

whose deterministic trajectories are illustrated in Figure 5.3. The figure illustrates that in the deterministic setting corresponding to $\epsilon_x = \epsilon_y = 0$, the process wanders off to infinity for all initial conditions with $X_0, Y_0 \neq 0$. Both axes consist of a continuum of equilibrium points. If the process starts off of the axes, then whichever quadrant the process begin in, it remains in that quadrant for all time.

One's first guess might be that the addition of noise in both directions would induce stability by allowing the process to jump over the axes, connecting the four unstable deterministic trajectories and forming a quasi-periodic orbit. However, (5.2) is actually a Hamiltonian system with the Hamiltonian function $H(x, y) = \frac{1}{2}x^2y^2$, i.e.,

$$\begin{aligned} dX_t^\epsilon &= \frac{\partial H(X_t^\epsilon, Y_t^\epsilon)}{\partial y} dt + \epsilon_x \, dW_t^x \\ dY_t^\epsilon &= -\frac{\partial H(X_t^\epsilon, Y_t^\epsilon)}{\partial x} dt + \epsilon_y \, dW_t^y \, . \end{aligned}$$

By direct computation of the adjoint of the associated generator, it can be shown that Lebesgue measure is invariant for any Hamiltonian system of this form. Thus, the system does not exhibit noise-induced stabilization. While the addition of noise does allow the process to traverse all four quadrants rather than zooming off to infinity along one particular axis within the initial quadrant, the perturbed system wanders off to infinity by spiraling outward. The location where the process hits the axes as it traverses the four quadrants is a random variable that varies over time, and its distribution does not decay sufficiently quickly.

The system given by (5.2) has many similarities to the system given by (5.1) and we would like to more precisely pinpoint why the latter exhibits noise-induced stabilization while the former does not. Both systems wander off to infinity in the deterministic setting and, in both cases, noise allows the processes to traverse between regions of deterministic instability. Yet in one, the noise creates a quasi-periodic behavior, but for the other, the system still wanders off to infinity. Our hope is to develop criteria to distinguish between the two cases. We are also interested in analyzing what minimal changes could be made to (5.2) in order to preserve the deterministic phase portrait, but allow for the addition of noise to have a stabilizing effect.

Since it is often of interest to determine the minimum amount of noise necessary for stabilization, we also wish to utilize the systematic algorithm in order to determine the minimum restriction on the values of ϵ_x and ϵ_y in order for a system which exhibits noise-induced stabilization to be stabilized. This minimum restriction would certainly depend upon the particular example under consideration, but the key idea would be to analyze the diffusion terms in the generator in order to determine which ones are dominant under certain scalings.

It is often helpful to simulate the solution to an SDE in order to gain intuition into its behavior. The simulations can help determine whether or not a particular system exhibits noise-induced stabilization before attempting to construct a Lyapunov function. The simulations can also aid in pinpointing where the critical diffusive regions are located and possible restrictions on the noise coefficients. In addition, for systems exhibiting noise-induced stabilization, simulations can provide insight into the properties of the invariant probability measure. The model problem presented in Chapter 4, as well as the example system defined by (5.1) and described in Section 5.2, were simulated in Matlab using the basic Forward Euler-Maruyama method. We would like to develop more sophisticated simulation techniques for these types of problems.

5.4 Higher Dimensions and Multiplicative Noise

The systematic algorithm outlined in Chapter 2 was presented exclusively for planar systems. However, we wish to extend the algorithm to three-dimensional (or even higher dimensional) systems where the regions of deterministic instability may be more difficult to characterize. This extension would not be elementary since Brownian motion has different properties in three and higher dimensional systems and, heuristically, there are more possible routes to infinity.

It is interesting to note that in one dimension it is impossible for additive noise to stabilize an unstable deterministic system. This can easily be shown using one-dimensional Feller theory [22]. Moreover, Scheutzow proved in [32] that a perturbation by additive white noise of any one-dimensional deterministic system that blows up in finite time results in a perturbed system that blows up in finite time with strictly positive probability. However, multiplicative noise can indeed stabilize an unstable one-dimensional system, which is evident in the classic example of geometric Brownian motion: if $\mathbb{X}_t^\epsilon \in \mathbb{R}$ is the solution to the SDE

$$d\mathbb{X}_t^\epsilon = a\mathbb{X}_t^\epsilon + \epsilon\mathbb{X}_t^\epsilon dW_t$$

where $\epsilon \geqslant 0$ and W_t is one-dimensional Brownian motion, then $\mathbb{X}_t^\epsilon$ is called *geometric Brownian motion.* If $a > 0$ and $\epsilon = 0$, $\mathbb{X}_t^\epsilon$ is unbounded for any $\mathbb{X}_0^\epsilon \neq 0$, and hence unstable. It can be shown using the law of the iterated logarithm (see [30]) that if $\epsilon^2 > 2a$, then $\mathbb{X}_t^\epsilon$ is not only stable, but in fact

$$\lim_{t\to\infty} \mathbb{X}_t^\epsilon = 0 \quad \text{with probability one}$$

for any initial condition.

Our algorithm, which was presented with only additive noise, could easily be extended to the case of multiplicative noise in two-dimensional systems. Thus, we could extend the mathematical setting for $\mathbb{X}_t^\epsilon$ described in Section 1.1 to SDEs of the form

$$d\mathbb{X}_t^\epsilon = \varphi(\mathbb{X}_t^\epsilon)dt + \sigma(\mathbb{X}_t^\epsilon)\ \epsilon\ dW_t \tag{5.3}$$

where $\sigma \in C^2(\mathbb{R}^2 \to \mathbb{R}^2 \times \mathbb{R}^2)$.

The phenomenon of noise-induced stabilization is a very rich and intriguing area of study, and it is hoped that the contributions made in this thesis, such as the systematic Lyapunov construction, scaling techniques, and super Lyapunov properties, will aid relevant applications and inspire further research in this area.

Appendix A

Model Problem in Polar Coordinates

We analyze the complex-valued SDE

$$dZ_t = Z_t^2 dt + \epsilon dW_t \tag{A.1}$$

where $Z_t = X_t + iY_t \in \mathbb{C}$ and $W_t = W_t^x + iW_t^y$, with W_t^x and W_t^y independent standard one-dimensional Brownian motions. Written in Cartesian coordinates, this SDE is equivalent to the model problem presented in Chapter 4 given by (4.2) with $\epsilon = \epsilon_x = \epsilon_y$. We now consider the SDE written in polar coordinates; hence, we analyze the system (R_t, θ_t) where $Z_t = R_t e^{i\theta_t}$, $R_t \geqslant 0$ for all t, and $\theta_0 \in (-\pi, \pi]$. We may also write R_t and θ_t in terms of Z_t and $\bar{Z}_t = R_t e^{-i\theta_t}$:

$$R_t = \sqrt{Z_t \bar{Z}_t} \text{ and } \theta_t = \frac{1}{2i} \ln\left(\frac{Z_t}{\bar{Z}_t}\right).$$

Note that $\bar{Z}_t$ satisfies the SDE:

$$d\bar{Z}_t = \bar{Z}_t^2 dt + \epsilon d\bar{W}_t\,.$$

The polar coordinates help elucidate many properties of the dynamics since the deterministic trajectories are circular. In particular, the polar coordinates enable us to easily compute an asymptotic expansion of the invariant probability density, which we calculate in Section A.3. We first analyze the deterministic system, corresponding to $\epsilon = 0$, in polar coordinates.

