5-2018

Physical Applications of the Geometric Minimum Action Method

George L. Poppe Jr.

The Graduate Center, City University of New York

How does access to this work benefit you? Let us know!

Follow this and additional works at: https://academicworks.cuny.edu/gc_etds

Part of the Applied Statistics Commons, Dynamical Systems Commons, Fluid Dynamics Commons, Non-linear Dynamics Commons, Numerical Analysis and Scientific Computing Commons, Optics Commons, Ordinary Differential Equations and Applied Dynamics Commons, Partial Differential Equations Commons, Statistical Models Commons, and the Statistical, Nonlinear, and Soft Matter Physics Commons

Recommended Citation

https://academicworks.cuny.edu/gc_etds/2739

This Dissertation is brought to you by CUNY Academic Works. It has been accepted for inclusion in All Dissertations, Theses, and Capstone Projects by an authorized administrator of CUNY Academic Works. For more information, please contact deposit@gc.cuny.edu.
PHYSICAL APPLICATIONS OF THE GEOMETRIC MINIMUM ACTION METHOD

by

GEORGE L. POPPE JR.

A dissertation submitted to the Graduate Faculty in Physics in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2018
Physical Applications of the Geometric Minimum Action Method

by

George Poppe

This manuscript has been read and accepted by the Graduate Faculty in Physics in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.
Physical Applications of the Geometric Minimum Action Method

George L. Poppe Jr.

Physics Program
CUNY Graduate Center
City University of New York

Abstract

This thesis extends the landscape of rare events problems solved on stochastic systems by means of the geometric minimum action method (gMAM). These include partial differential equations (PDEs) such as the real Ginzburg-Landau equation (RGLE), the linear Schrödinger equation, along with various forms of the nonlinear Schrödinger equation (NLSE) including an application towards an ultra-short pulse mode-locked laser system (MLL). Additionally we develop analytical tools that can be used alongside numerics to validate those solutions. This includes the use of instanton methods in deriving state transitions for the linear Schrödinger equation and the cubic diffusive NLSE. These analytical solutions are shown to be in good agreement with the numerics. Lastly this thesis investigates the relationship between such PDEs and associated ordinary differential equation (ODE) reductions. We find that while there is good agreement for certain properties of the two systems, the ODE model can have difficulty reproducing some aspects of the PDE solution.
Acknowledgments

First, I would like to thank everyone. Parents, grandparents, siblings and friends. Your influence on me was not always good, but with the blame comes a share of the credit.

My fellow students and colleagues at the Graduate Center. Alongside whom I had the great opportunity to learn and grow, and without whom I would not be who I am today.

Lastly, both my official and unofficial advisors, Tobias Schäfer and Vadim Oganesyan. To whom I can not overstate my gratitude. . .
# Contents

<table>
<thead>
<tr>
<th>List of Figures</th>
<th>viii</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1 Introduction</strong></td>
<td></td>
</tr>
<tr>
<td>1.1 Stochastic Methods and Rare Events</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Optics and Lasers</td>
<td>3</td>
</tr>
<tr>
<td><strong>2 Stochastic Modeling</strong></td>
<td>11</td>
</tr>
<tr>
<td>2.1 Stochastic Differential Equations</td>
<td>11</td>
</tr>
<tr>
<td>2.2 The Fokker-Planck Equation</td>
<td>20</td>
</tr>
<tr>
<td>2.3 Path Integrals</td>
<td>28</td>
</tr>
<tr>
<td><strong>3 Large Deviation Theory</strong></td>
<td>35</td>
</tr>
<tr>
<td>3.1 The Freidlin-Wentzell Action Functional</td>
<td>35</td>
</tr>
<tr>
<td>3.2 Instantons</td>
<td>39</td>
</tr>
<tr>
<td>3.3 Example: Diffusive Ornstein Uhlenbeck</td>
<td>43</td>
</tr>
<tr>
<td><strong>4 Numerical Methods</strong></td>
<td>45</td>
</tr>
<tr>
<td>4.1 Filtering</td>
<td>45</td>
</tr>
<tr>
<td>4.2 String Method</td>
<td>47</td>
</tr>
<tr>
<td>4.3 gMAM</td>
<td>49</td>
</tr>
<tr>
<td>4.4 gMAM for PDEs</td>
<td>56</td>
</tr>
<tr>
<td>4.5 Linear Schrödinger Equation</td>
<td>59</td>
</tr>
<tr>
<td>4.5.1 Analytic Solution</td>
<td>59</td>
</tr>
<tr>
<td>4.5.2 Applying gMAM to the linear Schrödinger SPDE</td>
<td>60</td>
</tr>
<tr>
<td><strong>5 Applications</strong></td>
<td>65</td>
</tr>
<tr>
<td>5.1 Maier-Stein</td>
<td>65</td>
</tr>
<tr>
<td>5.1.1 Maier-Stein: String Method</td>
<td>65</td>
</tr>
<tr>
<td>5.1.2 Maier-Stein(SDE): gMAM</td>
<td>68</td>
</tr>
<tr>
<td>5.1.3 Maier-Stein (PDE): gMAM</td>
<td>71</td>
</tr>
<tr>
<td>5.2 Real Ginzburg Landau Equation</td>
<td>75</td>
</tr>
<tr>
<td>5.2.1 Applying the String Method (RGLE)</td>
<td>76</td>
</tr>
<tr>
<td>5.2.2 Applying the gMAM (RGLE)</td>
<td>78</td>
</tr>
</tbody>
</table>
6 Instantons of the cubic nonlinear Schrödinger equation 82
6.1 Mode-Locked Laser PDE/Introduction 82
6.2 NLSE and Solitons 82
6.3 Chernykh-Stepanov Step applied towards the cubic diffusive NLSE 84
6.4 Nonlinear OU comparison using Instantons 89
6.5 Kerr Lens Mode-Locked Laser SPDE 92

7 ODE reduction and comparison to PDE model 94
7.1 NLSE with Cosine Potential 94
7.2 Moore’s Model 98
7.3 Deterministic Comparison of Laser PDE vs ODE reduction 100
7.4 gMAM for Moore’s Model (SODE) 102
7.5 Comparing the PDE to the ODE reduction (Simple Case: Diffusive NLSE) 105
7.6 PDE-ODE comparison for the full MLL PDE model 109

8 Conclusion 111

9 Appendix 113
9.1 Mode-Locked Laser SPDE to ODE Reduction 113
9.1.1 Writing the Lagrangian 114
9.1.2 Finding the Equations of Motion 115
9.1.3 Resolving $\Xi$ 118

10 References 121
List of Figures

1. Velocity field corresponding to the Ornstein-Uhlenbeck drift vector $b = -\alpha x$. Arrows indicate direction and relative magnitude of the velocities in space that a particle would have if it occupied that position. ................. 12

2. Vector field corresponding to circular drift vector $\dot{u} = -\sin(\theta), \dot{v} = \cos(\theta)$. Wherever particle is placed, it stays at that radius and orbits the origin. ... 13

3. Velocity field corresponding to the combined circular and Ornstein-Uhlenbeck drift vector $\dot{u} = -\sin(\theta) - \alpha u, \dot{v} = \cos(\theta) - \alpha v$. In this case $\alpha = 1$. All initial conditions lead to particle having a spiral trajectory towards origin. .... 14

4. Velocity field corresponding to the combined Maier-Stein drift vector $\dot{u} = u - u^3 - \beta uv^2, \dot{v} = -(1 + u^2)v$. In this case $\beta = 1$. There are two real stable fixed points at (1,0) and (-1,0). ......................... 15

5. Velocity field corresponding to the “classical pendulum” drift vector with some dissipation caused by an Ornstein-Uhlenbeck term. Namely: $\dot{u} = v, \dot{v} = -\sin(u) - \alpha v$. There are stable fixed points at any position $(2\pi n, 0)$ with $n$ being any integer value. ......................... 17

6. Displays the probability density $p = p(x,t)$ as a function of space $x$ and time $t$. This is a numerical simulation of the Fokker-Planck equation where only the deterministic component ($b = -.01x$) is active. The distribution is initialized to an approximate delta function centered at $x = 10$. As time goes on, the peak of the distribution moves from $x = 10$ at $t = 0$ (background) to $x = 9.89$ at $t = .4$ (foreground) due to the deterministic effects of the Ornstein-Uhlenbeck drift vector. ......................... 23

7. Displays the probability density $p = p(x,t)$ as a function of space $x$ and time $t$. This is a numerical simulation of the Fokker-Planck equation with $b = 0$ and $\sigma = .1$. The distribution initial begins approximately as a delta function (background) and broadens (foreground) due to the stochastic effects of the Wiener process. ......................... 24

8. Displays the probability density for the full Ornstein-Uhlenbeck system as a function of time. System begins in a near-delta function centered at position $x = 10$ (background). The system evolves in time such that the width of the distribution increases and the location of the peak drifts towards the origin ending at $x = 9.05$. ......................... 27

9. The image on the left shows the density of paths taken by a particle which begins at $x = 0$ and ends at $x = 1$ under the influence of the Ornstein Uhlenbeck potential. Red indicates a high path density, blue indicates low path density. The image on the right shows the path given by the most common numerical trajectory (blue) as compared to the theoretical result calculated by the path integral method (green). ......................... 33
10 A color plot of the results of applying the filtering method to the Ornstein-Uhlenbeck damped pendulum. Dark blue represents low probability density, and red represents very high probability density. The particle begins at the fixed point \((u, v) = (0, 0)\) and ends at the fixed point \((u, v) = (2\pi, 0)\). The trajectory is a spiral path out of the potential well centered at \((0, 0)\) followed by a similar spiral path into the neighboring well centered at \((2\pi, 0)\).

11 Displays the gMAM solution to the linear Schrödinger equation that describes the transition from \([u(T_{\text{min}}, x), v(T_{\text{min}}, x)] = [0, 0]\) to \([u(0, x), v(0, x)] = [a \exp(-|x|/\sqrt{\beta}), 0]\). The V component is roughly an order of magnitude smaller than the U component. This corresponds to having a weak dispersion term.

12 Comparison between a slice of the the analytic instanton calculation for the linear Schrödinger equation (red) against the gMAM result (blue) for arc length parameter \(\alpha = .78\). The left panel shows the real part \(u\), the right panel the imaginary part \(v\).

13 Initial conditions for Maier-Stein string method. The initial condition is a curve that connects the fixed points \((u, v) = (-1, 0)\) and \((u, v) = (1, 0)\).

14 The final state of the system after relaxation for Maier-Stein string method. The string Method accurately reveals the minimizing trajectory that represents a straight line between \((u, v) = (-1, 0)\) and \((u, v) = (1, 0)\).

15 Plot of the results for the Maier-Stein model for \(\beta = 1\) and \(\beta = 10\).

16 Figure shows \(u(\alpha, x)\) where each line in the figure represents a line of constant \(\alpha\). The trajectory proceeds from the bottom of the image at \(u(0, x) = -1\) towards the top of the image \(u(1, x) = 1\). This result is in agreement with the result produced in [1].

17 Displays \(\lambda \) versus \(\alpha\) corresponding to the minimizing solution. The root at \(\alpha = .5\) signals the existence of a saddle point and therefore suggests that our solution has captured the minimizer.

18 Displays the initial condition for the string method. We attempt to find the transition between \(u = -1\) and \(u = 1\) by creating a line of fixed endpoints corresponding to those two values at the start and the end of the transition.

19 Displays the string method solution to the RGLE for the transition between \(u = -1\) and \(u = 1\).

20 Displays the gMAM solution to the RGLE for the transition between \(u=-1\) and \(u=1\). As can be seen, it is identical to that of figure (19) showing good agreement between the two methods.

21 Displays \(\lambda \) vs \(\alpha\) corresponding to the minimizing solution. The root at \(\alpha = .5\) signals the existence of a saddle point and therefore suggests that our solution has captured the minimizer.
22 Soliton solutions for NLSE for $E[\phi^2] = 0$ and $E[\phi^2] = .002^2$ case. The small noise in the second case has the effect of destabilizing the soliton and displacing it from its natural equilibrium position. ........................................... 83
23 Displays the minimizer produced by the gMAM for the dissipative cubic NLSE. .................................................. 86
24 Here we see a plot of the difference between the two solutions of the position fields corresponding to the to the diffusive cubic NLSE (One solution resulting from the gMAM, the other resulting from exponential integration according to Chernykh-Stepanov step). The difference between the two solutions is very small relative to the peak height ($\approx .35$). ................................. 88
25 Here we see a plot of the difference between the two solutions of the momentum fields corresponding to the diffusive cubic NLSE PDE (One solution resulting from the gMAM, the other resulting from exponential integration according to Chernykh-Stepanov step). For the momentum fields, the CS step prescribes that we integrate backwards in time towards $t = -\infty$. The agreement between the two solutions is not as strong as we saw in figure 24. (peak height of $\theta_u \approx .6$) ................................. 89
26 Displays the solution of the $u$ and $v$ fields generated via gMAM for the PDE laser model. This figure was generated using parameters $c_1 = .01$, $c_2 = .1$, $c_3 = 1$, $c_4 = 1$, $d_1 = .3$, $d_2 = .1$, $b = .85$, $\omega = 1.5$. .................................................. 93
27 Comparison of simulated solutions of the deterministic NLSE with a cosine potential vs the simulated solution of the ODE that results from the effective Hamiltonian. The cos potential is shown in red; the PDE solution is in blue with "x" markers; the ODE solution is in green. The agreement between the two is very good. For this particular case, the soliton is in a bound state such that it bounces back and forth between the peaks of the cosine potential. .................................................. 96
28 Simulated solution to the NLSE cosine potential with additive Gaussian white noise. The parameters of the problem are tuned so that the destabilizing effects of the noise are nearly in balance with the stabilizing effects of the potential well. This creates a tunneling phenomenon with a metastable bound state. The top image displays the soliton at time $t = 2$, while the bottom shows the position at time $t = 40$. The soliton has "tunneled" from one metastable bound state into a nearby one. .................................................. 97
29 Comparison of simulated solutions of the deterministic phenomenological MLL PDE with vs the simulated solution of the ODE that results from the reduction carried out in [2]. The cosine potential is shown in red; the PDE solution is in blue with "x" markers; the ODE solution is in green. Just as with the NLSE with a cosine potential, there is good agreement between the solutions of the two cases. .................................................. 101
30 Displays the minimizer produced by the gMAM 3D ODE model for the transition from the point $U_1 = 1.5$ $U_2 = 0$ $U_3 = 0$ to the point $U_1 = 1.5$ $U_2 = 0$ $U_3 = 2$ with parameters: $c_1 = 1$ $c_2 = 1$ $d_1 = 3$ $d_2 = 1$ $b = .85$ $\omega = 1.5$ . . . . . . . . . . . 104