A.1 Deterministic System in Polar Coordinates

Setting $\epsilon = 0$, we can explicitly solve for the deterministic solution in polar coordinates. Using the fact that

$$Z_t = \frac{Z_0}{1 - Z_0 t} \text{ and } \bar{Z}_t = \frac{\bar{Z}_0}{1 - \bar{Z}_0 t}$$

we obtain that R_t satisfies

$$\begin{aligned} R_t &= \sqrt{Z_t \bar{Z}_t} \\ &= \sqrt{\left(\frac{R_0 e^{i\theta_0}}{1 - R_0 e^{i\theta_0} t}\right)\left(\frac{R_0 e^{-i\theta_0}}{1 - R_0 e^{-i\theta_0} t}\right)} \\ &= \frac{R_0}{\sqrt{R_0^2 t^2 - 2R_0 \cos(\theta_0) t + 1}} \end{aligned}$$

and that θ_t satisfies

$$\begin{aligned} \theta_t &= \frac{1}{2i} \ln\left(\frac{Z_t}{\bar{Z}_t}\right) \\ &= \frac{1}{2i}\left(\ln(e^{2i\theta_0}) + \ln\left(\frac{1 - R_0 e^{-i\theta_0} t}{1 - R_0 e^{i\theta_0} t}\right)\right) \\ &= \theta_0 + \frac{1}{2i} \ln\left(\frac{1 - R_0 \cos(\theta_0) t + iR_0 \sin(\theta_0) t}{1 - R_0 \cos(\theta_0) t - iR_0 \sin(\theta_0)}\right) \\ &= \theta_0 + \arctan\left(\frac{R_0 \sin(\theta_0) t}{1 - R_0 \cos(\theta_0) t}\right). \end{aligned}$$

Note that we take the range of arctan to be $[0, \pi)$ if $\theta_0 \geqslant 0$ and we take the range to be $(-\pi, 0]$ if $\theta < 0$. We observe that if $R_0 = 0$, then $R_t \equiv 0$ for all t; hence, the origin is an equilibrium point. We also observe that if $\theta_0 \neq 0$, then we obtain the following relationship between R_t and θ_t:

$$R_t = \left(\frac{R_0}{\sin(\theta_0)}\right)\sin(\theta_t)\,.$$

Hence, the process follows a circle with

$$\text{radius } = \frac{R_0}{2|\sin(\theta_0)|} \quad \text{and center } = \left(0, \frac{R_0}{2\sin(\theta_0)}\right)\,.$$

In this case, we also note that the following limits show that the process approaches the origin, but does not reach it in finite time:

$$\lim_{t\to\infty}\theta_t = \theta_0 + \tan^{-1}(\tan(-\theta_0)) = \theta_0 + (\mathrm{sgn}(\theta_0)\pi - \theta_0) = \mathrm{sgn}(\theta_0)\pi \quad \text{and}$$

$$\lim_{t\to\infty} R_t = \lim_{t\to\infty}\frac{R_0}{\sqrt{R_0^2t^2 - 2R_0\cos(\theta_0)t + 1}} = 0\,.$$

The instability of the deterministic system occurs when $\theta_0 = 0$, evidenced by the fact that $\theta_0 = 0$ implies $\theta_t \equiv 0$ for all t and

$$R_t = \frac{R_0}{1 - R_0 t}\,.$$

Hence R_t blows up in finite time, where the blow up time is $t = \frac{1}{R_0}$. We now consider the perturbed system in polar coordinates.

A.2 Perturbed System in Polar Coordinates

We use Ito's formula to rewrite the complex-valued SDE given by (A.1) in polar coordinates. We first note the following quadratic variations and quadratic covariation of W_t and $\bar{W}_t$:

$$\langle W_t, W_t\rangle = \langle W_t^x + iW_t^y, W_t^x + iW_t^y\rangle = t - t = 0$$

$$\langle \bar{W}_t, \bar{W}_t\rangle = \langle W_t^x - iW_t^y, W_t^x - iW_t^y\rangle = t - t = 0$$

$$\langle W_t, \bar{W}_t\rangle = \langle W_t^x + iW_t^y, W_t^x - iW_t^y\rangle = t + t = 2\,.$$

Note that if we considered $\epsilon_x W_t^x + i\epsilon_y W_t^x$ rather than ϵW_t, the quadratic variations and quadratic covariation would not exhibit the simplification found in the above calculations. Hence for simplicity, we consider noise added equally in both directions in the polar setting. We now apply Ito's formula to

$$R_t = f(Z_t, \bar{Z}_t) \text{ where } f(x,y) = \sqrt{xy}\,.$$

Thus we obtain that

$$\begin{aligned} dR_t &= \frac{1}{2}\sqrt{\frac{\bar{Z}_t}{Z_t}}dZ_t + \frac{1}{2}\sqrt{\frac{Z_t}{\bar{Z}_t}}d\bar{Z}_t + \frac{1}{4\sqrt{Z_t\bar{Z}_t}}(dZ_t)(d\bar{Z}_t) \\ &= \frac{e^{-i\theta_t}}{2}(R_t^2 e^{2i\theta_t}dt + \epsilon dW_t) + \frac{e^{i\theta_t}}{2}(R_t^2 e^{-2i\theta_t}dt + \epsilon d\bar{W}_t) + \frac{\epsilon^2}{2R_t}dt \\ &= R_t^2\left(\frac{e^{i\theta_t} + e^{-i\theta_t}}{2}\right)dt + \frac{\epsilon^2}{2R_t}dt + \frac{\epsilon}{2}(e^{-i\theta_t}dW_t + e^{i\theta_t}d\bar{W}_t) \\ &= R_t^2\cos(\theta_t)dt + \frac{\epsilon^2}{2R_t}dt + \epsilon(\cos(\theta_t)dW_t^x + \sin(\theta_t)dW_t^y)\,. \end{aligned}$$

We next apply Ito's formula to

$$\theta_t = g(Z_t, \bar{Z}_t) \text{ where } g(x,y) = \frac{1}{2i}\ln\left(\frac{x}{y}\right).$$

Thus we obtain that

$$\begin{aligned} d\theta_t &= \frac{1}{2iZ_t}dZ_t - \frac{1}{2i\bar{Z}_t}d\bar{Z}_t \\ &= \frac{e^{-i\theta_t}}{2iR_t}(R_t^2 e^{2i\theta_t}dt + \epsilon dW_t) - \frac{e^{i\theta_t}}{2iR_t}(R_t^2 e^{-2i\theta_t}dt + \epsilon d\bar{W}_t) \\ &= R_t\left(\frac{e^{i\theta_t} - e^{-i\theta_t}}{2i}\right)dt + \epsilon\left(\frac{e^{-i\theta_t}}{2iR_t}dW_t - \frac{e^{i\theta_t}}{2iR_t}d\bar{W}_t\right) \\ &= R_t\sin(\theta_t)dt + \frac{\epsilon}{R_t}\left(-\left(\frac{e^{i\theta_t} - e^{-i\theta_t}}{2i}\right)dW_t^x + \left(\frac{e^{i\theta_t} + e^{-i\theta_t}}{2}\right)dW_t^y\right) \\ &= R_t\sin(\theta_t)dt + \frac{\epsilon}{R_t}(-\sin(\theta_t)dW_t^x + \cos(\theta_t)dW_t^y)\,. \end{aligned}$$

We may simplify the SDEs for R_t and θ_t by defining B_t^R and B_t^θ from the following equations

$$
\begin{aligned}
dB_t^R &= \cos(\theta_t) dW_t^x + \sin(\theta_t) dW_t^y \\
dB_t^\theta &= -\sin(\theta_t) dW_t^x + \cos(\theta_t) dW_t^y \,.
\end{aligned}
$$

By [30], B_t^R and B_t^θ are independent standard one-dimensional Brownian motions. Hence, we may rewrite the SDEs for R_t and θ_t as

$$
\begin{aligned}
dR_t &= R_t^2 \cos(\theta_t) dt + \frac{\epsilon^2}{2R_t} dt + \epsilon dB_t^R \\
d\theta_t &= R_t \sin(\theta_t) dt + \frac{\epsilon}{R_t} dB_t^\theta \,.
\end{aligned}
\tag{A.2}
$$