31 Displays comparative heat maps of $u$ for the PDE gMAM solution (left) vs the ODE gMAM solution (right) with parameters: $c_1 = .1$, $c_2 = .1$. We recognize the similarity of the two images. . . . . . . . . . . . . . . . . . . . . . . . . . . . . 106

32 Displays comparison of the heat map of $v$ for the PDE gMAM solution (left) against that of the ODE gMAM solution (right) with parameters: $c_1 = .1$, $c_2 = .1$. As in figure 31, we see partial agreement in the $v$ field as well. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 107

33 Left panel: plot of the amplitude $A(\alpha)$ of the ODE minimizer (red) in comparison to the amplitude of the PDE minimizer (blue) for the transition of a soliton with $|A|_1 = 0.5$ to $|A|_2 = 2.5$. Right panel: same comparison but for the exit from zero to the soliton with an amplitude $|A|_2 = 2.5$. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 108

34 Displays heat maps of $u$ for the ODE gMAM solution (left) vs the PDE gMAM solution (right) with parameters: $c_1 = .1$, $c_2 = .1$, $d_1 = 0.1$, $d_2 = 0.1$, $b = 0.1$, $\omega = 1.0$. There does appear to be broad correspondence between the two images. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 109

35 Displays heat maps of $v$ for the ODE gMAM solution (left) vs the PDE gMAM solution (right) with parameters: $c_1 = .1$, $c_2 = .1$, $d_1 = 0.1$, $d_2 = 0.1$, $b = 0.1$, $\omega = 1.0$. We see agreement similar to that seen in Figure 34 . . . . 110

36 Example plots of phase vs time (in blue) plotted on the same axes with a magnified plot of the amplitude vs time (in red) for $\alpha=0.180$ (left) and $s=0.624$ (right). We can see that the phase shows parabolic behavior within the time frame where the amplitude is large which could justify a using a parabolic phase model to characterize this system. As the system evolves along the minimizing path ($\alpha$), we see that the parabolic phase model starts to break down as an approximation of the waveform. . . . . . . . . . . . . . . . . . . . . . . . . . . . . 111
1 Introduction

1.1 Stochastic Methods and Rare Events

A rapidly growing field lying at the interface of physics and applied mathematics is the study of extreme events. This term refers to the study of improbable events which have a large effect on the state or dynamics of the system in question. Given the probabilistic nature of the definition, extreme events therefore exist in the context of stochastic systems where one or many aspects of the system can be characterized by a random variable. Famous examples include rogue waves ([3-6]), the weather [7], the financial market [8, 9], quantum mechanics, statistical mechanics [10, 11], turbulent fluid mechanics [12-14], or even population dynamics [15].

The term rare events is used to denote configurations of a system whose likelihood of occurrence is typically very far from the mean or median of the probability distribution associated with the state of the system. This typically implies that the timescale between two rare events occurring is going to be much larger than the timescale associated with more likely events. Simply because these states are unlikely on their own, does not imply that they are unimportant. There are many physical and mathematical systems where these rare events have drastic effects on the state of the system. This is especially true in metastable systems where the system has several equilibrium states and noise-driven transitions from one equilibrium to another are rare. In the context of stochastic differential equations (SDEs) one can imagine a randomly fluctuating particle that is hopping from one potential well to another where, in this analogy, the particle’s position actually represents some description of the state of a larger system. Conceptually this is similar to a quantum tunneling event. In the context of extreme events, one is usually interested

1 Strictly speaking, any scientific measurement is part of a stochastic system because any measurement will always have some random error associated with it, however, we are interested in cases where the randomness has a significant effect on the dynamics of the system.
in events that are not only low in likelihood but also large in impact. A typical example is a system failure, where the noise in a system triggers a transition from a desirable state to a failing state. In fiber optical signal transmission, for example, such an event is the transition from a pulse (representing a *one* in a bit stream) to a *zero* or vice versa. For a laser system it can mean the transition from a correctly operating state to a failing state.

To understand how these small fluctuations affect the system in question, an important mathematical framework is the Freidlin-Wentzell theory of large deviations [16]. Stated simply, the rarer the event in question is, the more likely it is to follow the classical path (the minimizer) of the associated Freidlin-Wentzell action. Broadly speaking, this allows us to localize the particle into some small area of phase space with exponential likelihood. There are only very few examples, where this minimizer can be computed analytically as the related Euler-Lagrange equations are rather complicated. Wishing to solve these problems numerically, it is important to select a method that allows us to do so with a certain degree of expediency. There are a number of available methods to solve these types of problems such as the so called *string method* [17] or the *Minimum Action Method* (MAM) [18]. For many physically relevant systems, there are numerical issues that arise when one tries to compute the path that minimizes the action directly. Many PDEs represent high dimensional systems that would test the limits of many brute force methods. Additionally the choice of proper boundary conditions can be very tricky. Lastly, the action is generally defined by the integral of the Lagrangian over all times. The minimum action path describing the exit from a stable fixed point requires an infinite amount of time, so numerically evaluating this integral is not feasible in many cases.

Recently M. Heymann and E. Vanden-Eijnden [19, 20] published a method for finding the solution paths associated with rare event transitions in stochastic differential equations (SDE). They call this method the Geometric Minimum Action Method (gMAM). As its name suggests it is a relaxation method that makes use of a geometric reparametrization
of the action.

In this thesis, we first give some background on stochastic systems and introduce gMAM using some ‘classic’ and novel examples. One novel result is the formulation of gMAM for linear and nonlinear Schrödinger systems. We then implement this formulation first for simple systems (i.e. the dissipative nonlinear Schrödinger equation, Real Ginzburg Landau Equation) and then show how to apply gMAM in the context of a highly complex model describing pulses in a femtosecond laser. We investigate the solution space of soliton systems and the relationship between PDE and ODE systems. Additionally we generate novel analytical solutions to these problems using techniques such as the instanton method.

A further goal of this work is to make gMAM and its numerical implementation more accessible to physicists. The original paper was written using notation that is more common among mathematicians that most physicists are unfamiliar with. As a practical consequence, it is somewhat inaccessible to many physicists who otherwise would find value in the numerical techniques. To this end, the document is written to be understood by all those with a basic education in physics.

1.2 Optics and Lasers

Some of the important novel applications of the geometric minimum action method will be discussed in the context of nonlinear optical systems, in particular stochastic laser systems. Often, such systems can be modeled by a stochastic nonlinear Schrödinger equation or equations with similar properties. There are numerous examples of such models in optics, for some examples we refer to [21, 2, 22]. So for the purposes of understanding some of the topics and terminology used later, we will briefly review the basics of these physical topics.
The field of optics, and by extension lasers, describes the behavior of light in specific physical situations. Starting as simply as possible, we can ask what it is that we mean when we say "light"? If we wish to be pedantic, light is an electromagnetic wave that propagates through space with a frequency that corresponds to the visible spectrum (∼400-700 nm). On the other hand, electromagnetic waves at other frequencies behave according to the same physics. Consequently, the field of optics could be more correctly stated as the study of the behavior of electromagnetic waves generally. All electromagnetic waves are subject to the physical laws known as Maxwell’s equations:

\[
\begin{align*}
\nabla \cdot \mathbf{D} &= 4\pi \rho_f \\
\nabla \times \mathbf{E} &= -\frac{1}{c} \frac{\partial \mathbf{B}}{\partial t} \\
\nabla \cdot \mathbf{B} &= 0 \\
\nabla \times \mathbf{H} &= \frac{1}{c} \frac{\partial \mathbf{D}}{\partial t} + \frac{4\pi}{c} \mathbf{J}_f
\end{align*}
\] (1)

These relations place a tight constraint on the dynamics of wave propagation. For example, in uniform and isotropic media, the electric and magnetic fields are always going to be orthogonal and will always be in phase with one another. It is possible to polarize the light along different direction (or even circularly polarize it), but the dynamics of propagation stay the same. Another set of constraints imposed by Maxwell’s equations are the boundary conditions associated with an electromagnetic wave passing from one medium to another:

\[
\begin{align*}
0 &= \hat{n} \cdot (\mathbf{B}^{(2)} - \mathbf{B}^{(1)}) \\
0 &= \hat{n} \times (\mathbf{E}^{(2)} - \mathbf{E}^{(1)}) \\
4\pi \rho_s &= \hat{n} \cdot (\mathbf{D}^{(2)} - \mathbf{D}^{(1)}) \\
\frac{4\pi}{c} \mathbf{J}_s &= \hat{n} \times (\mathbf{H}^{(2)} - \mathbf{H}^{(1)})
\end{align*}
\] (2)

These equations give the relationship of the electric and magnetic fields on one side of the boundary to the electric and magnetic fields on the other side of a boundary. A quick inspection of these equations shows that the boundary conditions depend on \(\varepsilon\) and \(\mu\).
The parameters $\varepsilon$ and $\mu$, are the permittivity and permeability of a medium respectively. These physical constants give coarse grained information about the distribution of electrons and protons associated with a medium and consequently give us information about the way electromagnetic waves can propagate inside that a medium.

It is well known that these constants determine how fast an electromagnetic wave can propagate through a given medium. In so doing, they also determine the index of refraction. The index of refraction (usually denoted with an "n") is a physical constant associated with a medium that is traditionally defined as the ratio of the phase velocity of light in the vacuum relative to the velocity of light that medium. Much of the field of optics is devoted to the study of systems with various indices of refraction (and therefore the study of the behavior of light in various dielectric materials).

\[ n_1 = \frac{\text{velocity of light in vacuum}}{\text{velocity of light in medium 1}} = \frac{\sqrt{\varepsilon_1 \mu_1}}{\sqrt{\varepsilon_0 \mu_0}} = \sqrt{\varepsilon_r \mu_r} \quad (3) \]

Here $\varepsilon_r$, $\mu_r$ are the relative permittivity and permeability. It is useful to note that the index of refraction is almost always frequency dependent and that it is often greater than one.

Given that the velocity of propagation of light depends on the properties of the medium, it follows that suddenly changing the medium of propagation is going to have an effect on the path that the light travels. In geometric optics, this effect is given by Snell’s law:

\[ n_1 \sin(\theta_i) = n_2 \sin(\theta_r) \quad (4) \]

Snell’s law also predicts that for certain media there will be a phenomenon called total internal reflection. This is the foundational idea for fiber optics. Unfortunately, however, understanding the details of sending light signals through fiber optic tubes requires a more advanced theory than geometric optics.
The signal in a fiber optic system is produced by a laser. Let us briefly review basic concepts of laser theory and design. Broadly speaking a laser is a device that produces a beam of light through stimulated emission. There is a large variety of lasers that one can study. These are often lumped into two main categories: Continuous wave lasers and pulse lasers. The former operates with the goal of producing a focused beam of monochromatic light often with a very narrow frequency spectrum. Pulse lasers can be either intended to be long pulse or shot pulse lasers. We will focus on the concepts associated with short pulse lasers for now. The adjective "short" in short pulse refers to the time of duration of the pulses as they pass by a particular position. This is directly proportional to the spatial width of a pulse via the group velocity as the proportionality constant. Electromagnetic waves (actually, waves in general) have a property that requires the spatial pulse width to be inversely proportional to the frequency bandwidth associated with the wave. Because of this constraint, short pulse lasers do not exhibit the monochromatic properties that continuous lasers often do. Techniques have been developed that allow lasers to produce high intensity pulses of light that have a temporal width on the order of femtoseconds. Naturally, such a short pulse implies having a correspondingly broad frequency spectrum. These are called ultra-short pulse lasers. One of our main interests in this paper will involve this type of laser.

Let us now turn to how lasers are made (for more detailed discussions see [23-25]). In general, the design of a laser can typically be described in terms of three physical aspects: The gain medium, the excitation process that creates the stimulated emission, and the feedback mechanism.

The gain medium of a laser is a material that has specific properties, due to its shape, or atomic structure, that amplifies the strength of the beam through stimulated emission. The idea behind stimulated emission is that the atoms in the gain medium can be raised to a meta-stable excited state, and when the beam passes through the medium, it causes
the atoms in the medium to emit photons at a frequency associated with the energy of the excited state. This process can be repeated again and again, amplifying the intensity of the laser each time. The gain medium needs to be specially selected to emit light at a particular frequency as well as the particular phase of the incident photons. For example, an ultra-short pulse laser will require a gain medium that can emit photons across a wide frequency spectrum. Otherwise it will not be able to amplify all the modes associated with the laser pulse.

There are a number of methods that have been developed to excite the atoms in the gain medium to induce population inversion and therefore cause stimulated emission to occur. Both continuous wave lasers and pulse lasers require a pump to keep the atoms in the excited state. This has been accomplished through various designs including gas discharges (e.g. hollow cathodes, arc discharges), optical pumping (e.g. flash-lamps, arc-lamps, or other lasers), chemical reactions (such as flash photolysis, or combustion), nuclear, as well as plasma pumps. Any method that delivers energy to the gain medium can be effective. In the case of pulse lasers there is often a desire to have a high power output, a technique known as Q-switching is often used to accomplish this end.

As noted above, the laser operates by exciting atoms in the gain medium such that the medium emits more light than it absorbs. This is called population inversion. However, there is a limit to the number of atoms that can be reasonably excited at one time. This is known as gain saturation. Q-switching is in the category of techniques whereby the laser is prevented from lasing until the gain medium has been fully gain saturated at which point the "switch" is thrown and the laser outputs a high energy pulse until the gain medium is depleted and needs to be recharged. If we are interested in creating ultra-short pulses, we need to make use of a technique called mode-locking. The concept of mode-locking is to align the phases of the longitudinal modes such that the phases all peak together at the same time and place. This causes them to constructively interfere for a short time,
and destructively interfere elsewhere. There are a few ways to implement a mode-locked system that fall into two categories: active and passive. We are primarily interested in the passive forms of mode-locking - specifically, the Kerr lens. A Kerr lens is a special kind of dielectric material whose index of refraction, \( \tilde{n} \), has a nonlinearity such that it depends on the intensity (the square of the magnitude of the electric field) of the beam passing through the lens. The nonlinearity in \( \tilde{n} \) has a self-focusing effect on the beam so that the power is focused more and more within a shorter time frame until it reaches some limit where the beam cannot be focused anymore. This limit is imposed by dispersion effects in the material as well as physical constraints on the bandwidth of the gain medium. Such effects result in the nonlinear Schrödinger equation (NLSE) appearing as a description of the wave propagating through an optical fiber. For certain choices of parameters, the solution of this equation will be given by a solitary wave (soliton). Since we are interested in studying these phenomena, we will now briefly outline of how the NLSE arises in these systems.

We can write the index of refraction as

\[
\tilde{n}(z, \omega) = \tilde{n}(\omega) + n_2 |E|^2 + \frac{i\tilde{\alpha}}{2k_0} \tag{5}
\]

\[
\equiv \tilde{n}(\omega) + \Delta n \tag{6}
\]

where the tilde is used to denote functions of frequency \( \omega \) (as opposed to functions of time \( t \)). In other words, \( \tilde{n}(z, \omega) \) is the time domain Fourier transform of \( n(z, t) \). \( \tilde{\alpha} \) corresponds to the imaginary part of the electric susceptibility \( \chi \) and is called the absorption coefficient.

We are interested in describing the electromagnetic wave propagating through an optical fiber with a nonlinear index of refraction given by equation (5). Therefore we need to solve the Helmholtz equation:
\[ \nabla^2 \tilde{E} - \tilde{n}(z, \omega)^2 k_0^2 \tilde{E} = 0 \]  

where \( \tilde{E} \) denotes the time domain Fourier transform of the electric field, \( E(t, z) \). We suppose that the magnitude of \( \Delta n \) is small relative to \( \tilde{n} \) so that we can make the following approximation:

\[ \tilde{n}(z, \omega)^2 \approx \tilde{n}(\omega)^2 + 2\tilde{n}(\omega)\Delta \tilde{n} \]

We take the following ansatz solution to the Helmholtz equation:

\[ \tilde{E}(z, \omega) = F(x, y)[\tilde{A}(z, \omega - \omega_0)e^{i\beta_0 z}] \]  

Here, \( \tilde{A} \) is a slowly varying function of \( z \) and \( \beta_0 \). \( F(x, y) \) is a function which takes into account properties of the fiber that only depend on the \( x \) and \( y \) coordinates. Combining Equations (7), (1.2), and (8) and separating variables leads to the equations:

\[ \frac{\partial^2}{\partial x^2} F + \frac{\partial^2}{\partial y^2} F + \left( \frac{\tilde{n}^2(z, \omega) \omega^2}{c^2} - \tilde{\beta}^2 \right) F = 0 \]  

\[ 2i\beta_0 \frac{\partial}{\partial z} \tilde{A} + (\tilde{\beta}^2 - \beta_0^2) \tilde{A} = 0 \]

Where \( \tilde{\beta} \) is given by

\[ \tilde{\beta}(\omega) = \beta(\omega) + \Delta \beta \]

and \( \Delta \beta \) is:

\[ \Delta \beta = \frac{k_0 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Delta n_1 |F(x, y)|^2 dxdy}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |F(x, y)|^2 dxdy} \]
Since we are interested in constructing the nonlinear Schrödinger equation we can put the equation for $A$ into a more suitable form:

$$\frac{\partial \tilde{A}}{\partial z} = i(\beta(\omega) + \Delta \beta - \beta_0)\tilde{A}$$

$\beta(\omega)$ can be expanded according to a Taylor expansion:

$$\beta(\omega) = \beta_0 + (\omega - \omega_0)\beta_1 + \frac{1}{2}(\omega - \omega_0)^2\beta_2 + \frac{1}{6}(\omega - \omega_0)^3\beta_3 + ...$$

In this case, we are only interested in the second order Taylor expansion. We convert this result to the time domain with an inverse Fourier transform. After this operation, the $\omega - \omega_0$ terms become $i\frac{\partial}{\partial t}$ terms. This is where the time derivatives will come from. Finally, we can unpack the $\Delta \beta$ and arrive at the solution ([26], [27]):

$$\frac{\partial A}{\partial z} + \beta_1 \frac{\partial A}{\partial t} + i\beta_2 \frac{\partial^2 A}{2\partial t^2} + iA = i\gamma|A|^2A$$

with

$$\gamma = \frac{n_2\omega_0}{cA_{eff}}$$

and $A_{eff}$ given by:

$$A_{eff} = \left(\frac{\int \int |F(x,y)|^2dx\,dy}{\int \int |F(x,y)|^4dx\,dy}\right)^2$$

And this gives us the NLSE, which we will study closely later in the thesis. In particular, we will discuss properties of this system and show how the geometric minimum action method can be used to model and study it.
2 Stochastic Modeling

2.1 Stochastic Differential Equations

As mentioned in the introduction, stochastic models have an interesting history in mathematics, physics, and finance. There have been a number of books written that discuss stochastic differential equations (SDEs) from both an analytical and a numerical point of view [28-30]. As mentioned before, on particular problem we are concerned with is to find the likeliest transition path between two stable states in a stochastic system described by a stochastic differential equation. Specifically, we are interested in SDEs that are driven by Brownian motion. Such SDEs take the form

\[
    dx(t) = b(x(t))dt + \sigma(x(t))dW(t).
\]

Here, \(x\) is the vector that we wish to solve for. It has an arbitrary number of components in principle, or could even be a field in the context of stochastic partial differential equations. For now, we will consider the finite-dimensional case. SDEs are characterized by a deterministic component and a random component. The vector field \(b(x)\) represents the deterministic component of the system’s dynamics. It is often called the drift vector. The random component is given by the term \(\sigma(x(t))dW(t)\). Here, \(dW\) denotes a vector of independent Brownian increments. In this work we will assume that the dimension of \(dW\) is the same as the dimension of the \(x\) vector. Sometimes, the corresponding stochastic process \(W(t)\) is called a Wiener process and \(dW\) represents its differential increment that can be related to white noise. We will address the meaning of these terms in a moment. The matrix \(\sigma\) corresponds to the diffusion in the system and represents the strength of the random component. In the general case, \(\sigma\) can depend on \(x\).

For now, let us look at the deterministic aspects of this system. Since we are going to
be considering cases where noise is small, we can get a good idea of the dynamics of these systems by temporarily ignoring the stochastic component and just computing the solution to the deterministic system given by

\[ \dot{x} = b(x). \]  

This is simply an ordinary differential equation. Let us take a look at some examples of drift vectors to see how such systems behave. Consider the following ODE in one dimension

\[ \dot{x} = -\alpha x. \]  

This drift vector \( b = -\alpha x \) is, in the context of SDEs, called Ornstein-Uhlenbeck problem [31]. Typically, we are going to be looking at problems with more than one spatial dimension and the above Ornstein-Uhlenbeck drift can be easily generalized to systems with
higher dimensionality. In two spatial dimensions, we can write this generalization of eq. (13) explicitly, and we get:

\[
\begin{bmatrix}
\dot{u} \\
\dot{v}
\end{bmatrix} = \begin{bmatrix}
-\alpha u \\
-\alpha v
\end{bmatrix}
\]  (14)

In principle, we could also choose different decay constants \( \alpha_1 \) and \( \alpha_1 \) for the different directions, but let us assume for now that both constants are the same. Examining this simple equation reveals the idea that whatever place I choose as my starting position, my particle is going to be drawn towards the origin with a total velocity \( V_t = \alpha \sqrt{u^2 + v^2} \). This implies that once my particle reaches the origin, it no longer moves. Therefore we can say that the origin represents a stable fixed point corresponding to the drift vector \( b = -\alpha x \).

Another possible drift vector we can look at corresponds to counterclockwise circular motion.
Figure 3: Velocity field corresponding to the combined circular and Ornstein-Uhlenbeck drift vector $\dot{u} = -\sin(\theta) - \alpha u$, $\dot{v} = \cos(\theta) - \alpha v$. In this case $\alpha = 1$. All initial conditions lead to particle having a spiral trajectory towards origin.

\[
\begin{bmatrix}
\dot{u} \\
\dot{v}
\end{bmatrix} =
\begin{bmatrix}
-\sin(\theta) \\
\cos(\theta)
\end{bmatrix}
\]  

(15)

Here $\theta$ is the angle measured from the positive x axis. We can convert from polar coordinates into Cartesian to make the left hand side consistent with the right hand side of the equation. So $\cos(\theta) = \frac{u}{\sqrt{u^2+v^2}}$ and $\sin(\theta) = \frac{v}{\sqrt{u^2+v^2}}$, so $b(\theta)$ becomes:

\[
\begin{bmatrix}
\dot{u} \\
\dot{v}
\end{bmatrix} =
\begin{bmatrix}
-\frac{v}{\sqrt{u^2+v^2}} \\
\frac{u}{\sqrt{u^2+v^2}}
\end{bmatrix}
\]  

(16)

In this case nothing is drawn either towards or away from the origin.

As it is, there is no stable fixed point. We can imagine that if we put in some additive noise, that the particle would be kicked from radius to radius where it would encircle the origin as if in orbit. However, we can imagine ways to alter our system such that we do have a fixed point: For example, we can combine the drift vectors of the previous two
Figure 4: Velocity field corresponding to the combined Maier-Stein drift vector $\dot{u} = u - u^3 - \beta uv^2$, $\dot{v} = -(1 + u^2)v$. In this case $\beta = 1$. There are two real stable fixed points at (1,0) and (-1,0).

examples.

$$\begin{bmatrix} \dot{u} \\ \dot{v} \end{bmatrix} = \begin{bmatrix} -\frac{v}{\sqrt{u^2 + v^2}} - \alpha u \\ \frac{u}{\sqrt{u^2 + v^2}} - \alpha v \end{bmatrix}$$

(17)

One can see in figure (3) that the combined system is can be interpreted as a superposition of the fields in figures (1) and (2).

A particle placed anywhere will spiral downward counterclockwise towards the location of the stable fixed point (here in the origin). If we imagine adding some stochastic noise to this system, we can see that wherever the noise kicks the particle, for sufficiently small noise, it will "eventually" be drawn back to the origin (eventually, in this case, is an infinite amount of time). An example of particular interest to us is the following drift vector:

$$\begin{bmatrix} \dot{u} \\ \dot{v} \end{bmatrix} = \begin{bmatrix} u - u^3 - \beta uv^2 \\ -(1 + u^2)v \end{bmatrix}$$

(18)

In the context of SDEs this is called the Maier-Stein system [32]. This system of ODEs is a bit more complicated than the ones we have looked at so far. However, it is easy
to see that, in this case, there are actually two stable fixed points where the particle can be drawn toward depending on the initial condition of the particle. Additionally, if we add stochastic noise to the system, there is a possibility that our particle can move from one stable fixed point to the other. This corresponds to a state transition, and represents the particular phenomenon that we are interested in understanding. In particular, we are interested in answering the following question: "For sufficiently small noise, what is the most likely transition path between these two stable states?"

Note that, in principle, we can consider any system from classical mechanics and add random perturbations. For example, consider the classical pendulum. The classical pendulum conserves energy and is therefore a Hamiltonian system. Setting some constants equal to 1, its Hamiltonian is written as

$$H(q, p) = \frac{1}{2} p^2 - \cos(q)$$

(19)

We know from classical mechanics that the equations of motion for a system like this are given by Hamilton’s equations:

$$\dot{q} = \frac{\partial H}{\partial p}, \quad \dot{p} = -\frac{\partial H}{\partial q}$$

(20)

For the pendulum, these equations are

$$\begin{bmatrix} \dot{q} \\ \dot{p} \end{bmatrix} = \begin{bmatrix} p \\ -\sin(q) \end{bmatrix}.$$  

(21)

The phase space associated with this problem is similar to our second example. We can add the Ornstein-Uhlenbeck drift term to the equations of motion to cause the pendulum...
George Poppe

Applications of gMAM

Figure 5: Velocity field corresponding to the "classical pendulum" drift vector with some dissipation caused by an Ornstein-Uhlenbeck term. Namely: $\dot{u} = v$, $\dot{v} = -\sin(u) - \alpha v$. There are stable fixed points at any position $(2\pi n, 0)$ with $n$ being any integer value.

to lose energy just as we had done before. Thus we will have a Hamiltonian system with added dissipation. Rewriting these equations of motion replacing $(q,p)$ with $(u,v)$, we get:

$$
\begin{bmatrix}
\dot{u} \\
\dot{v}
\end{bmatrix} =
\begin{bmatrix}
v \\
-\sin(u) - \alpha v
\end{bmatrix}.
$$

This drift vector gives a similar spiral shape as our last example, however this system has multiple fixed points. At every period along the x-axis ($u = 2\pi n$, $v = 0$), there is a stable fixed point into which a particle may be drawn. Thus, as with the Maier-Stein case, since we have more than one stable fixed point we can talk sensibly about the concept of a state transition from one fixed point and the other if the system has sufficiently small random noise.

Forgoing the discussion of a precise mathematical meaning of the term "sufficiently small noise" for now, let us switch gears and learn a more about the specific properties of noise. The easiest place to start is with the idea of a Wiener process (also commonly referred to as a random walk or Brownian motion). The Wiener process was invented
as a mathematical construction many years after the discovery of Brownian motion (the physical observation which the Wiener process aimed to model). The empirical fact that was observed by Robert Brown [33] is essentially that, when a small particle is placed in water, it moves about with an unpredictable motion. The physical explanation for what was happening (unknown to Robert Brown) is that water molecules (whose velocities are given by the probability distribution associated with the temperature of the water) are bumping into the particle and transferring momentum to it.

Because the Wiener process is modeling a statistical system where the precise micro-state is unknown, it must itself be probabilistic in nature. Usually we write

\[ X(t) = W(t) = \int_0^t dW(t). \]  

(23)

Here, \( W(t) \) is called a Wiener process. In order to be considered a Wiener process, \( W(t) \) must have both the property that \( W(t + \Delta t) - W(t) \) is independent of \( W(t) - W(t - \Delta t) \), furthermore, for arbitrary \( t \), the probability distribution of \( W(t + \Delta t) - W(t) \) must be a Normal distribution centered at 0. In particular in the physics literature, it is common to consider the case that \( \frac{dW}{dt} \) is white noise henceforth to be represented by the Greek letter \( \xi \).