A.3 Asymptotic Expansion of Invariant Density

The generator, $\mathcal{L}$, for the two-dimensional SDE defined by (A.2), in polar coordinates, is given by:

$$
(\mathcal{L}\rho)(r,\theta) = r^2 \cos\theta \frac{\partial \rho(r,\theta)}{\partial r} + r \sin\theta \frac{\partial \rho(r,\theta)}{\partial \theta} + \kappa[\Delta\rho](r,\theta)
$$

where $\kappa = \frac{\epsilon^2}{2}$ and Δ is the polar Laplacian, i.e.,

$$
(\Delta\rho)(r,\theta) = \frac{1}{r}\frac{\partial \rho(r,\theta)}{\partial r} + \frac{\partial^2 \rho(r,\theta)}{\partial r^2} + \frac{1}{r^2}\frac{\partial^2 \rho(r,\theta)}{\partial \theta^2} \,.
$$

We consider the adjoint operator, $\mathcal{L}^*$, given by:

$$
\begin{aligned}
(\mathcal{L}^*\rho)(r,\theta) &= -\frac{\cos\theta}{r}\frac{\partial}{\partial r}[r^3 \rho(r,\theta)] - r\frac{\partial}{\partial \theta}[\sin\theta \rho(r,\theta)] + \kappa[\Delta\rho](r,\theta) \\
&= -\left[4r\cos\theta \rho(r,\theta) + r^2 \cos\theta \frac{\partial \rho(r,\theta)}{\partial r} + r \sin\theta \frac{\partial \rho(r,\theta)}{\partial \theta}\right] + \kappa[\Delta\rho](r,\theta) \,.
\end{aligned}
$$

It is well known (see [30]) that the invariant probability density, $\rho(r,\theta)$, for (A.2) satisfies the PDE

$$
(\mathcal{L}^*\rho)(r,\theta) = 0 \,.
$$

We now calculate an asymptotic expansion for $\rho(r,\theta)$ of the form

$$\rho(r,\theta) = \sum_{n=0}^{\infty} \frac{1}{r^{n+3}}\rho_n(\theta) \tag{A.3}$$

to consider the asymptotic decay of the invariant probability density. Note that in order for $\rho(r,\theta)$ to be a valid probability density, we must have

$$\int_0^{2\pi}\int_0^{\infty} \rho(r,\theta) r dr d\theta = 1\,.$$

Our expansion given in (A.3) is not integrable at zero; however, the expansion will still be useful for analyzing the decay of the invariant probability density for large radii. Note that we start the expansion with $\frac{1}{r^3}$ because in order for the density to be integrable at infinity, it must decay faster than $\frac{1}{r^2}$. We define $\mathcal{L}_0^*$ by:

$$(\mathcal{L}_0^*\rho)(r,\theta) = 4r\cos\theta\rho(r,\theta) + r^2\cos\theta\frac{\partial\rho(r,\theta)}{\partial r} + r\sin\theta\frac{\partial\rho(r,\theta)}{\partial\theta}\,.$$

Then $\mathcal{L}^* = -\mathcal{L}_0^* + \kappa\Delta$. Applying $\mathcal{L}_0^*$ and Δ to the terms in the asymptotic expansion given by (A.3), we observe that

$$\begin{aligned}
\mathcal{L}_0^*\left[\frac{1}{r^{n+3}}\rho_n(\theta)\right] &= \frac{1}{r^{n+2}}((1-n)\cos\theta\rho_n(\theta) + \sin\theta\rho_n'(\theta)) \text{ and} \\
\Delta\left[\frac{1}{r^{n+3}}\rho_n(\theta)\right] &= \frac{1}{r^{n+5}}((n+3)^2\rho_n(\theta) + \rho_n''(\theta))\,.
\end{aligned}$$

Thus, in order for

$$\mathcal{L}^*\left[\sum_{n=0}^{\infty}\frac{1}{r^{n+3}}\rho_n(\theta)\right] = 0,$$

we must have that

$$\begin{aligned}
&(1-n)\cos\theta\rho_n(\theta) + \sin\theta\rho_n'(\theta) = 0 \quad \text{for } n = 0,1,2 \\
&(1-n)\cos\theta\rho_n(\theta) + \sin\theta\rho_n'(\theta) = n^2\rho_{n-3}(\theta) + \rho_{n-3}''(\theta) \quad \text{for } n \geqslant 3\,.
\end{aligned}$$

The general solution to the ODE

$$(1-n)\cos\theta\rho_n(\theta)+\sin\theta\rho_n'(\theta)=0$$

is given by

$$\rho_n(\theta)=c_n\sin^{(n-1)}\theta$$

where c_n is a constant. Now we know a priori that $\rho_n(\theta)$ must be an even function due to the symmetry properties of the dynamics. Hence, $c_n = 0$ for all even n. Therefore, the leading term in the asymptotic expansion is $\frac{c_1}{r^4}$, which indicates the decay rate of the invariant probability density for large radii.

Appendix B

Model Problem with Degenerate Noise

In this appendix, we consider the model problem described in Chapter 4 given by (4.2) with degenerate noise in the x-direction; i.e., we consider $\epsilon_x = 0$ and $\epsilon_y > 0$. In Section B.1 we characterize the positivity of the density function for the model problem in this case. The positivity result fulfills the assumptions required in Section 3.4.2 in order to guarantee that a minorization condition holds for the model problem. Hence, the positivity result, along with the existence of a super Lyapunov function proven in Chapter 4, completes the proof that the model problem converges to its invariant probability measure at a uniform, exponential rate. In Section B.2, we use the properties of the density to characterize the positivity of the invariant probability measure.

B.1 Positivity of Density

From Section 3.4.2, we know that a continuous density function, $p_t(\mathbb{x}, \mathbb{x}')$, exists for the model problem in the case of degenerate noise with $\epsilon_x = 0$ and $\epsilon_y > 0$. However, unlike the nondegenerate noise setting, it is no longer true that $p_t(\mathbb{x}, \mathbb{x}') > 0$ for all

$t > 0$ and $\mathbb{x}, \mathbb{x}' \in \mathbb{R}^2$. Below, we state and prove a theorem which characterizes the positivity of the density in the case of degenerate noise.

Theorem B.1. *If $\epsilon_x = 0$ and $\epsilon_y > 0$, then for any $\mathbb{x}_* = (x_*, y_*) \in \mathbb{R}^2$ with $x_* < 0$, there exists T_* so that $p_t(\mathbb{x}, \mathbb{x}_*) > 0$ for all $t \geqslant T_*$ and for all $\mathbb{x} \in \mathbb{R}^2$. Furthermore, if $\mathbb{x}_0 = (x_0, y_0) \in \mathbb{R}^2$ with $x_0 < 0$ and $\mathbb{x}_1 = (x_1, y_1) \in \mathbb{R}^2$ with $x_1 \geqslant 0$, then $p_t(\mathbb{x}_0, \mathbb{x}_1) = 0$ for all $t > 0$.*

The proof of this theorem relies heavily on results from Malliavin calculus and control theory, as described in [4, 6, 7]. We give an overview of these results and how they apply to our setting in the following section. Then in Section B.1.2 we state and prove two intermediary lemmas before finally proving Theorem B.1 in Section B.1.3.

B.1.1 Results from Malliavin Calculus and Control Theory

For the process $\mathbb{X}_t^\epsilon = (X_t^\epsilon, Y_t^\epsilon)$ satisfying (4.2) with $\epsilon_x = 0$ and $\epsilon_y > 0$, i.e.,

$$\begin{aligned} dX_t^\epsilon &= ((X_t^\epsilon)^2 - (Y_t^\epsilon)^2)dt \\ dY_t^\epsilon &= 2X_t^\epsilon Y_t^\epsilon dt + \epsilon_y dW_t^y \end{aligned}$$

and any $U_t \in L^2(\mathbb{R})$, we can associate the process $\mathbb{X}_t^U = (X_t^U, Y_t^U)$ which solves the following ODE:

$$\begin{aligned} dX_t^U &= ((X_t^U)^2 - (Y_t^U)^2)dt \\ dY_t^U &= 2X_t^U Y_t^U dt + U_t\,. \end{aligned} \tag{B.1}$$

Note that $\mathbb{X}_t^U$ is a continuous function of t. We will refer to U_t as the control and denote by $\mathbb{X}_t$ the solution to the deterministic system corresponding to either $\epsilon_y = 0$ or $U_t = 0$. For $T > 0$ and $\mathbb{x} \in \mathbb{R}^2$, we consider the function $\Phi_{T,\mathbb{x}}$ defined by

$$\Phi_{T,\mathbb{x}}(U) = (X_T^U, Y_T^U) \quad \text{where} \quad (X_0^U, Y_0^U) = \mathbb{x}\,.$$

Translating [4] to our current setting, we obtain the following proposition.

Proposition B.1. $p_T(\mathbb{x}, \mathbb{x}') > 0$ *if and only if there exists a control* $U \in L^2([0,T], \mathbb{R})$ *so that* $\Phi_{T,\mathbb{x}}(U) = (\Phi^{(1)}_{T,\mathbb{x}}(U), \Phi^{(2)}_{T,\mathbb{x}}(U)) = \mathbb{x}'$ *and*

$$M_{T,\mathbb{x}}(U) = \begin{pmatrix} \|D\Phi^{(1)}_{T,\mathbb{x}}(U)\|^2_{L^2} & \langle D\Phi^{(1)}_{T,\mathbb{x}}(U), D\Phi^{(2)}_{T,\mathbb{x}}(U)\rangle_{L^2} \\ \langle D\Phi^{(1)}_{T,\mathbb{x}}(U), D\Phi^{(2)}_{T,\mathbb{x}}(U)\rangle_{L^2} & \|D\Phi^{(2)}_{T,\mathbb{x}}(U)\|^2_{L^2} \end{pmatrix} \tag{B.2}$$

is nondegenerate, where D *represents the Frechet derivative.*

It is straightforward to deduce that the matrix $M_{T,z}$ is nondegenerate if there exists a time $t_0 \in [0,T]$ so that $\Phi_{t_0,\mathbb{x}}(U) \neq 0$. Thus, as a consequence of this proposition, to prove the positivity of $p_t(\mathbb{x}, \mathbb{x}')$ for two given points $\mathbb{x}, \mathbb{x}' \in \mathbb{R}^2$, we simply need to find a control U for which $\Phi_{t,\mathbb{x}}(U) = \mathbb{x}'$ and for which the path does not spend all of its time at the origin. Since the path is continuous in time, this last condition poses no restriction if either $\mathbb{x}$ or $\mathbb{x}'$ is not the origin. If both $\mathbb{x}$ and $\mathbb{x}'$ are the origin, it is still elementary to find a control satisfying the second condition which still also satisfies $\Phi_{t,\mathbb{x}}(U) = \mathbb{x}'$.

We construct the required control functions through the two lemmas in Section B.1.2. In the first lemma, we construct a control that maps any point $\mathbb{x} \in \mathbb{R}^2$ to any point in the left-half plane of the form $(-a, 0)$ for some $a > 0$. In the second lemma, we construct a control that can map a point of the form $(-a, 0)$ with $a > 0$ to a point in the left-half plane. In Section B.1.3, we complete the proof of Theorem B.1 by combining the two lemmas to form a control that can map any point $\mathbb{x} \in \mathbb{R}^2$ to any point $\mathbb{x}'$ in the left-half plane. We prove that the control constructed has the desired properties.

All the controls we design will take the form $U_t = u(X^U_t, Y^U_t, t)$ for some piecewise smooth $u : \mathbb{R}^2 \times [0,T] \to \mathbb{R}$. At first glance, this might seem ill-defined since (X^U_t, Y^U_t) depends on the function U_t. However, (X^U_t, Y^U_t) can be defined in a self-contained way as the solution to the ODE

$$\begin{aligned} dX_t^U &= ((X_t^U)^2 - (Y_t^U)^2)dt \\ dY_t^U &= 2X_t^U Y_t^U dt + u(X_t^U, Y_t^U, t)\,. \end{aligned}$$

Then, with this solution in hand, one can define $U_t = u(X_t^U, Y_t^U, t)$.

B.1.2 Intermediary Lemmas

We now state and prove the two intermediary lemmas which are used to construct the control required in Theorem B.1.

Lemma B.1. *For any $a > 0$, there exists a positive $T_* = T_*(a)$ such that for all $\mathbb{x}_0 \in \mathbb{R}^2$ and for all $T > T_*$, there exists a control $U \in L^2([0,T],\mathbb{R})$ for which $\Phi_{T,\mathbb{x}_0}(U) = (-a, 0)$, i.e.,*

$$(X_0^U, Y_0^U) = \mathbb{x}_0 \quad \textit{and} \quad (X_T^U, Y_T^U) = (-a, 0)\,.$$

Proof of Lemma B.1. Fix $a > 0$ and $\mathbb{x}_0 = (x_0, y_0)$ where $y_0 \geqslant 0$. By symmetry, the case $y_0 < 0$ follows analogously. As discussed above, we implicitly construct $U_t = u(X_t^U, Y_t^U, t)$ where $u\colon \mathbb{R}^2 \times [0,T] \to \mathbb{R}$ is an explicit function which is continuous and piecewise smooth. This in turn implies that the resulting U_t is a continuous, piece-wise smooth control which produces a path from $\mathbb{x}_0$ at time 0 to $(-a, 0)$ at time T.

We construct the control to produce a path that passes through two intermediate points $\mathbb{x}_1 = (x_1, y_1)$ and $\mathbb{x}_2 = (x_2, y_2)$, which will be specified below. The time to traverse this path from $\mathbb{x}_0$ to $\mathbb{x}_1$ is denoted T_0, the time to traverse from $\mathbb{x}_1$ to $\mathbb{x}_2$ is denoted T_1, and the time to traverse from $\mathbb{x}_2$ to $(-a, 0)$ is denoted T_2. Thus, $T = T_0 + T_1 + T_2$.

We will see that once the endpoints $\mathbb{x}_0 = (x_0, y_0)$ and $(-a, 0)$ are fixed, the only free parameter in the construction will be y_1. Once it is fixed, all other variables needed to define the intermediate points, namely x_1, x_2, and y_2, and the time necessary to traverse the path, are completely determined. Hence we view $T = T_0+T_1+T_2$ as a function of y_1. It will be clear from the construction that T is a continuous function of y_1; hence by the intermediate value theorem, any time between $\min_{y_1}\{T(y_1)\}$ and $\max_{y_1}\{T(y_1)\}$ can be achieved. We estimate the minimum and show that the maximum is infinity.

We start by constructing our trajectory. As already mentioned, we utilize a control $U_t = u(X_t^U, Y_t^U, t)$ where

$$u(x,y,t) = \begin{cases} u_0(x,y) & \text{for } t \in [0,T_0] \\ u_1(x,y) & \text{for } t \in (T_0, T_0+T_1] \\ u_2(x,y) & \text{for } t \in (T_0+T_1, T] \end{cases}$$

with the $u_i(x,y)$ as specified below.

Path from $\mathbb{x}_0$ *to* $\mathbb{x}_1$: Let $u_0(x,y) = -2xy+1$. Then for $t \in [0,T_0]$, $dY_t^U = dt$, which implies that $Y_t^U = t + y_0$. For any $y_1 > y_0$, we choose $T_0 = y_1 - y_0$ so that $Y_{T_0} = y_1$. Thus T_0 can be chosen arbitrarily large based upon the choice of y_1 and is clearly a continuous function of y_1. Note that the value of x_1 is uniquely determined by $\mathbb{x}_0$ and the choice of y_1.