The above demands that the function \( \frac{dW}{dt} \) must have values that are distributed according to a normal probability distribution. Therefore one can imagine that for each differential time step \( \Delta t \), \( dW(t) \) draws a random number that is uncorrelated to the previous number drawn (although there have been attempts at modeling correlated noise as well [34]). Hence the time derivative of \( W(t) \) is discontinuous. This implies that \( W(t) \) actually is not differentiable. Still, we will often write the time derivative of \( W \) as a shorthand with this understanding in mind.

Just as we examined SDEs with only the deterministic part (the drift vector \( b \)), let us look
at some SDEs where there is only a stochastic part hence

\[ \dot{x}(t) = \sigma(x(t)) \xi(t) , \]  

(24)

where \( \xi(t) \) is a white noise process. If we think of a two dimensional case, \( \xi(t) \) is generally going to represent a two component vector that has independent white noise processes for each dimension.

\[ \xi(t) = \begin{bmatrix} \xi_u(t) \\ \xi_v(t) \end{bmatrix} . \]

(25)

Other than that, is not much more to say about \( \xi(t) \). \( \sigma(x) \), on the other hand has some important degrees of freedom that we can discuss. \( \sigma \) is a matrix that is generally defined in terms of the diffusion tensor \( a = \sigma \sigma^T \). As far as we are concerned for now, it can be interpreted as the strength of the stochastic part of the equation. The simplest case is where \( \sigma \) is some constant times the identity matrix.

\[ \sigma = A \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} . \]

(26)

This \( \sigma \) corresponds to the case where the stochastic aspects of the two dimensions are independent of each other and are independent of space. We can also consider a different system by linking the white noise processes of the two dimensions:

\[ \sigma = A \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} . \]

(27)

This \( \sigma \) effectively constrains the system so that the \( u \)-dimension is always going to feel the same stochastic kick as the \( v \)-dimension. In this way the stochastic aspects of the system are no longer independent from each other, hence we have correlation between
the components.

The last example we are going to discuss for now contains one additional complication. The matrix $\sigma$ can be, in principle, a function of $x$. There are situations where the strength of the randomness is going to depend on the state of the system. When this happens, the system is said to have "multiplicative noise" because the state of the system $x$ is multiplying the noise terms. This presents an interesting ambiguity in the meaning of the noise. There are two common interpretations, the Ito and the Stratonovich. For the purpose of this paper, we are going to assume the Ito interpretation. For example, consider that $x$ represents temperature of a system. As the temperature of a system goes to zero, it becomes more and more susceptible to being affected by random noise. Even the smallest bit of energy has a large effect on the temperature when it nears 0 Kelvin.

In the spirit of that idea, consider:

$$\sigma = \frac{A}{\sqrt{u^2 + v^2}} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$  \hspace{1cm} (28)

This $\sigma$ has the effect of keeping the particle from staying near zero for any extended period of time. As the particle approaches zero, the noise approaches infinity which will tend to kick the particle far away. In this way we see that having a position dependent noise is going to bias the particles position such that it tends to stay in areas where noise is small.

### 2.2 The Fokker-Planck Equation

In the previous section, we addressed the occurrence of randomness in terms of stochastic processes, as solutions of stochastic differential equations. Alternatively, we can consider also the evolution of the probability density of the stochastic process. For this pur-
pose, it is easiest to look at the Fokker-Planck form of an SDE. It is always possible to write an equation associated with an SDE like equation (11) as a Fokker-Planck equation in terms of the associated probability density. The general form of the one-dimensional Fokker-Planck equation is as follows:

$$\frac{\partial p}{\partial t} = -\frac{\partial}{\partial x} [b(x,t)p(x,t)] + \frac{\partial^2}{\partial x^2} \left( \frac{\sigma^2}{2} p(x,t) \right)$$  \hspace{1cm} (29)$$

Here $p(x,t)$ represents the probability density that the particle (or system) will be found at position $x$ and time $t$. The functions $b(x,t)$ and $\sigma(x,t)$ correspond to the same drift vector and diffusion matrix from Equation (11). It is beyond the scope of this paper to prove that in fact (29) describes the same dynamics as (11). Instead, we refer to [35-37].

The best way to understand the meaning of this equation, is to look at the behavior of simple examples as we did in the previous section. Let us start by assuming that there is no stochastic effect on the system, i.e. $\sigma = 0$. Additionally, for simplicity, we assume a constant drift vector $b(x,t) = c$. The Fokker-Planck equation reduces then to a simple first-order partial differential equation

$$\frac{\partial p(x,t)}{\partial t} = -\frac{\partial}{\partial x} [c p(x,t)]$$  \hspace{1cm} (30)$$

$$p_t(x,t) = -cp_x(x,t)$$  \hspace{1cm} (31)$$

Remember, according to equation (11), this drift vector means that the particle is going to move along the $x$-axis with a constant velocity given by $c$. If we know the exact starting position of the particle, let us say at $x_0$, this means the probability density corresponding to the initial condition is going to be a delta function centered at $x_0$. In other words, $p(x,0) = \delta(x - x_0)$. In many cases, we can assume without loss of generality $x_0 = 0$ for simplicity. Since there is no noise in the system, the particle is going to move deterministically, which
means that the shape of the probability distribution is going to stay as a delta function: The
delta function moves along x axis with velocity given by $c$. This intuition is also reflected
in the systematic solution of the above Fokker-Planck equation. Clearly, the solution of
the above equation is given by

$$p(x,t) = A \cdot h(x - ct). \quad (32)$$

Here $h$ is some function of $x - ct$. Since we know that the initial condition is a delta function
of $x$ (and as we discussed, we already know how the solution should look), we need to
pick

$$p(x,t) = A \cdot \delta(x - ct), \quad (33)$$

and the constant $A$ needs to be one because of normalization. If we take the derivatives
on this function we will see that it satisfies equation (30). When only the deterministic
component is active, a delta function initial condition is going to mimic a particle’s traject-
ory. This makes solving the purely deterministic case a bit easier.

Consider again Equation (13) that called the Ornstein-Uhlenbeck drift vector. This will
give us the corresponding Fokker-Planck equation:

$$\frac{\partial p}{\partial t} = -\frac{\partial}{\partial x}[(-\alpha x)p(x,t)]$$

Let us assume we are taking a one-dimensional case for the sake of expediency. Taking
the derivative, we get:

$$p_t = \alpha(p + xp_x)$$

as before we may wish to set our initial condition to be a delta function in position. i.e.
$p(x,0) = A\delta(x-x_0)$ (This time we do not want $x_0 = 0$, we want to start somewhere far away
from the fixed point). Again assuming that we are thinking of a one-dimensional system,
Figure 6: Displays the probability density $p = p(x,t)$ as a function of space $x$ and time $t$. This is a numerical simulation of the Fokker-Planck equation where only the deterministic component ($b = -0.01x$) is active. The distribution is initialized to an approximate delta function centered at $x = 10$. As time goes on, the peak of the distribution moves from $x = 10$ at $t = 0$ (background) to $x = 9.89$ at $t = .4$ (foreground) due to the deterministic effects of the Ornstein-Uhlenbeck drift vector.

we already have a good idea of what is going to happen from looking at the vector fields.

If I place a delta function anywhere in space, for the Ornstein-Uhlenbeck drift vector $b = -\alpha x$, it is going to move towards the origin with a velocity equal to $-\alpha x$. Now, if we were talking about the probability distribution as an initial condition of the Fokker-Planck equation, it would not be so simple because the position dependence of $b$ tends to squeeze the probability distribution together. However, since we are talking about a delta function, all the squeezing has already been done; it behaves as a particle and thus we can solve Equation (13) directly. The delta function must propagate according to:

$$x(t) = x_0 e^{-\alpha t}$$  \hspace{1cm} (34)
George Poppe

Applications of gMAM

Figure 7: Displays the probability density \( p = p(x,t) \) as a function of space \( x \) and time \( t \). This is a numerical simulation of the Fokker-Planck equation with \( b = 0 \) and \( \sigma = 0.1 \). The distribution initial begins approximately as a delta function (background) and broadens (foreground) due to the stochastic effects of the Wiener process.

and therefore the description of our probability distribution must look like:

\[
p(x,t) = \delta (x - x_0 e^{-\alpha t})
\]

(35)

Let us look now at the other special case where \( b(x,t) = 0 \) and we have only stochastic effects controlling our system. Say that \( \sigma \) is a constant that does not depend on \( x \). Then we get:

\[
dx(t) = \sigma dW(t)
\]

(36)

which gives us the Fokker-Planck equation:

\[
\frac{\partial p}{\partial t} = \frac{\sigma^2}{2} \frac{\partial^2 p}{\partial x^2}.
\]

(37)

This equation, which corresponds to Brownian motion in our SDE, is famously known as the heat equation. The solution to this is well known [38]. Assuming that we again make
our initial condition a delta function $p(x,0) = \delta(x-x_0)$, one can show that the probability that, starting from position $x_0 = 0$ and time $t_0 = 0$ that the particle will be found at position $x$ at time $t$ is

$$p(x,t) = \frac{1}{\sqrt{2\pi \sigma^2 t}} \exp\left(\frac{-x^2}{2\sigma^2 t}\right)$$

(38)

This solution satisfies the initial condition in the sense that it approaches a delta function as $t$ tends to zero. As time passes, the delta function broadens into a Gaussian probability distribution. This equation essentially describes the fact that the variance of the distribution is a linear function of time, more precisely equal to $\sigma^2 t$:

$$\text{var}(x) = \int x^2 p(x,t) \, dx = \sigma^2 t$$

One more example we can think about briefly is the case where $\sigma$ is a function of $x$. If, for example we have:

$$\sigma = \kappa x$$

(39)

where $\kappa$ is some positive real constant and $x$ is the position, our experience with the Brownian motion case ($b = 0$, $\sigma = \kappa$) as contrasted with the deterministic Ornstein-Uhlenbeck case ($\sigma = 0$, $b = -\alpha x$) tells us that the higher the value of $\sigma$ is, the more broadening is going to occur in the probability distribution. That means that when $\sigma = \kappa x$, the regions where $x$ is higher are going to suffer greater dispersion from the stochastic process. So the symmetry of our initial distribution will be broken and the part that is further from zero will be broader than the part that is closer to zero. Looking at the Fokker-Planck Equation, we get:

$$\dot{p} = \kappa^2 (p + 2xp_x + x^2 p_{xx})$$

(40)

This equation can be solved by making use of spectral methods. It corresponds to a special case of a system called Geometric Brownian Motion. It gets its name because
of the fact that the logarithm of the process is Brownian [39]. This type of system is commonly used in stock market models. The equation is solved by:

$$ p(x,t) = \frac{1}{x\sigma\sqrt{2\pi t}} \exp\left( -\frac{\ln\left(\frac{x}{x_0}\right) + \frac{1}{2}\sigma^2 t}{2\sigma^2 t} \right) $$ \hspace{1cm} (41)

In our discussion thus far we have only been looking at drift and diffusion terms of the Fokker-Planck equation separately. A simple example of a case where both terms are non-zero is the full Ornstein-Uhlenbeck system. This combines the deterministic and stochastic aspects of Equation (11). Recall that in this case we have $b = -\alpha x$ and now we will add in the stochastic portion: $\sigma = \text{const.}$ Plugging this into the Fokker-Planck equation:

$$ \dot{p} = \alpha(xp)_x + \frac{\sigma^2}{2} p_{xx} $$ \hspace{1cm} (42)

Let us again assume our initial condition to be a delta function as in previous cases $p(x,0) = \delta(x-x_0)$. If we take the Fourier transform where $\hat{p}(\lambda,t)$ represents the Fourier-transformed variable associated with $p(x,t)$ we find as equation for $\hat{p}$

$$ \hat{p}_t + \alpha \lambda \hat{p}_\lambda = -\frac{\sigma^2 \lambda^2}{2} \hat{p} $$ \hspace{1cm} (43)

This equation can be solved explicitly using the Method of Characteristics [40]. Consider a curve $\lambda = \lambda(t)$ and write

$$ \hat{p}_t = \hat{p}_t + \frac{\partial p}{\partial \lambda} \frac{d\lambda}{dt} = \hat{p}_t + \alpha \lambda \hat{p}_\lambda $$ \hspace{1cm} (44)

Therefore, we find as equation for $\lambda$:

$$ \frac{\partial \lambda}{\partial t} = \alpha \lambda $$ \hspace{1cm} (45)
Figure 8: Displays the probability density for the full Ornstein-Uhlenbeck system as a function of time. System begins in a near-delta function centered at position $x = 10$ (background). The system evolves in time such that the width of the distribution increases and the location of the peak drifts towards the origin ending at $x = 9.05$.

and this equation tells us that $\lambda$ is an exponential function of the form $\lambda = Be^{\alpha t}$. Plugging that in yields

$$\tilde{\rho}_t = -\frac{\sigma^2 B^2}{2} e^{2\alpha t} \tilde{\rho}$$

(46)

Now, since we are assuming a delta function as our initial condition in the position space, in the Fourier space we have $p(\lambda, 0) = e^{-i\lambda x_0}$. Acknowledging that, direct integration of this equation yields for $\tilde{\rho}$:

$$\tilde{\rho}(t, \lambda) = \tilde{\rho}(\lambda, 0) \exp \left( -\int_0^t \frac{\sigma^2 B^2}{2} e^{2\alpha t'} dt' \right) = \exp \left( -i\lambda x_0 - \frac{\lambda^2 \sigma^2}{2\alpha} e^{-2\alpha t (e^{2\alpha \alpha t} - 1)} \right)$$

(47)

Finally we take the Inverse Fourier Transform on $\tilde{\rho}$ and arrive at the solution for the probability density for the Ornstein-Uhlenbeck system [31, 41]:

$$p(t, x) = \left( \sqrt{\frac{2\pi \sigma^2}{2\alpha} e^{-2\alpha (e^{2\alpha t} - 1)}} \right)^{-1} \left( \exp \frac{(x - x_0 e^{-\alpha t})^2}{\sigma^2 e^{-2\alpha (e^{2\alpha t} - 1)}} \right)$$

(48)
2.3 Path Integrals

One possible avenue to understanding SDEs is through the concept of a path integral [42-45]. Many readers will be familiar with the concept from numerous fields including quantum mechanics [38], quantum electrodynamics [46], polymer physics, financial markets [8], optics [47], etc. [48]. The idea behind the path integral is very general so it is not surprising that we see the path integral appear in a wide variety of applications.

If we look at a slightly modified form (Langevin form) of Equation (11):

\[ \dot{x}(t) = b(x(t)) + \sigma(x(t))\zeta(t) \]  

(49)

We can see that the velocity \( \dot{x} \) is governed by a deterministic component \( b \) and a stochastic component \( \sigma \zeta \). Consider that \( \zeta \equiv \frac{dW}{dt} \) is a standard white noise process. A standard white noise process is a process which corresponds to a normal probability distribution with zero mean. As we discussed this also implies that the noise is delta correlated in time:

\[ \langle \zeta(t)\zeta(t+\tau) \rangle = \sigma^2 \delta(\tau) \]  

(50)

Essentially this means that for each differential time step, the \( \zeta \) process chooses a random number that is completely independent of the number it chose in any other differential time step. So, if we imagine that we are talking about a particle undergoing Brownian motion, we can imagine that each time step, the particle gets a small "kick" from the \( \zeta \) process and has an instantaneous change in its velocity. If we look at an individual time step \( \Delta t \) the probability distribution of "kicks" is given by:

\[ P(\Delta W) = A \exp \left[ -\frac{\Delta W^2}{2\Delta \tau} \right] \]  

(51)

Knowing this, we can think of a good trick. Let us rewrite Equation (49) as
One can see that by considering Equations (51) and (52) we can now relate our change in position $\Delta x$ and time $\Delta t$ to our probability distribution by substitution:

$$P(\Delta W) = \exp \left[ -\frac{1}{2\sigma^2} \left( \frac{\Delta x_k}{\Delta t} - b(x_k,t_k) \right)^2 \Delta t \right]$$

Here, we have the probability that the particle is kicked a small amount $\pm \Delta W$ in the $x$ direction. Our goal is to frame this as a path integral, so we can then ask: What is the probability that the particle takes a particular path going from point $x_0$ at time $t = 0$ to point $x_N$ at time $t_N$? We know that the "kick" in each time step is supposed to be independent of all the other kicks at all the other times. Therefore, the probability of taking a particular path is going to be equal to the product of the probabilities that the particle will move incrementally along that path.

$$P_{path}(x_N,t_N|x_0,0) = P(x_N,t_N|x_{N-1},t_{N-1})P(x_{N-1},t_{N-1}|x_{N-2},t_{N-2})...P(x_1,t_1|x_0,0)$$

$$= A \exp \left[ \frac{(\dot{x}_N - b_N)^2}{2\sigma^2} \Delta t \right] \ast \exp \left[ \frac{(\dot{x}_{N-1} - b_{N-1})^2}{2\sigma^2} \Delta t \right] \ast \cdots \ast \exp \left[ \frac{(\dot{x}_1 - b_1)^2}{2\sigma^2} \Delta t \right]$$

$$= A \exp \left[ \sum_n \frac{(\dot{x}_n - b_n)^2}{2\sigma^2} \Delta t \right]$$

Here we have replaced the $\frac{\Delta x}{\Delta t}$ with an $\dot{x}$ for visual simplicity. Now that we have the probability of taking a particular path, we can write an expression that gives the total probability that the particle takes any path that begins at point $(x_0,t_0)$ and ends at point $(x_N,t_N)$. It is easy to see that this total probability is given by the sum of the probabilities of the individual paths.
\[ P(x_N, t_N | x_0, 0) = \sum_{\text{paths}} P_{\text{path}}(x_N, t_N, x_0, 0) = \int_{-\infty}^{\infty} dx_{N-1} \int_{-\infty}^{\infty} dx_{N-2} \cdots \int_{-\infty}^{\infty} dx_1 A \exp \left[ -\sum_{i=1}^{N} \frac{(\dot{x}_i - b)^2}{2\sigma^2} \right] \equiv \int_{C(x,t|x_0,0)} Dx e^{-S} \]

Where in the last step we have let \( \tau \) go to the continuum limit and written the probability as a classical path integral. Under this notation, \( S \) must be:

\[ S = \int_0^t \frac{1}{2\sigma^2} (\dot{x} - b(x, \tau))^2 d\tau \] (53)

If our action is given by Equation (53), that means we can define a Lagrangian that corresponds to the integrand of that equation. Namely:

\[ L(x, \dot{x}, \tau) = \frac{1}{2\sigma^2} (\dot{x} - b(x, \tau))^2 \] (54)

The appearance of an effective Lagrangian is very fortuitous for us. The implication is that maximizing the probability of our path integral implies minimizing the action associated with the Lagrangian. Remember that we are considering rare state transitions on systems with small noise. Intuitively, if a transition almost never happens under the effects of this small noise, then it stands to reason that when transitions do occur, they’re likely going to fall very near to the trajectory that maximizes the probability of transition (the minimum action trajectory). Trajectories far away from the classical trajectory will have transitions in the path integral whose probability approaches zero and therefore will be much less likely to occur than the transitions that correspond to the action minimizing path. Thus
for sufficiently rare transitions with sufficiently small noise, we should be able to say that there is a very high probability that the transition is going to fall nearby the classical path. This argument is made rigorous in the Freidlin-Wentzell theory of large deviations which we discuss later in the thesis. We can minimize the action by utilizing the Euler-Lagrange equations.

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} = 0$$

Following the reasoning above, this minimum action path can be used as an approximation to the trajectory of the particle during a rare transition. This approximation is only valid for small levels of noise, although there are certain special cases where it will correctly predict the exact solution. Let us have a look at those cases.

The first one we are going to look at is the case of simple Brownian motion. In this case the drift vector from Equation (11) is given by $b = 0$, and $\sigma = \text{const.}$. We already know what the answer should look like having solved it using the Fokker-Planck equation. This gives us a chance to verify that the path integral method is actually functional.

Starting from the Lagrangian with $b = 0$

$$L = \frac{1}{2\sigma^2} \dot{x}^2$$

(55)

This Lagrangian corresponds to the action in Equation (53) and can be used to find the most likely trajectory from initial point $x=a$ to final point $x=b$. If we plug this Lagrangian into the Euler-Lagrange Equation, we get the equation of motion corresponding to the minimum action:

$$\ddot{x} = 0$$
We can integrate this equation twice to find the function \( x(t) \).

\[
\dot{x}(t) = c \\
x(t) = ct
\]

In other words, the particle moves in a straight line with constant velocity between the two points. This corresponds exactly to the solution we found using the Fokker-Planck equation (Equation (33)). Many will view this result as somewhat trivial, but what is important is that it is the correct result.

Now that we have some experience, let us consider what is called an "Escape Problem". Imagine there exists a particle trapped in a potential well that is subjected to a thermal bath. The presence of a thermal bath implies that the particle’s velocity will correspond to Brownian motion. Suppose further that the potential well that traps the particle represents a metastable state. The term metastable is a condition on the relative strength of the random motion versus the height of the potential well. It implies that at some finite time in the future, the particle will leave the well and find a new equilibrium in some other state. One example of such a situation can be modeled by the Ornstein Uhlenbeck process. We can define "escape" by imagining a that there is a certain position such that once the particle touches that point, it will leave the well that it is trapped in. Our SDE tells us:

\[
\Delta W(t) = \frac{1}{\sigma} \left( \frac{\Delta x(t)}{\Delta t} + kx \right)
\]

One thing that will become apparent now that might not have been apparent before is that this process has similarities to a classical harmonic oscillator. The Lagrangian corresponding to (54) becomes:
Taking this Lagrangian, we can attempt to find the equations of motion that minimize the action by applying the Euler-Lagrange equation:

\[
\frac{\partial L}{\partial x} = \frac{1}{\sigma^2} \alpha^2 x \\
\frac{d}{dt} \frac{\partial L}{\partial \dot{x}} = \frac{1}{\sigma^2} (\ddot{x} + \alpha \dot{x})
\]

The equation of motion becomes:

\[
\ddot{x} = \alpha \dot{x} - \alpha^2 x
\]
at position \( x=a \), we get:

\[
x(\tau) = a e^{k\tau} \left( \frac{1 - e^{2k(T_{\text{min}} - \tau)}}{1 - e^{2kT_{\text{min}}}} \right)
\]

And with this, we have found the most likely exit path for the particle. Note that the term \( T_{\text{min}} \) appears in this solution. This term refers to the "start time" of the process. We assume that the process ends at time \( t=0 \) so that \( T_{\text{min}} < 0 \). If we don’t specify a valued for \( T_{\text{min}} \), we also need to minimize the action over all times. It turns out that the minimum among these is given by the path that corresponds to \( T_{\text{min}} \to \infty \). It is in this case that the classical path best characterizes the most likely escape trajectory.

Of course, in some cases solving the Lagrangian equations of motion will be difficult. It is good to keep in mind that if we have the Lagrangian, we can always transform to the Hamiltonian frame and solve the equations of motion using Hamilton’s equations. We make the transformation to the Hamiltonian frame with the Legendre transformation:

\[
H(x, p, t) = \dot{x}p - L(x, \dot{x}, t)
\]

Remember that the canonical momentum is defined as \( \frac{\partial L}{\partial \dot{x}} \), which in this case gives us \( p = \frac{1}{\sigma^2}(\dot{x} + \alpha x) \). If we temporarily assume \( \sigma = 1 \) for visual simplicity and apply the transformation, this will give us:

\[
H = bp + \frac{1}{2}p^2 = -\alpha xp + \frac{1}{2}p^2 \quad (62)
\]

The Hamiltonian will be familiar to all physicists. The Hamiltonian is defined as the object which obeys Hamilton’s Equations. Namely:
\[ \frac{\partial H}{\partial x_n} = -p_n \quad , \quad \frac{\partial H}{\partial p_n} = x_n \]  

(63)

In the case of our escape problem, the particle starts in a state with \( H = 0 \), so in order to conserve energy, we have two possible solutions. The momentum must be either \( p = -2b \) or \( p = 0 \). With that in mind, Hamilton’s equations tell us that

\[ \dot{x} = b + p = -b = \alpha x \]

Solving this we get that \( x(t) = Ae^{\alpha t} \). After that we can plug in initial conditions as usual. Figure (9) displays a comparison of the escape trajectory generated numerically via the filtering method (explained in section 4.1) vs the path calculated using the above assuming that the particle starts at \( x = 0 \), and the particle ends at \( x = 1 \).

3 Large Deviation Theory

3.1 The Freidlin-Wentzell Action Functional

One important set of concepts that we are going to be applying in the gMAM is the Freidlin-Wentzell theory of large deviations [49-51]. The primary result of interest tells us something about the trajectories that systems will take when the transition in question is very rare. Let us make this statement more precise by considering an SDE of the following form:

\[ dX^\varepsilon(t) = b(X^\varepsilon(t)) dt + \sqrt{\varepsilon} \sigma dW(t). \]  

(64)

Here, we assume that the noise is small and its strength is controlled through the factor \( \sim \sqrt{\varepsilon} \) with \( \varepsilon \ll 1 \). For a system given by eq. (64) we can look for a rate function \( I \)
associated with the solutions $X^\varepsilon(t)$. This rate function $I$ is defined on the Banach space $C_0 = C_0([0, T]; \mathbb{R}^d)$ as

$$I(\omega) = \frac{1}{2} \int_{-T}^0 |\dot{\omega} - b(\omega)|^2 \, dt.$$  

(65)

The intuitive meaning of rate function is related to the statistics of the solutions $X^\varepsilon(t)$ of the underlying SDE. If we attempt to characterize the probability, as a function of the number of trials, that a stochastic system will be found far from the most likely state (in other words, the probability that the state of a system will represent a rare event), then the rate function is the function that characterizes how quickly the probability decays. We can say that, for $\varepsilon$ and $\delta$ sufficiently small, in a $\delta$-neighborhood of a path $\psi$, the probability of deviations from this path is given by [20]

$$P_x \{ \sup_{0 \leq t \leq T} |X^\varepsilon(t) - \psi(t)| \leq \delta \} \sim e^{-\frac{1}{\varepsilon} I(\psi)}.$$  

(66)

Here, we additionally assume that the trajectory started at $t = 0$ at $\psi(0) = x$. From this, one can see the similarity this rate function has to the action that appeared in section 2.3 on path integral methods for SDEs. What this theorem essentially says is that for stochastic systems that occur on a real space where noise is small and the transition is rare, the paths which do not correspond to minimizer of this action $I$ are going to be exponentially less likely to occur. Figure 9 exemplifies this idea: The color map image shows that the frequency of sample paths dies off very quickly with distance from the action minimizing path. We can imagine that there is a small cylindrical tube whose local axis is given by the minimizing path such that nearly all transition trajectories will fall inside of the tube.

This idea is important for what we are aiming to study and there are a few techniques that follow from it that we will look at now.
To start, we consider the above rate function as an action and define it according to:

\[
S_T(\psi) = \begin{cases} 
\int_{-T}^{0} L(\psi, \dot{\psi}) \, dt & \text{if the integral converges} \\
+\infty & \text{otherwise}
\end{cases}
\]

Note that, in this integral, we have shifted the starting point of the trajectory to the time \(-T\) and are ending our path at the time \(t = 0\). While this is a mere question of definition, it will prove to be a convenient choice in what follows. The action \(S_T\) has the property that if a path is deterministically allowed for given drift vector \(b(X)\), then such a path has an action \(S_T = 0\).

The corresponding Lagrangian in this case is formally defined in terms of the Legendre Transform:

\[
L(x, y) = \sup_{\theta \in \mathbb{R}^n} \left( \langle \dot{x}, \theta \rangle - H(x, \theta) \right)
\]

The \(\langle , \rangle\) notation is used to represent a Euclidean inner product (a generalization of the dot product) between two vectors \(a\) and \(b\) (similar to the bra-ket notation of quantum mechanics). Remember that in general \(x\) is a vector of arbitrary dimension. Also, here the variable \(\theta\) is analogous to the canonical momentum. From here on, assume that \(\theta\) corresponds to the \(\theta\) which maximizes the function in Equation (67).

We will now show explicitly that this formulation represents a generalization of the path integral discussed before. In particular, we show that, given an SDE of the form of equation (64), we can write a corresponding Hamiltonian as:

\[
H(x, \theta) = \langle b, \theta \rangle + \frac{1}{2} \langle \theta, a\theta \rangle
\]

To prove this statement, we will show using a Legendre transformation that the corre-
The corresponding Lagrangian is given by
\[
L(x, \dot{x}) = \frac{1}{2} \langle \dot{x} - b(x), a^{-1}(\dot{x} - b(x)) \rangle \equiv \frac{1}{2} |\dot{x} - b|^2_a,
\]
with the norm \(|\cdot|_a\) defined via the inner product \(|f|_a^2 = \langle f, a^{-1} f \rangle\).