Path from $\mathbb{x}_1$ *to* $\mathbb{x}_2$: Once $\mathbb{x}_1$ is determined, $\mathbb{x}_2$ is chosen to be the unique point at the intersection between the deterministic trajectory emanating from $\mathbb{x}_1$ and the ball of radius $2a$ about the origin, in the upper left-half plane. Since we want the trajectory to follow the deterministic curve, we set $u_1(x,y) \equiv 0$. The coordinates of $\mathbb{x}_2$ can be calculated explicitly in terms of the coordinates of $\mathbb{x}_1$:

$$x_2 = -\sqrt{4a^2 - y_2^2} \text{ and } y_2 = \frac{4a^2 y_1}{x_1^2 + y_1^2}. \tag{B.3}$$

In Proposition 4.1, we proved that under the deterministic dynamics, the time to reach a ball of radius R about the origin is uniformly bounded by $\frac{2}{R}$, independent of the initial condition. Thus, $T_1 \leqslant \frac{1}{a}$ for all choices of $\mathbb{x}_0 \in \mathbb{R}^2$ and y_1, which as already discussed completely determines $\mathbb{x}_1 \in \mathbb{R}^2$. In addition, T_1 is a continuous function of y_1, since the trajectories deform smoothly as y_1 is varied and do not coincide at any point with a trajectory requiring infinite time to travel between $\mathbb{x}_1$ and $\mathbb{x}_2$.

It will be useful in what follows to have bounds on x_2 and y_2. Hence we further restrict our choice of y_1 by assuming that $y_1 \geqslant \max(y_0, 8a)$. In light of (B.3), this choice implies that $y_2 \in (0, \frac{a}{2}]$ and $-2a < x_2 < -\frac{\sqrt{15}a}{2}$.

Path from $\mathbb{x}_2$ *to* $(-a, 0)$: We use a linear path to connect $\mathbb{x}_2$ and $(-a, 0)$. More specifically, we set $m = \frac{y_2}{a+x_2} < 0$ and $u_2(x, y) = -2xy + m(x^2 - y^2)$. Then for $t \in (T_0 + T_1, T]$, $dY_t^U = mX_t^U dt$, and hence $Y_t^U = m(X_t^U - x_2) + y_2$. Since over this time interval, (X_t^U, Y_t^U) is moving linearly from a point on the circle of radius $2a$ to a point on the circle of radius a, we know that $X_t^U \in [-2a, -a]$ and $Y_t^U \in [0, a/2]$ for all $t \in [T_0 + T_1, T]$. From this, a bound on T_2 is easily obtained: along this path,

$$\dot{X}_t^U = (X_t^U)^2 - (Y_t^U)^2 \geqslant a^2 - \frac{a^2}{4} = \frac{3a^2}{4}.$$

Thus $X_t^U \geqslant \frac{3a^2}{4}[t - (T_0 + T_1)] + x_2$. Hence, $-a = X_T^U \geqslant \frac{3a^2}{4}T_2 + x_2$. When combined with the previous fact that $|x_2| \leqslant 2a$, this estimate in turn implies that

$$T_2 \leqslant \frac{4}{3a^2}(|x_2| - a) \leqslant \frac{4}{3a}$$

for all $\mathbb{x}_0 \in \mathbb{R}^2$ and all choices of y_1. With this bound in hand, it is clear that T_2 depends continuously on y_1 since T_2 depends continuously on the point $\mathbb{x}_2$, which depends continuously on T_1 and $\mathbb{x}_1$, both of which depend continuously on y_1.

In summary, we have constructed a family of controls U, depending on a parameter y_1, which form a path from $\mathbb{x}_0$ to $(-a, 0)$ in time T, where T is a continuous function of y_1. Figure B.1 gives an illustration of an example path. Since $T = T_0 + T_1 + T_2$ and T_0 increases to infinity as y_1 increases to infinity, we are as-

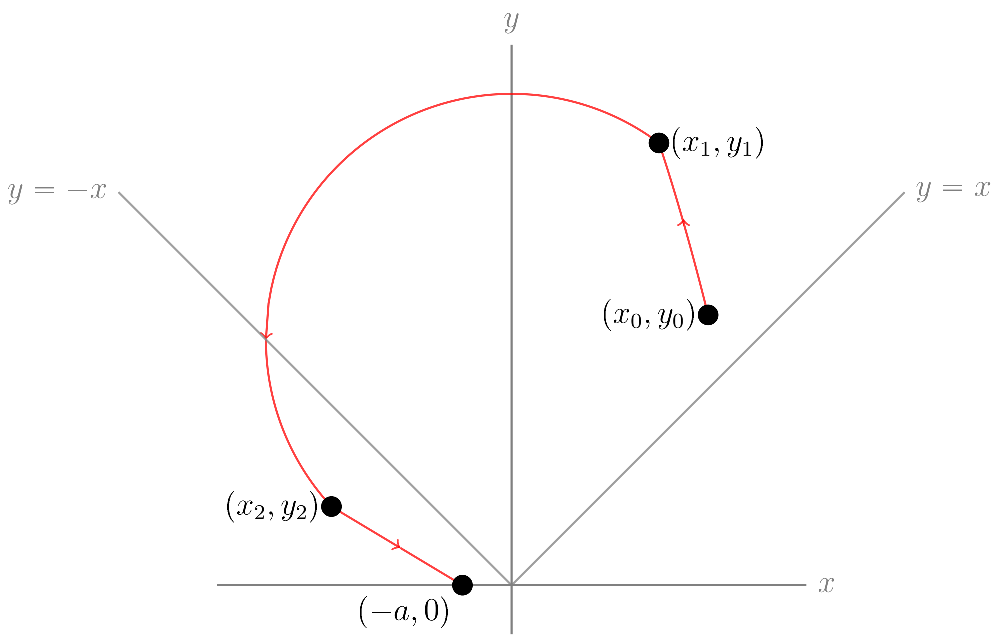

FIGURE B.1: An example control path connecting (x_0, y_0) and $(-a, 0)$.

sured that $\max_{y_1}\{T(y_1)\} = \infty$. To conclude the proof, we need an upper bound on $\min_{y_1}\{T(y_1)\}$ which is independent of $\mathbb{x}_0$.

Recall that we have assumed that $y_1 \geqslant \max\{8a, y_0\}$ and deduced that $T_1 \leqslant \frac{1}{a}$ and $T_2 \leqslant \frac{4}{3a}$ for all $\mathbb{x}_0 \in \mathbb{R}^2$ and choices of y_1. Since $T_0 = y_1 - y_0$, we see that

$$\min\{T(y_1) : y_1 \geqslant \max(8a, y_0)\} \leqslant 8a + \frac{1}{a} + \frac{4}{3a}\,.$$

Hence if we set $T_*(a) = 8a + \frac{7}{3a}$, we ensure that the range of $T(y_1)$ includes $[T_*(a), \infty)$. Hence for any value $T \in [T_*(a), \infty)$, the intermediate value theorem guarantees the existence of some point y_1 with $T(y_1) = T$. □

In the following lemma, we show that for any fixed terminal point $\mathbb{x}^*$ in the left-half plane, the process can traverse from a point of the form $(-a, 0)$ with $a > 0$ to $\mathbb{x}^*$ by a suitably constructed control.

Lemma B.2. *Let $z^* = (x^*, y^*)$ with $x^* < 0$ be fixed. Then there exist an $a > 0$, a finite time T^*, and a control $U \in L^2([0, T^*] \to \mathbb{R})$ such that it takes exactly time T^**

for the process $\mathbb{X}_t^U$ *to traverse from the point* $(-a, 0)$ *to* $\mathbb{x}^*$*, i.e.,*

$$(X_0^U, Y_0^U) = (-a, 0) \quad \textit{and} \quad (X_{T^*}^U, Y_{T^*}^U) = \mathbb{x}^* \,.$$

Proof of Lemma B.2. By symmetry, it suffices to consider the case where $\mathbb{x}^* = (x^*, y^*)$ in the left-half plane lies on or above the x-axis. Here, we examine three different possibilities, depending upon whether $\mathbb{x}^* \in B_1, B_2,$ or B_3, where

$$B_1 = \{(x, y) : x < 0, 0 \leqslant y < |x|\}$$

$$B_2 = \{(x, y) : x < 0, y > |x|\}$$

$$B_3 = \{(x, y) : x < 0, y = |x|\} \,.$$

Our control will again be of the form $U_t = u(X_t^U, Y_t^U, t)$.