For this Hamiltonian, we find the equations of motion (similar to eqs.(63)) as:
\[
H_\theta(x, \theta) = \dot{x} = b + a \theta \tag{69}
\]
\[
H_\theta(x, \theta) = -\dot{\theta} = \langle b_n, \theta \rangle + \frac{1}{2} \langle \theta, a_n \theta \rangle \tag{70}
\]

So for a two-dimensional case we would have
\[
-H_x = \begin{bmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \end{bmatrix} = - \begin{bmatrix} H_{x_1} \\ H_{x_2} \end{bmatrix} = - (\nabla b)^T \theta - \frac{1}{2} \begin{bmatrix} \langle \theta, a_{x_1} \theta \rangle \\ \langle \theta, a_{x_2} \theta \rangle \end{bmatrix}, \quad \nabla b \equiv \begin{pmatrix} b_{1u} & b_{1v} \\ b_{2u} & b_{2v} \end{pmatrix} \tag{71}
\]

Carrying out the Legendre transformation on \(H(x, \theta)\) using eq. (69), we obtain indeed
\[
L(x, \dot{x}) = \langle \dot{x}, \theta \rangle - H(x, \theta) \quad \text{(Legendre Transform)}
\]
\[
= \langle \dot{x}, \theta \rangle - \langle b(x), \theta \rangle - \frac{1}{2} \langle \theta, a(x) \theta \rangle \\
= \sum_n \dot{x}_n \theta_n - \sum_n b_n \theta_n - \frac{1}{2} \sum_n \theta_n (\dot{x}_n - b_n) \quad \text{(from (69))}
\]
\[
= \sum_n \dot{x}_n \theta_n - \sum_n b_n \theta_n - \frac{1}{2} \sum_n \theta_n (\dot{x}_n - b_n) \\
= \frac{1}{2} \sum_n \theta_n (\dot{x}_n - b_n) \\
= \frac{1}{2} \sum_n (\dot{x}_n - b_n) \sum_m a_{mn}^{-1} (\dot{x}_m - b_m) \\
= \frac{1}{2} \langle \dot{x} - b(x), a^{-1}(\dot{x} - b(x)) \rangle \\
\equiv \frac{1}{2} |\dot{x} - b|^2_a
\]
This expression gives us the Lagrangian in terms of the generalized coordinate vectors $x$ and $\dot{x}$. Notice that it has the same form as the rate function, $I(\omega)$, in equation (65). This proves that the form of the Hamiltonian above correct for describing this system.

### 3.2 Instantons

The term instanton is simply a different term for the "classical path" or the minimizer of the Freidlin-Wentzell action. As shown in the previous section, given the law of large deviations, for rare transitions, the transition path is likely to be close to this the classical path. Instantons have been considered in the physics literature starting from stochastic equations written in Langevin form:

$$\dot{u} = b(u) + \eta, \quad (72)$$

where $\eta$ represents a random process that can be characterized as mean-zero Gaussian noise with correlation:

$$\langle \eta(x,t), \eta(x+r,t+\tau) \rangle = a(r)\delta(\tau). \quad (73)$$

Where in this instance the $\langle \cdot, \cdot \rangle$ operator represents the autocorrelation function. The $a(r)$ in this case is not quite the same thing as the diffusion tensor $a$ that we see in the Hamiltonian. The basic idea of an autocorrelation function is that it will give a high number if, on average, the function on the left side of the comma has similar values to the function on the right hand side of the comma. In the case of Equation (73) the correlation function is equal to a delta function in time. This means that there is no correlation between the value of the noise at present time $t$ and the value of noise at some other time either earlier or later than $t$. This is analogous to having independent dice rolls to decide the noise at every time $t$. Given an SDE like (72), suppose we want to calculate the probability given
some function $F(u)$, that at final time $t = 0$, starting from $t = -T$, we have $F(x) = A$, where $A$ is some arbitrary constant. Often, the function $F$ is called *observable*, and we are interested in the events $u$ such that $F(u(t = 0, x)) = A$. This condition on $u$ results in a particular boundary condition of the associated instanton equations: It can be shown (see [12]) that for this Lagrangian, there is a set of corresponding instanton equations given by:

$$
\dot{u} = b + a\theta, \quad \dot{\theta} = -(\nabla b)^T \theta + \lambda \nabla T F(u) \delta(t) \tag{74}
$$

These are, in principle, the same equations derived in the previous section using the diffusive Hamiltonian. For example, the first equation is clearly equivalent to eq. (69).

However, the last boundary term deserves a more detailed discussion. To see where this term comes from, we can begin by considering the path integral formulation discussed in section 2.3.

$$
P(x_N,t_N,x_0,0) = \int_{C(x_0 | u_0,0)} D\mathcal{X} e^{-S} \tag{75}
$$

If we take the expression we found for the Lagrangian at the end of last section, we can rewrite the action:

$$
S_T(x) = \frac{1}{2} \int_{-T}^{0} \langle \dot{u} - b, a^{-1}(\dot{u} - b) \rangle dt \tag{75}
$$

We should be able to find the probability that our system ends up in the final state of our choice by summing the probabilities of the possible system paths. However, if one of the paths is overwhelmingly more likely than all the others, then the probability that a transition will occur at all is about the same as the probability that the transition will occur given that the particle takes the most likely path. This is the basic idea of the instanton calculation.

There are a few ways that we can use to show that equation (74) is true. We also mention that, in the physics literature, one of them is most commonly known as the Martin-Sigma-
Rose (MSR) formalism. The other path is to follow the Freidlin-Wentzell method by considering large deviations. In the following we sketch briefly both methods and show that they yield equivalent results. First, we apply the MSR formalism, then we look at the Freidlin-Wentzell formulation of the problem.

**MSR formalism**

By a similar argument to the one used in section 2.3, we can write the expectation value of an arbitrary operator given equation (72) starting with the correlation function of $\eta$ given in equation (73).

The expectation value of some arbitrary function of $u$, $Y(u)$, is given by $\int Y(u')P(u')du'$, where $P(u')$ is the probability that I will find $u$ in a state $u = u'$. So, since we know the probability distribution of $\eta$ we can write the expectation value in terms of that probability distribution:

$$\langle Y(\eta) \rangle = \int \mathcal{D} \eta Y(\eta) e^{-\int \langle \eta, a^{-1}\eta \rangle / 2} dt.$$  

Where here we’re now using the $\langle , \rangle$ notation to indicate the inner product (and not the autocorrelation function). Now, we can ”uncomplete the square” using a Hubbard-Stratonovich transformation. This technique introduces an auxiliary scalar field $\mu$ and linearizes the $\eta$ in the exponent; but it is most useful to us because it uninverts the diffusion tensor $a$ yielding

$$\langle Y(\eta) \rangle = \int \int \mathcal{D} \mu \mathcal{D} \eta Y(\eta) e^{-\int \langle \mu, a\mu \rangle / 2 - i\langle \mu, \eta \rangle} dt$$

From here we can use equation (72) to put the integral in terms of $u$, rather than $\eta$:

$$\langle Y(u) \rangle = \int \int \mathcal{D} \mu \mathcal{D} u Y(u) J[u] e^{-\int \langle \mu, a\mu \rangle / 2 - i\langle \mu, \dot{u} - b \rangle} dt$$
With $J[u] = \det \left( \partial_t - \frac{db}{du} \right)$. Here, $J[u]$ is the Jacobian that results from transforming $\eta$ to $u$.

So, we have the equation for the expectation value of $Y(u)$. Remember, our goal is to find the probability that a function $F(u) = A$ at time $t=0$. This condition corresponds to:

$$Y(u) = \delta(F(u(x, t = 0)) - A) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\lambda e^{-i\lambda(F(u)\delta(t)-A)} \quad (76)$$

The delta function is expanded in terms of $\lambda$ which will act as a Lagrange multiplier in the instanton equations of motion that we ultimately arrive at. Plugging this into the expectation value equation gives:

$$\langle Y(u) \rangle = \int D\mu D\lambda J[u] \frac{1}{2\pi} \int_{-\infty}^{\infty} d\lambda e^{-\frac{1}{2}[\langle \mu, a\mu \rangle]} e^{-i\lambda(F(u)\delta(t)-A)} dt \quad (77)$$

This results in an effective action that looks like:

$$S_{\text{eff}} = -\int \langle \mu, a\mu \rangle / 2 - i\langle \mu, \dot{u} - b \rangle - i\lambda(F(u)\delta(t)-A)dt \quad (78)$$

If we take the functional derivatives on this, we find the instanton equations of motion that correspond to the saddle point trajectory. This trajectory is a critical point of the action, and therefore represents a large contribution to the probability in the path integral. Extremizing the action yields:

$$0 = \frac{\delta S}{\delta \mu} = -i(\dot{u} - b[u]) + a\mu, \quad 0 = \frac{\delta S}{\delta u} = i\dot{\mu} + i(\nabla_u b)^T \mu - i\lambda \nabla_u F(u)\delta(t) \quad (78)$$

This becomes equivalent to equation (74) if we make the substitution: $\mu = -i\theta$.

The Freidlin-Wentzell formulation of the instanton begins by assuming that by assuming the validity of expression (66). Remember, we’re interested in finding the expression for the probability that a function $F(u)$ will be found at to have a particular value at time $t = 0$. The Freidlin-Wentzell formulation of the instanton begins by assuming that by assuming the validity of expression (66).
such that $F(u(t = 0)) = a$. For this purpose, we can define a function $S^*(\lambda)$ such that:

$$e^{\frac{1}{\epsilon}S^*} = \langle e^{\frac{1}{\epsilon}F(u(t=0))} \rangle$$

Now the statement of the expectation value assumes large deviations and therefore:

$$e^{\frac{1}{\epsilon}S^*} = \int_{-\infty}^{\infty} e^{\frac{1}{\epsilon}a} e^{-\frac{1}{\epsilon}S^*_T} da$$

Just as we had done in the MSR version of the proof, we assume that the value of this integral is about equal to the value of the greatest contribution in the sum. This corresponds to the value that minimizes the action. This gives the near equality:

$$S^*(\lambda) \approx \lambda a_{min} - S(a_{min})$$

with $a_{min}$ being the value of $a$ that minimizes the action $S(a)$. This expression implies that $S^*(\lambda)$ is approximately equal to the Legendre transformation of $S(u)$. This means that the classical equations of motion associated with varying the action correspond to the instanton equations of motion. If we carry this out, we reproduce the equations in (74).

### 3.3 Example: Diffusive Ornstein Uhlenbeck

To see an example of how to apply the instanton method, we can look at Ornstein Uhlenbeck problem with a diffusive term added.

$$u_t = \kappa u_{xx} - \alpha u + \eta(x,t)$$

(79)

For simplicity, assume that the noise is Gaussian white noise (uncorrelated). Suppose
we want to know what is the path that the system will take if it begins at \( u(x, -T) = 0 \) at and ends at \( u(0, 0) = A \).

Our instanton equations (Eq. (74) can be written:

\[
\begin{align*}
    u_t(x, t) &= \kappa u_{xx} - \alpha u + \theta, \\
    \theta_t(x, t) &= -\kappa \theta_{xx} + \alpha \theta
\end{align*}
\]  

(80)

with initial conditions:

\[
\begin{align*}
    u(0, 0) &= A, \\
    \theta(x, 0) &= \lambda \delta(x)
\end{align*}
\]

We can solve this problem by taking the Fourier transform on the \( \theta_t \) equation in (80)

\[
\tilde{\theta}(w, t) = (\alpha + \kappa \omega^2) \hat{\theta} \quad \rightarrow \quad \tilde{\theta}(w, t) = \frac{\lambda}{2\pi} e^{(\alpha + \kappa \omega^2) t}
\]

We can apply this result into the Fourier transform of the equation for \( u_t \) in (80).

\[
\tilde{u}_t + (\alpha + \kappa \omega^2) \tilde{u} = \tilde{\theta} \\
\frac{d}{dt}(\tilde{u} e^{(\alpha + \kappa \omega^2) t}) = \frac{\lambda}{2\pi} e^{(\alpha + \kappa \omega^2) t}
\]

\[
\tilde{u}(0, \omega) - \tilde{u}(t, \omega) e^{(\alpha + \kappa \omega^2) t} = \int_t^0 \frac{\lambda}{2\pi} e^{(\alpha + \kappa \omega^2) t'} dt'
\]

Solving the system, we get:

\[
\tilde{u}(t, \omega) = \frac{\lambda}{\sqrt{2\pi} 2(\alpha + \kappa \omega^2)} e^{(\alpha + \kappa \omega^2) t}
\]  

(81)

The last step is to determine how the Lagrange multiplier \( \lambda \) relates to the ending condition
on \( u(0,0) = A \). We take the inverse transform on \( \tilde{u} \) and see that:

\[
u(0,0) \equiv A = \frac{\lambda}{4\sqrt{\alpha\kappa}} \quad \rightarrow \quad \lambda = 4\sqrt{\alpha\kappa}A
\]

With this the trajectory is determined.

### 4 Numerical Methods

On the numerical side, a number of techniques have been developed to find the transition pathways we are trying to find. Monte Carlo methods exist ([52] [53]) however these can be numerically difficult since the trajectories we are looking for almost never occur relative to other possible trajectories. This has given rise to a field of study called rare events sampling which aims to focus the simulation on the rare elements of the state space that we want to look at ([54], [55], [56], [57]). Other methods exist such as Instanton Filtering [58].

The particular category of solver that gMAM belongs to is called relaxation method. "Relaxation method" is a general term that refers to iterative numerical algorithms that tend to converge on the solution through repetition.

Before talking about relaxation methods, let us briefly examine the Filtering method.

#### 4.1 Filtering

The Filtering method is a numerical method for finding the most likely transition path. This technique approaches the problem in an obvious way. If we want to see the most likely trajectory for a particle to transition from point A to point B, we can just directly simulate the motion of the particle according to the SDE and see what path it takes. If it takes a path that ends sufficiently near point B, we record this path. If it fails to end nearby point
Figure 10: A color plot of the results of applying the filtering method to the Ornstein-Uhlenbeck damped pendulum. Dark blue represents low probability density, and red represents very high probability density. The particle begins at the fixed point \((u, v) = (0, 0)\) and ends at the fixed point \((u, v) = (2\pi, 0)\). The trajectory is a spiral path out of the potential well centered at \((0, 0)\) followed by a similar spiral path into the neighboring well centered at \((2\pi, 0)\).

B, we filter out this path. We can repeat this numerical experiment many many times such that eventually we will effectively have a representation of the probability distribution associated with the transition between those endpoints.

The reason we filter out transition paths which end sufficiently near point B as opposed to ending at point B is because, numerically speaking, the probability that the particle will land exactly on point B is effectively 0. Since we wish to solve problems in a finite amount of time, it makes sense to take some small range of points very near point B and say that those solutions which end within that range "count" as trajectories of interest.

Once we have the effective distribution associated with this transition, we can find where the peak probability occurs and therefore extract the most likely trajectory.

For an example of this method, we can look at one of our earlier example SDE's. Recall the drift vector associated with the classical pendulum with dissipation:
\[ \begin{bmatrix} \dot{u} \\ \dot{v} \end{bmatrix} = \begin{bmatrix} v \\ -\sin(u) - \alpha v \end{bmatrix} \]  

(82)

We can apply the filtering method to this system to find the trajectory that the particle will take given the initial condition that it exists in the well on the left and transitions into the well on the right.

The most likely trajectory is that the particle moves in a spiral pattern on its way out of the fixed point, then spirals downward into the endpoint on the right.

One possible drawback associated with this method is that the number of required repeat "experiments" grows very quickly for sufficiently rare transitions. This means that unless sophisticated sampling techniques are employed, the computation time can become intractable. However, if the interest is in broad qualitative results, the filtering method is quite useful. Another example problem that we can look at is the Maier-Stein laser model that we will study in more detail in a later section.

### 4.2 String Method

Before delving into the way gMAM functions, let us examine one of its conceptual predecessors. The string method, like the gMAM, is a relaxation method for finding the minimum action path between two states. The idea behind the string method is simple. It proposes that the space of paths itself can be looked at in a physical way. So, if we consider a system with a starting and ending point, we can imagine the path that connects the two existing as a string in some physical space. We can then try to define a corresponding potential that tends to "push" the string towards the equilibrium position that corresponds to the minimum energy path. In this way, as time goes to infinity, the system "relaxes" into the minimum that we assume is the solution to our problem. This section will review the theory associated with the string method. There are a number of
applications covered in this thesis that one can use to see how to implement the string method. Namely: Ginzburg-Landau problem (5.2.1) or the Maier-Stein system (5.1.1). The interested reader is instructed to visit those sections for examples of practical implementation.

The string method is originally described in [59], and improved upon in [17]. This section will summarize those ideas. First, we imagine a string, where the position along the state of the string is represented by the variable $\gamma$. Then, define a potential $V$ such that when the string is at the lowest point in the potential, the component of $\nabla V$ normal to the string (given by $\gamma$) is zero (i.e. there is no force on the string when it reaches the minimum action path):

$$\left(\nabla V\right)^\perp(\gamma) = 0$$  \hspace{1cm} (83)

Note that

$$\left(\nabla V\right)^\perp(\gamma) = \nabla V(\gamma) - \left(\nabla V(\gamma), \hat{\tau}\right)\hat{\tau}$$  \hspace{1cm} (84)

where $\hat{\tau}$ is the unit vector pointing tangent to the string. Now, as the string gets pushed, the length of the string may change. However, we want to make it so that the variable $\gamma$ can be parameterized according to $\gamma: \gamma = \varphi(\alpha): \alpha \in [0, 1]$. In words, this means that if the string has a normalized length of 1, $\alpha$ represents the distance along the string from the initial state (which itself is given by $\varphi(0)$), thus $\varphi(1)$ corresponds to the final state. Having chosen this parameterization, we can constrain the string to have a constant length by assuming: \[ \left|\frac{\partial \varphi}{\partial \alpha}\right| = \text{const}. \] The differential equation that the algorithm seeks to integrate is therefore:

$$\dot{\varphi} = -\nabla V(\varphi) + \lambda \hat{\tau}$$  \hspace{1cm} (85)

Where $\lambda$ is some Lagrange multiplier that will be used for enforcing our parameterization on the string in each time-step of the integration. There are many ODE solvers that can
integrate this equation (e.g. a 4th order Runge Kutta (RK4), or Euler methods).

In each step of the integration ($\phi^n \to \phi^{n+1}$), the chosen parameterization will have to be reinforced as well. There are a few choices on how to do this. In this paper, we’re going to stick with the so called equal arc length parameterization.

The method to enforce the equal arc length parameterization is straightforward. The first step is to calculate the local cumulative arc length at a point $n$ along the string given by:

$$\text{arclength}_n \equiv s_n = \sum_{i=2}^{n} |\varphi_i - \varphi_{i-1}|$$  \hspace{1cm} (86)

$$s_1 = 0$$  \hspace{1cm} (87)

Once $s_n$ is known, enforcing the condition $\alpha_N = 1$, requires that the function $s_n$ be normalized by dividing all the terms by $s_N$.

The second step is interpolation. The new reparametrized values of $\phi$ must be found by pointwise interpolation from the old values of $\phi$ using $s_n$.

We will see in later sections that this parameterization scheme can be generalized to higher dimensional systems (for PDE problems). In that case, the norm taken in step one has to be generalized to the L2 norm.

### 4.3 gMAM

One important aspect of the string method that was briefly mentioned in the last section, is that it only works in the case where the energy landscape can be written as the gradient of a potential. When dealing with a problem that cannot be written in terms of a potential gradient, the logic of the string method is no longer applicable, thus it will fail. For such situations we need more general tools. It was for this reason that the Minimum Action Method (MAM) was created [19].
The idea behind the MAM is that we want to take an SDE of the form of Equation (11), and numerically compute the most likely transition path between two stable fixed points.

\[
dx(t) = b(x(t))dt + \sqrt{\varepsilon} \sigma(x(t))dW(t)
\]

The strategy that MAM uses to achieve this goal is to represent this SDE as a Freidlin-Wentzell action functional and find the minimum action path. If we try to do this, we quickly run into a potential problem: Finding the action requires an integral of the Lagrangian over all time.

\[
S_T(x) = \int_0^T L(x, \dot{x}) dt
\]

There are many cases where MAM works just fine such as Ginzburg-Landau Equation [19], or the Kardar-Parisi-Zhang (KPZ) Equation [60]. However this can represent a serious numerical complication. If one of the points along the minimizing path contains a critical point, the trajectory requires an infinite amount of time. This makes MAM numerically intractable for certain systems. Therefore, for cases such as those, we need to modify our approach. geometric minimum action method (gMAM) does just that by re-parameterizing the system such that the action integral relies solely on the geometry of a potential, rather than explicitly on the time.

Let us examine how this is done. Freidlin-Wentzell Theory of large deviations as described in [16] shows that SDE’s of the type ((11)) can be assigned a quasi-potential.

\[
V(x_1, x_2) = \inf_{T > 0} \inf_{\psi \in C^2_T} S_T(\psi)
\]

When the noise is sufficiently small, the quasi-potential represents the difficulty of transitioning from one point to another. In this sense it is similar to the path integrals describing
the probability of transition in section 2.3 and for instantons in section 3.2. We can take this quasi-potential and substitute our Lagrangian for diffusive SDEs

$$V(x_1, x_2) = \inf_{T > 0} \inf_{\psi \in C_t^2} \int_0^T |\psi(t) - b(\psi(t))|_{a(\psi)}^2 dt$$ \hspace{1cm} (89)$$

Now, the sum of the squares of two positive real numbers is always going to be greater equal to twice the product of the two, so:

$$|u|_{a(x)}^2 + |v|_{a(x)}^2 \geq 2|u|_{a(x)}|v|_{a(x)}$$

We can use this fact to make an inequality on $V(x_1, x_2)$:

$$V(x_1, x_2) \geq \inf_{T \psi} \int_0^T |\psi(t)||b(\psi(t))|_{a(\psi)} - (\psi(t), b(\psi(t)))_{a(\psi)}) dt$$

$$= 2\inf_{T \psi} \int_0^T |\psi(t)||b(\psi(t))|_{a(\psi)} \sin^2 \frac{1}{2} \eta(t) dt$$

where in the last step we make an analogy to the normal dot product and use the identity $\sin^2(x) = \frac{1}{2}(1 - \cos(2x))$. $\eta$ is the angle between $\dot{\psi}(t)$ and $b(\psi(t))$ in the space of the scalar product $(\cdot, \cdot)_{a(\psi)}$. At this point, we can put a constraint on $\psi$ such that $|\dot{\psi}|_{a} = |b(\psi(t))|_{a}$ and get:

$$V(x_1, x_2) = 2\inf_{\gamma} \int_{\gamma} |b(\psi(t))|_{a(\psi)} \sin^2 \frac{1}{2} \eta(t) ds$$ \hspace{1cm} (90)$$

This expression of the quasi-potential removes the need to explicitly deal with $\psi$ and therefore the need to integrate over time. Similarly to the string method, here $\gamma$ represents the curve marking the trajectory of the transition. With this parameterization, $\eta$ represents the angle between $b$ and $\gamma$ at position $s$ along the curve. This is the key idea.
that separates gMAM from other methods. Now, just like the string method, we recognize that we can choose our parameterization of the curve $\gamma$ such that $\gamma = \varphi(\alpha) : \alpha \in [0, 1]$. With this parameterization set, we can now reformulate the quasi-potential:

$$V(x_1, x_2) = \inf_{\varphi \in C^2([0, 1])} S(\varphi)$$

(91)

Thus our new action, which is no longer in terms of time, is:

$$S(\varphi) \equiv \sup_{\vartheta : [0, 1] \to \mathbb{R}^n} \int_0^1 \langle \varphi', \vartheta \rangle d\alpha$$

(92)

Where $\vartheta$ is defined as the particular value of $\theta$ that extremizes the Hamiltonian ($H(\varphi, \vartheta) = 0$) for a particular value of $\varphi$. Equivalently, we can choose to write the action in the following ways:

$$S(\varphi) = \int_0^1 \langle \varphi', \hat{\vartheta}(\varphi, \varphi') \rangle d\alpha$$

(93)

or

$$S(\varphi) = \int_0^1 \frac{L(\varphi, \lambda \varphi')}{\lambda} d\alpha$$

(94)

Here $\lambda$ is a non-negative function of $\varphi$ and $\varphi'$ that carries the constraint that corresponds to the extremizing of the Hamiltonian with $\vartheta$. It should be noted that the action in Equation (94) must be zero if $\lambda = 0$, and is also zero if $\varphi' = 0$. $\lambda(\varphi, \varphi')$ is defined according to:

$$H(x, \vartheta) = 0, H(x, \vartheta)_{\theta} = \lambda y$$

(95)

It can be seen from this equation how this $\lambda$ gets passed to the Lagrangian through a Legendre transformation. We will see how that looks in a moment. For now our action, as we have defined it, is invariant under a change in the parameterization of $\varphi$. Therefore,
we have an additional constraint available to us. Just like the string method, one attractive parameterization is to simply demand that $|\varphi'| = \text{const.}$.

Consider that for SDE’s that can take the form of equation (11). In that case the Lagrangian is given by

$$L(x, y) = \frac{1}{2} \langle y - b(x), a^{-1}(y - b(x)) \rangle \equiv \frac{1}{2} |y - b(x)|_a^2.$$ 

In that case we can easily relate $\vartheta$ to $\lambda$ and $y$.

$$\vartheta(x, y) = a^{-1}(x) \left( \frac{|b(x)|_a}{|y|_a} y - b(x) \right) \quad (96)$$

Given the definition of $\lambda$ in Equation (95) we can conclude that $\lambda$ must be given by:

$$\lambda = \frac{|b(x)|_a}{|y|_a} \quad (97)$$

In this case we can also simplify our action $S(\varphi)$ like so:

$$S(\varphi) = \int_0^1 |\varphi'|_a |b(\varphi)|_a - \langle \varphi', b(\varphi) \rangle_a d\alpha \quad (98)$$

Thus far we have covered the basic theory behind the gMAM. However our interest lies in applying these ideas towards constructing a numerical algorithm.

One intuitive way to see how they arise is presented in [61]. The idea is that if we take $\psi(t)$ to be the function that describes the path of the system as a function of time and $\varphi(\alpha)$ describes the path of the system as a function of alpha then we can imagine a function $G(\alpha)$ such that $\psi(G(\alpha)) = \varphi(\alpha)$. Further we can write the Hamilton’s equations associated with $\psi$. $-H_x = \dot{\theta}, H_\theta = \psi$ If we suppose this to be the case, then we can write the equations:

53
\[ \lambda \varphi' = \psi(G(\alpha)) \quad \text{and} \quad \lambda^2 \varphi'' + \lambda \lambda' \varphi' = \psi(G(\alpha)) \]

with \( \lambda = 1/G' \). Then we can use the Hamiltonian to rewrite it like so:

\[ \ddot{\psi} = H_{\theta x} \dot{\psi} + H_{\theta \theta} \dot{\theta} \]
\[ = H_{\theta x} \dot{\psi} - H_{\theta \theta} H_x \]
\[ \lambda^2 \varphi'' + \lambda \lambda' \varphi' = H_{\theta x} \lambda \varphi' - H_{\theta \theta} H_x \]
\[ \lambda^2 \varphi'' - \lambda H_{\theta x} \varphi' + H_{\theta \theta} H_x + \lambda \lambda' \varphi' = 0 \]

Therefore, the most likely trajectory is the one that makes the left hand side of the above equation equal to zero. Thus the equations of motion are given by [20]:

\[
\begin{cases}
\dot{\varphi} = \lambda^2 \varphi'' - \lambda H_{\theta x} \varphi' + H_{\theta \theta} H_x + \lambda \lambda' \varphi', \\
\varphi(\tau, 0) = x_1, \quad \varphi(\tau, 1) = x_2 \quad \varphi(0, \alpha) = \varphi^0(\alpha),
\end{cases}
\]

Note that now we have a time derivative appearing in the equation. Thus far \( \varphi \) has been a function of \( \alpha \) only and we have done a lot of work to make sure \( \varphi \) was independent of time. This new “time” \( \tau \) is an artificial construction that corresponds to the relaxation time of the system. gMAM, like the string method, imagines that the trajectory \( \gamma \) can be thought of as a string being pushed by a potential into a relaxed position that corresponds to the most likely trajectory. The differential equation above describes the “motion” of the trajectory as a function of relaxation time. \( \mu \varphi' \) is a Lagrange multiplier term that enforces some constraint on the parameterization.