Case 1: $\mathbb{x}^* = (x^*, y^*) \in B_1$: We choose any $a > |x^*|$ and construct a linear path from $(-a, 0)$ to $\mathbb{x}$ by setting

$$u(x, y) = -2xy + m(x + a) \quad \text{where} \quad m = \frac{y^*}{x^* + a} > 0 \,.$$

It is elementary to check that the x-derivative along this linear path from $(-a, 0)$ to $\mathbb{x}^*$ is strictly positive and bounded from below. Therefore the time T^* to move between $(a, 0)$ and $\mathbb{x}^*$ is finite.

Case 2: $\mathbb{x}^* = (x^*, y^*) \in B_2$: We choose $0 < a < 1$ so that $a < |x^*|$. This guarantees that $(-a, 0)$ lies to the right of the terminal point $\mathbb{x}^*$. We construct the control in three parts by specifying two intermediate points, denoted $\mathbb{x}_1 = (x_1, y_1)$ and $\mathbb{x}_2 = (x_2, y_2)$, between $(-a, 0)$ and $\mathbb{x}^*$. The point $\mathbb{x}_1$ is chosen to lie in B_1, to the right of $(-a, 0)$ and just below the line $y = -x$. The point $\mathbb{x}_2$ is chosen to lie just above the line $y = -x$, to the right of $\mathbb{x}^*$. We employ a linear path to move between $(-a, 0)$ and $\mathbb{x}_1$ in time T_1, a path along which the y-derivative is strictly positive to cross the line $y = -x$ in time T_2, and another linear path to travel from $\mathbb{x}_2$ to $\mathbb{x}^*$ in time T_3. Each of these times can be easily bounded and is completely determined by the choices of $\mathbb{x}_1$ and $\mathbb{x}_2$. We will then set $T^* = T_1 + T_2 + T_3$.

We specify the points $\mathbb{x}_1$ and $\mathbb{x}_2$ shortly. Given these points, we let

$$m_1 = \frac{y_1}{x_1 + a} \quad \text{and} \quad m_2 = \frac{y^* - y_2}{x^* - x_2}$$

be the slopes of the linear paths. We then define the control by

$$u(x,y,t) = \begin{cases} u_1(x,y) = -2xy + m_1(x+a) & \text{for } t \in (0, T_1] \\ u_2(x,y) = -2xy + 1 & \text{for } t \in (T_1, T_1 + T_2] \\ u_3(x,y) = -2xy + m_2(x - x_2) + y_2 & \text{for } t \in (T_1 + T_2, T^*] \end{cases} .$$

Since in region B_1, the x-derivative is always positive and in region B_2, the x-derivative is always negative, T_1 and T_3 will be finite if and only if $x_1 > -a$ and $x^* < x_2$. Our choices of x_1 and x_2 will satisfy these restrictions. The first restriction is satisfied by setting $\mathbb{x}_1 = (x_1, y_1) = \left(-\frac{a}{2}, \frac{a}{4}\right)$. Note that for $t \in (T_1, T_1 + T_2]$, the controlled system $\mathbb{X}_t^U = (X_t^U, Y_t^U)$ is the solution to the ODE

$$dX_t^U = \left((X_t^U)^2 - (Y_t^U)^2\right) dt$$
$$dY_t^U = dt$$

with initial condition $(X_{T_1}^U, Y_{T_1}^U) = (x_1, y_1)$. Set $y_2 = \frac{3a}{4}$. Then $Y_{T_2}^U = y_2 = y_1 + T_2$ implies that $T_2 = \frac{a}{2}$. Moreover, for $t \in (T_1, T_1 + T_2]$,

$$dX_t^U \geqslant -(Y_t^U)^2 \geqslant -\left(\frac{3a}{4}\right)^2 .$$

This in turn implies that

$$X_{T_2}^U = x_2 \geqslant x_1 - \left(\frac{3a}{4}\right)^2 T_2 = -\frac{a}{2} - \frac{9a^3}{32} > -a$$

since $a < 1$. Thus, $x^* < x_2$. Therefore, the total time T^* is finite. Figure B.2 illustrates an example path connecting $(-a, 0)$ and $\mathbb{x}^* = (x^*, y^*)$ when $\mathbb{x}^* \in B_2$.

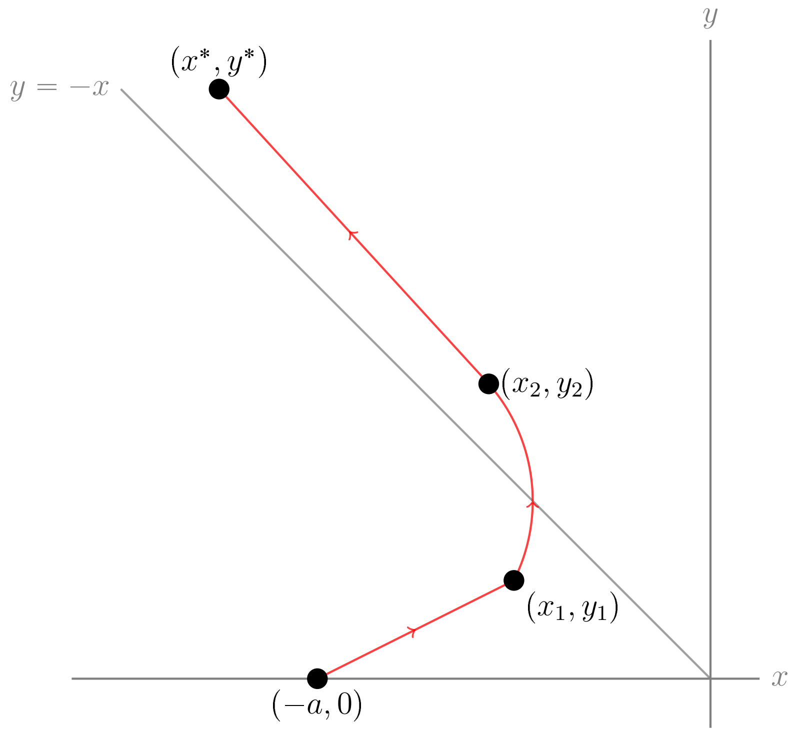

FIGURE B.2: An example control path connecting $(-a, 0)$ and $(x^*, y^*) \in B_2$.

Case 3: $\mathbb{x}^* = (x^*, y^*) \in B_3$: We note that every point on the line $y = -x$, except the origin, intersects with a single deterministic trajectory. We use the arguments in the previous case to construct a control to reach an intermediate point $\mathbb{x}_2 \in B_2$ such that $\mathbb{x}_2$ lies on the deterministic trajectory through $\mathbb{x}^*$, and from here we use the deterministic dynamics to reach $\mathbb{x}^*$. In particular, for $\mathbb{x}^* = (x^*, -x^*)$, we need only choose a point $\mathbb{x}_2 = (x_2, y_2)$ in the left-half plane and above $y = -x$ such that $\frac{x_2^2 + y_2^2}{2y_2} = -x^*$. For instance, we can choose a point $y_2 = -2x_2$ with $x_2 = 4/5x^*$. $\square$

B.1.3 Proof of Theorem B.1

We now invoke the two lemmas proven above, as well as the results from Malliavin calculus and control theory, to prove the theorem regarding the positivity of the density function.

Proof of Theorem B.1. Let $\mathbb{x}^* = (x^*, y^*)$ with $x^* < 0$ be fixed. Then by Lemma B.2, there exist an $a > 0$, a finite time $\bar{T}^*$, and a control $\bar{U} \in L^2([0, \bar{T}^*] \to \mathbb{R})$ such that it takes exactly time $\bar{T}^*$ for the process $\mathbb{X}_t^{\bar{U}}$ to traverse from the point $(-a, 0)$ to $\mathbb{x}^*$. Now by Lemma B.1, there exists a finite time $\hat{T}^*$ such that for all $T \geqslant \hat{T}^*$ and for all $\mathbb{x} \in \mathbb{R}^2$, there exists a control $\hat{U} \in L^2([0, T] \to \mathbb{R})$ such that it takes exactly time T for the process $\mathbb{X}_t^{\hat{U}}$ to traverse from the point $\mathbb{x}$ to the point $(-a, 0)$. We set $T^* = \bar{T}^* + \hat{T}^*$. Then for any $T > T^*$ and any $\mathbb{x} \in \mathbb{R}^2$, we construct the following control:

$$U_t = \begin{cases} \hat{U}_t & \text{for } t \in [0, T - \bar{T}^*) \\ \bar{U}_t & \text{for } t \in [T - \bar{T}^*, T] \end{cases} .$$

Then $U \in L^2([0, T] \to \mathbb{R})$ and U induces a path from $\mathbb{x}$ to $\mathbb{x}^*$ in exactly time T. Thus, $\Phi_{T,\mathbb{x}}(U) = \mathbb{x}^*$. Combining this fact with Proposition B.1 proves the positivity claim in Theorem B.1.

The fact that $p_t(\mathbb{x}_0, \mathbb{x}_1) = 0$ when $\mathbb{x}_0$ is in the left-half plane (i.e., $\mathbb{x}_0 = (x_0, y_0)$ with $x_0 < 0$) and $\mathbb{x}_1$ is in the right-half plane (i.e., $\mathbb{x}_1 = (x_1, y_1)$ with $x_1 > 0$) follows from Proposition B.1, provided we can show that there is no control which moves the process from $\mathbb{x}_0$ to $\mathbb{x}_1$. To see this, observe that except for the fixed point at the origin, the vector field for any control always points toward the left half-plane along the y-axis. Hence it is impossible to leave the left-half plane. The fact that $p_t(\mathbb{x}_0, \mathbb{x}_1) = 0$ when $x_1 = 0$ follows from the continuity of $p_t(\mathbb{x}_0, \mathbb{x}_1)$. □

B.2 Positivity of Invariant Measure

In this section we give a characterization of the positivity of the invariant probability measure for the model problem, which is summarized in the theorem below.

Theorem B.2. *If $\epsilon_y > 0$, then the invariant probability measure μ for the model problem given by* (4.2) *has a smooth density with respect to Lebesgue measure, denoted by m. If $\epsilon_x > 0$ and $\epsilon_y > 0$, then $m(x, y) > 0$ for all $(x, y) \in \mathbb{R}^2$. If $\epsilon_x = 0$ and*

$\epsilon_y > 0$, *then* $m(x,y) = 0$ *if* $x \geqslant 0$ *and* $m(x,y) > 0$ *if* $x < 0$.

Proof of Theorem B.2. The existence of a smooth invariant probability density, m, follows from the hypoellipticity of the generator, and hence hypoellipticity of the adjoint generator, when $\epsilon_y > 0$ since m satisfies $\mathcal{L}^* m = 0$. Now for any $t \geqslant 0$,

$$\begin{aligned}\mu(A) &= \int_{\mathbb{R}^2} P_t(\mathbb{x}', A)\mu(d\mathbb{x}') \\ &= \int_{\mathbb{R}^2}\int_A p_t(\mathbb{x}', \mathbb{x}) m(\mathbb{x}') d\mathbb{x} d\mathbb{x}' \\ &= \int_A \int_{\mathbb{R}^2} p_t(\mathbb{x}', \mathbb{x}) m(\mathbb{x}') d\mathbb{x}' d\mathbb{x} \quad \text{by Fubini's Theorem.}\end{aligned}$$

Hence, m satisfies the equality

$$m(\mathbb{x}) = \int_{\mathbb{R}^2} p_t(\mathbb{x}', \mathbb{x}) m(\mathbb{x}') d\mathbb{x}' \tag{B.4}$$

for all $t \geqslant 0$.

We now consider the positivity of the invariant probability density. Suppose there exists a point $\mathbb{x}$ such that for some $t > 0$,

$$p_t(\mathbb{x}', \mathbb{x}) > 0 \quad \text{for all} \quad \mathbb{x}' \in \mathbb{R}^2 \,. \tag{B.5}$$

Then since m integrates to one and is smooth, there must exist some open set A such that $m(\mathbb{x}') > 0$ for all $\mathbb{x}' \in A$. Combining this observation with (B.4), we obtain

$$m(\mathbb{x}) \geqslant \int_A p_t(\mathbb{x}', \mathbb{x}) m(\mathbb{x}') d\mathbb{x}' > 0 \,.$$

Hence we deduce that $m(\mathbb{x})$ is positive at any point $\mathbb{x}$ which satisfies (B.5). Applying this result to the information on the positivity of p_t in the case of nondegenerate noise and the case of degenerate noise from Theorem B.1, we obtain the conclusions about the positivity of $m(\mathbb{x})$ stated in Theorem B.2.

To deduce the statement that $m(x, y) = 0$ for $x \geqslant 0$ when $\epsilon_x = 0$ and $\epsilon_y > 0$, we use Theorem B.1, which states that if $\mathbb{x} \in H_+ = \{(x, y) \in \mathbb{R}^2 : x \geqslant 0\}$ and $\mathbb{x}' \in H_-$, where H_- is defined to be the complement of H_+, then $p_t(\mathbb{x}', \mathbb{x}) = 0$ for all $t > 0$. This implies that if $\mathbb{x} \in H_+$, then

$$m(\mathbb{x}) = \int_{H_+} p_t(\mathbb{x}', \mathbb{x}) m(\mathbb{x}') d\mathbb{x}' .$$

Integrating this expression over H_+ and interchanging the order of integration, we obtain

$$\begin{aligned} \int_{H_+} m(\mathbb{x}) d\mathbb{x} &= \int_{H_+} \int_{H_+} p_t(\mathbb{x}', \mathbb{x}) m(\mathbb{x}') d\mathbb{x}' d\mathbb{x} \\ &= \int_{H_+} m(\mathbb{x}') \int_{H_+} p_t(\mathbb{x}', \mathbb{x}) d\mathbb{x} d\mathbb{x}' \\ &= \int_{H_+} P_t(\mathbb{x}', H_+) m(\mathbb{x}') d\mathbb{x}' . \end{aligned}$$

This implies that

$$\int_{H_+} (1 - P_t(\mathbb{x}, H_+)) m(\mathbb{x}) d\mathbb{x} = \int_{H_+} P_t(\mathbb{x}, H_-) m(\mathbb{x}) d\mathbb{x} = 0 .$$

Since $m(\mathbb{x}) \geqslant 0$ and $P_t(\mathbb{x}, H_-) \geqslant 0$ for all $\mathbb{x} \in \mathbb{R}^2$, this in turn implies that for Lebesgue-almost-every $\mathbb{x} \in H_+$, either $m(\mathbb{x}) = 0$ or $P_t(\mathbb{x}, H_-) = 0$. Yet from Theorem B.1, we know that given any $\mathbb{x}' \in H_-$, there exists a time $t > 0$ so that $p_t(\mathbb{x}, \mathbb{x}') > 0$ for all $\mathbb{x} \in \mathbb{R}^2$. Because $p_t(\mathbb{x}, \mathbb{x}')$ is continuous, we know $P_t(\mathbb{x}, H_-) > 0$ for every $\mathbb{x} \in H_+$. This implies that $m(\mathbb{x}) = 0$ for almost every $\mathbb{x} \in H_+$. Since $m(\mathbb{x})$ is continuous, this forces $m(\mathbb{x}) = 0$ for all $\mathbb{x} \in H_+$. This completes the proof of Theorem B.2. □

Bibliography

[1] M. Abramowitz and I. A. Stegun, editors. *Handbook of mathematical functions with formulas, graphs, and mathematical tables.* Dover Publications Inc., New York, 1992. Reprint of the 1972 edition.

[2] L. Arnold and W. Kleimann. Qualitative theory of stochastic systems. In *Probabilistic analysis and related topics, Vol. 3*, pages 1–79. Academic Press, New York, 1983.

[3] A. Athreya, T. N. Kolba, and J. C. Mattingly. Propagating Lyapunov functions to prove noise-induced stabilization. *ArXiv e-prints*, Nov. 2011.

[4] M. T. Barlow and D. Nualart. *Lectures on probability theory and statistics*, volume 1690 of *Lecture Notes in Mathematics.* Springer-Verlag, Berlin, 1998. Lectures from the 25th Saint-Flour Summer School held July 10–26, 1995, Edited by P. Bernard.

[5] R. F. Bass. *Diffusions and elliptic operators.* Probability and its Applications (New York). Springer-Verlag, New York, 1998.

[6] G. Ben Arous and R. Léandre. Décroissance exponentielle du noyau de la chaleur sur la diagonale. I. *Probab. Theory Related Fields*, 90(2):175–202, 1991.

[7] G. Ben Arous and R. Léandre. Décroissance exponentielle du noyau de la chaleur sur la diagonale. II. *Probab. Theory Related Fields*, 90(3):377–402, 1991.

[8] J. Birrell, D. P. Herzog, and J. Wehr. Transition from ergodic to explosive behavior in a family of stochastic differential equations. *ArXiv e-prints*, May 2011.

[9] K. Bodová. *Topics in applied stochastic dynamics.* PhD thesis, University of Michigan, Ann Arbor, MI, 2009.

[10] K. Bodová and C. R. Doering. Noise-induced statistically stable oscillations in a deterministically divergent nonlinear dynamical system. *Communication Mathematical Science*, 10(1):137–157, 2012.

[11] B. Cooke, J. C. Mattingly, S. A. McKinley, and S. C. Schmidler. Geometric ergodicity of two–dimensional Hamiltonian systems with a Lennard–Jones–like repulsive potential. *ArXiv e-prints*, Apr. 2011.

[12] G. Da Prato. *An introduction to infinite-dimensional analysis*. Universitext. Springer-Verlag, Berlin, 2006. Revised and extended from the 2001 original by Da Prato.

[13] R. Douc, G. Fort, and A. Guillin. Subgeometric rates of convergence of f-ergodic strong Markov processes. *Stochastic Process. Appl.*, 119(3):897–923, 2009.

[14] P. Dupuis and R. J. Williams. Lyapunov functions for semimartingale reflecting Brownian motions. *Ann. Probab.*, 22(2):680–702, 1994.

[15] G. Fort, S. Meyn, E. Moulines, and P. Priouret. The ODE method for stability of skip-free Markov chains with applications to MCMC. *Ann. Appl. Probab.*, 18(2):664–707, 2008.

[16] K. Gawędzki, D. P. Herzog, and J. Wehr. Ergodic properties of a model for turbulent dispersion of inertial particles. *Comm. Math. Phys.*, 308(1):49–80, 2011.

[17] M. Gitterman. *The noisy oscillator*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2005. The first hundred years, from Einstein until now.

[18] M. Hairer. P@W course on the convergence of Markov processes. http://www.hairer.org/notes/Convergence.pdf, 2010.

[19] M. Hairer and J. C. Mattingly. Yet another look at Harris' ergodic theorem for Markov chains. *ArXiv e-prints*, Oct. 2008.

[20] M. Hairer and J. C. Mattingly. Slow energy dissipation in anharmonic oscillator chains. *Comm. Pure Appl. Math.*, 62(8):999–1032, 2009.

[21] T. E. Harris. The existence of stationary measures for certain Markov processes. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. II*, pages 113–124, Berkeley and Los Angeles, 1956. University of California Press.

[22] R. Z. Has'minskiĭ. *Stochastic stability of differential equations*, volume 7 of *Monographs and Textbooks on Mechanics of Solids and Fluids: Mechanics and Analysis*. Sijthoff & Noordhoff, Alphen aan den Rijn, 1980. Translated from the Russian by D. Louvish.

[23] D. P. Herzog. *Geometry's fundamental role in the stability of stochastic differential equations*. PhD thesis, University of Arizona, Tucson, AZ, 2011.

[24] J. Huang, I. Kontoyiannis, and S. P. Meyn. The ODE method and spectral theory of Markov operators. In *Stochastic theory and control (Lawrence, KS, 2001)*, volume 280 of *Lecture Notes in Control and Inform. Sci.*, pages 205–221. Springer, Berlin, 2002.

[25] J. C. Mattingly, S. A. McKinley, and N. S. Pillai. Geometric ergodicity of a bead-spring pair with stochastic Stokes forcing. *ArXiv e-prints*, Feb. 2009.

[26] J. C. Mattingly, A. M. Stuart, and D. J. Higham. Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise. *Stochastic Process. Appl.*, 101(2):185–232, 2002.

[27] S. Meyn. *Control techniques for complex networks*. Cambridge University Press, Cambridge, 2008.

[28] S. Meyn and R. L. Tweedie. *Markov chains and stochastic stability*. Cambridge University Press, Cambridge, second edition, 2009. With a prologue by Peter W. Glynn.

[29] S. P. Meyn and R. L. Tweedie. *Markov chains and stochastic stability*. Communications and Control Engineering Series. Springer-Verlag London Ltd., London, 1993.

[30] B. Øksendal. *Stochastic differential equations*. Universitext. Springer-Verlag, Berlin, third edition, 1992. An introduction with applications.

[31] M. Scheutzow. Stabilization and destabilization by noise in the plane. *Stochastic Anal. Appl.*, 11(1):97–113, 1993.

[32] M. Scheutzow. An integral inequality and its application to a problem of stabilization by noise. *J. Math. Anal. Appl.*, 193(1):200–208, 1995.

[33] U. Thygesen. A survey of Lyapunov techniques for stochastic differential equations. Technical report, The Technical University of Denmark, 1997.

[34] A. Y. Veretennikov. On polynomial mixing bounds for stochastic differential equations. *Stochastic Process. Appl.*, 70(1):115–127, 1997.

[35] A. Y. Veretennikov. On polynomial mixing and the rate of convergence for stochastic differential and difference equations. *Teor. Veroyatnost. i Primenen.*, 44(2):312–327, 1999.

Biography

Tiffany Nicole Kolba was born on December 30, 1985 in Tacoma Park, Maryland to parents James and Deborah Tasky, and was raised in Bowie, Maryland. She graduated Phi Beta Kappa from Johns Hopkins University in May 2006 with a B.A./M.A. in Mathematics, as well as a second major in Applied Mathematics and Statistics. At Johns Hopkins University, Tiffany received the J.J. Sylvester Award for Excellence in Mathematics, the Applied Mathematics and Statistics Achievement Award, and the Applied Mathematics and Statistics Award for Excellence in Teaching.

Tiffany received her M.A. and Ph.D. degrees in Mathematics from Duke University in December 2007 and May 2012, respectively. She also completed a Certificate in College Teaching, administered through the Duke University Graduate School. While a graduate student at Duke University, she was a recipient of a James B. Duke fellowship and a fellowship from the Statistical and Applied Mathematical Sciences Institute. She also received the L.P. and Barbara Smith Award for Excellence in Teaching from the Duke University Mathematics Department in August 2009.

In June 2008, Tiffany married Mark Philip Kolba, who received his Ph.D. in Electrical Engineering from Duke University. They currently reside in Durham, North Carolina with their daughter Kayla and kitten Mystery. In August 2012, Tiffany will be joining the faculty at Valparaiso University in Indiana as a tenure-track Assistant Professor of Mathematics.