Clearly when \( \varphi = 0 \), that indicates that the system has reached a local extremal point and therefore that (hopefully) the algorithm has found the minimum of the action that we are looking for. If so, this corresponds to the most likely trajectory. One other useful way to
check if the trajectory is correct is by looking at the $\lambda$. As noted above, the $\lambda$ has a special relationship with the action. Specifically, when $\lambda = 0$ the action also must be zero. This means that we can check whether our solutions actually minimize the action by checking to make sure that our final trajectory contains one of these critical points where $\lambda = 0$.

So, let us see how the corresponding numerical algorithm looks. Consider a discretized space where there are a total of $N+1$ points connecting the $\phi(0)$ to the endpoint $\phi(1)$. This implies that the space between these points is given by: $\Delta \alpha = 1/N$. Then we can write the equation:

$$
\begin{align*}
\frac{\dot{\phi} - \Phi}{\Delta \tau} &= (\lambda_i)^2 \frac{\Phi_{i+1} - 2\Phi_i + \Phi_{i-1}}{1/N^2} - \lambda_i^k \Theta_x \phi'_i + \Theta_{\theta \theta} H_x + \lambda_i \lambda'_i \phi'_i, \\
\bar{\phi}_0 &= x_1, \quad \bar{\phi}_N = x_2
\end{align*}
$$

(100)

We have taken a linear approximation of the time derivative such that $\phi$ represents the trajectory evaluated at current time $\tau$, while $\bar{\phi}$ refers to the trajectory evaluated one time step into the future ($\tau + \Delta \tau$). $\lambda$ and $\phi'$ are evaluated at current time $\tau$. The subscript $i$ is the discrete variable that refers to the relative position along the $\alpha$ direction.

Finding the values of the terms in (100) is straightforward.

$$
\phi'_i = (\phi_{i+1} - \phi_{i-1}) \frac{N}{2}, \quad \dot{\phi} = \dot{\phi}(\phi_i, \phi'_i) \rightarrow (Equation : (96))
$$

$$
\lambda_i = \langle H_\theta(\phi_i, \dot{\phi}_i), \phi'_i \rangle / |\phi'_i|^2, \quad \lambda'_i = (\lambda_{i+1} - \lambda_{i-1}) \frac{N}{2}
$$

$$
\lambda_0 = 3\lambda_1 - 3\lambda_2 + \lambda_3, \quad \lambda_N = 3\lambda_{N-1} - 3\lambda_{N-2} + \lambda_{N-3}
$$

Where $\lambda_0$ and $\lambda_N$ are found through interpolation. Obviously the derivatives on the Hamiltonian: $H_\theta, H_x, H_{\theta \theta}, H_{\theta x}$ should be evaluated at the "present" time $\tau$. I.e. $H_x = H_x(\phi, \dot{\phi})$.

The particulars of these terms are going to depend on the specifics of the problem, however if we assume that the form of the SDE we're trying to solve is an ODE representing
George Poppe

Applications of gMAM

a diffusion process where the diffusion tensor is the identity, then we can derive a more explicit form of the algorithm:

We can write out the derivatives of the Hamiltonian in terms of \( b, \phi, \) and \( \vartheta \):

\[
H_{\theta x} = \nabla b, \quad H_{\theta \theta} = I, \quad H_x = (\nabla b)^T \vartheta.
\]  

(101)

Remember that eq. (96) tells us the relationship \( \theta \) has to \( \phi \). In this case, it reduces to:

\[
\vartheta = \lambda \phi' - b(\phi)
\]

Thus we can rewrite \( H_x = (\nabla b)^T \vartheta = (\nabla b)^T (\lambda \phi' - b) \). With this, we can write the entire relaxation equation in terms of \( \phi, \phi', \) and \( \lambda \):

\[
\phi = \lambda^2 \phi'' - \lambda (\nabla b - (\nabla b)^T) \phi' - (\nabla b)^T b + \lambda \lambda' \phi'.
\]  

(102)

After integrating one step of (100), now we have to have to ensure that our constraints on the parameterization still hold. To do this we reparametrize by interpolating a curve across \( \{ \tilde{\phi}_i \}_{i=0,...,N} \) and discretize the curve to find our new \( \phi(\tau + \Delta \tau) \).

Now we are back at the situation we started with, and the algorithm can be repeated until \( \phi \) is within convergence criteria.

4.4 gMAM for PDEs

In the last section, the gMAM was presented which could solve an ODE of the form:

\[
du = b(u)dt + \sigma(u) dW.
\]  

(103)

It is possible to extend the functionality of gMAM so that it can be applied to PDE prob-
lems as well as ODEs. This is done through what is known as an infinite dimensional generalization of the gMAM. Let us take the case that \( u \) is a function of \( x \), \( u = u(x,t) \), and that the partial derivative that appears in the SDE is taken on the \( x \) variable, \( \kappa \frac{\partial^2}{\partial x^2} u \), so that we can write:

\[
du = B(u) dt + \sigma dW \quad \text{with} \quad B(u) = \kappa u_{xx} + b(u)
\]  

(104)

In the ODE gMAM, the concept was that the trajectory connecting two stable fixed points will relax into the minimizing trajectory when the appropriate evolution equation is integrated over artificial time. In the infinite dimensional generalization, instead of merely having two endpoints which correspond to stable fixed points, we have a continuum of points spanning the \( x \) dimension. This essentially amounts to the fact that for PDEs, the gMAM evolution equation describes the evolution of a surface which connects two boundary lines that correspond to fixed points in \( u \). Another way to look at it is that it is necessary to simulate the \( x \) dimension in order to calculate the value of the partial derivative \( u_{xx} \).

The gMAM introduced in the previous section is only valid for the case where \( b(x(t)) \) does not contain partial derivative terms. This changes one major aspect of the derivation of the gMAM. :

\[
S(\phi) = \int_0^1 \| \phi' \|_{L^2} \| B(\phi) \|_{L^2} - \langle \phi', B(\phi) \rangle_{L^2} \, d\alpha
\]  

(105)

When this action is varied with \( B \) in the integrand, it results in the appearance of differential operators so that the equation of motion corresponding to the relaxation of the system can be written as:
\[ \phi = \lambda^2 \phi'' - \lambda (\partial B - (\partial B)^*) \phi' - (\partial B)^* B + \lambda \lambda' \phi'. \] (106)

Where

\[ \partial B = \kappa \partial_{xx} + \nabla b \] (107)

Also in this situation, the equation corresponding to \( \lambda \) is slightly different. Instead of being \( \frac{|b(\phi)|}{|\phi|} \), \( \lambda \) is now given by the L2 norm of B divided by the L2 norm of \( \phi' \):

\[ \lambda = \frac{||B||_{L^2}}{||\phi'||_{L^2}} \]

which results in the explicit evolution equation:

\[ \phi = \lambda^2 \phi'' - \lambda (\nabla b - (\nabla b)^T) \phi' - (\nabla b)^T b - \kappa (\nabla b + (\nabla b)^T) \phi_{xx} - \kappa^2 \phi_{xxxx} - \kappa \left( \frac{\langle \phi_x, \nabla b_1 \phi_x \rangle}{\langle \phi_x, \nabla b_2 \phi_x \rangle} \right) + \lambda \lambda' \phi'. \] (108)

In a case like this, there is an issue of stability that arises with the \( \phi_{xxxx} \) term. To stabilize the integration process, it is good to use a split-step where in the first part of the step, we have

\[ \phi = \lambda^2 \phi'' - \lambda (\nabla b - (\nabla b)^T) \phi' - (\nabla b)^T b - \kappa (\nabla b + (\nabla b)^T) \phi_{xx} - \kappa \left( \frac{\langle \phi_x, \nabla b_1 \phi_x \rangle}{\langle \phi_x, \nabla b_2 \phi_x \rangle} \right) + \lambda \lambda' \phi'. \]
and then in the second step we have only the 4th derivative term:

\[ \phi = -\kappa^2 \phi_{xxxx} \]

Note that this the second step can be solved implicitly as well (or by using fast Fourier transform and the associated propagator in frequency space).

### 4.5 Linear Schrödinger Equation

One interesting PDE application of the gMAM that can also serve as a tutorial for its implementation can be found in the linear Schrödinger equation. It is especially useful because it allows us to check the analytic theory against the numerical implementation (gMAM).

Consider the following equation:

\[ A_t(x, t) = idA_{xx} - a + A + \beta A_{xx} + \eta \quad (109) \]

Where d, a, \( \beta \), and \( \sigma \) are constants; and \( \eta \) is mean zero, Gaussian white noise. In this section, there will be two major topics:

- 1. Finding the instanton solution analytically.
- 2. Applying the gMAM to find the instanton of interest.

#### 4.5.1 Analytic Solution

The theory and background associated with finding the analytical solution to the instanton problem is given in Section 3.2. We are interested in finding the evolution of the diffusive linear Schrödinger equation supposing the condition that \( A(x, -T) = 0 \), and \( A(0, 0) = k \) where \( k \) is a constant.
The instanton path is governed by the following equations of motion which correspond to equation (74):

\[ A_t = idA_{xx} - aA + \beta A_{xx} + B, \quad B_t = idB_{xx} + aB - \beta B_{xx} \]  

(110)

Where \( B(x,t) \) is the momentum conjugate variable with respect to \( A(x,t) \) and \( B(x,0) = F \cdot \delta(x) \). \( "F" \) is a Lagrange multiplier whose value must be found via calculation. We construct the Fourier transform of \( B(x,t) \), \( \tilde{B}(\omega,t) \), which gives us the equality:

\[ \tilde{B}_t(\omega,t) = (a + (\beta - id)\omega^2)\tilde{B} \rightarrow \tilde{B} = Fe^{(a+(\beta-id)\omega^2)t} \]

And this can be solved by following the same technique as presented in section 3.3. The result in this case is:

\[ \tilde{A}(\omega,t) = \frac{F}{2(a+\beta\omega^2)}e^{(a+\beta\omega^2)t}e^{-id\omega^2t} \quad \text{with} \quad F = 4\sqrt{a\beta k} \]  

(111)

### 4.5.2 Applying gMAM to the linear Schrödinger SPDE

In this section we examine a novel application of the gMAM to the linear Schrödinger equation with dispersion. In order to apply gMAM to the linear Schrödinger equation (equation (109), we need to begin by recognizing that \( A(x,t) \) is a complex function. We can write the real and imaginary parts separately creating a pair of coupled PDE's. So we make the substitution \( A(x,t) = u(x,t) + iv(x,t) \).

\[ u_t + iv_t = d(\nu_{xx} - \nu) - a(u + iv) + \beta (u_{xx} + iv_{xx}) + \eta \]  

(112)
If we split the real and imaginary parts we get:

\[
\begin{align*}
\text{Real} : u_t &= -au + \beta u_{xx} - dv_{xx} + \eta_R \\
\text{Imag} : v_t &= -av + \beta v_{xx} + du_{xx} + \eta_I
\end{align*}
\]  

(113)  

(114)

This amounts to replacing \( i \), the imaginary number, with a rotation matrix \( \mathcal{J} \) with the form:

\[
\mathcal{J} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}
\]

In doing this expansion, there's an assumption made about the nature of the noise; that the real component of the noise, \( \eta_R \), is independent of the imaginary component, \( \eta_I \), and has the same properties. Specifically, both \( \eta \)'s represent independent mean zero Gaussian white noise processes.

With the understanding that the problem represents two coupled SPDE's, the gMAM can be constructed as outlined in section 4.4. To start, we need to clarify the initial and final condition of the problem. Remember that applying the gMAM to a PDE system with functions \([u(\alpha, x), v(\alpha, x)]\) requires that we specify the four lines given by \([u(0, x), v(0, x)]\) and \([u(1, x), v(1, x)]\). Suppose for simplicity that we choose the condition that \( u(0, x) = 0, v(0, x) = 0 \) and \( v(1, x) = 0 \). Then the condition on \( u(1, x) \) that corresponds to a fixed point is given by:

\[-au + \beta u_{xx} = 0\]

Therefore there is a fixed point at \( u(1, x) = ae^{-\frac{|x|}{\sqrt{\beta}}} \). Once the initial and final conditions are chosen, we now have to simulate the gMAM PDE equations of motion:
Figure 11: Displays the gMAM solution to the linear Schrödinger equation that describes the transition from $[u(T_{\text{min}},x), v(T_{\text{min}},x)] = [0,0]$ to $[u(0,x), v(0,x)] = [a \exp(-|x|/\sqrt{\beta}),0]$. The V component is roughly an order of magnitude smaller than the U component. This corresponds to having a weak dispersion term.

\[
\begin{bmatrix}
\dot{u} \\
\dot{v}
\end{bmatrix} \equiv \phi = \lambda^2 \phi'' - \lambda (\partial B - (\partial B)^*) \phi' - (\partial B)^* B + \lambda \lambda' \phi'.
\]

In order to evaluate this expression, we have write the problem in the proper form (see eq. (104)) with $B(\phi) = \kappa \phi_{xx} + b(\phi)$. In other words, we have to find $b(\phi)$ and $\kappa$. Examining equation (113) reveals that $B(\phi)$ is a two dimensional vector and that $b(\phi)$ and $\kappa$ must have the following forms:

\[
b(\phi) = -\begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix} \phi \quad \text{and} \quad \kappa = \begin{bmatrix} \beta & -d \\ d & \beta \end{bmatrix}
\]

The calculations below will be written for a general $b(\phi)$ so that they can be easily applied...
Figure 12: Comparison between a slice of the the analytic instanton calculation for the linear Schrödinger equation (red) against the gMAM result (blue) for arc length parameter $\alpha = .78$. The left panel shows the real part $u$, the right panel the imaginary part $v$.

to other problems. We can write $\partial B$ and $(\partial B)^T$:

$$\partial B = \kappa \partial_{xx} + \nabla b \quad \text{and} \quad (\partial B)^T = (\kappa \partial_{xx})^T + (\nabla b)^T$$

Explicitly working out the terms in (126) we have:

$$\partial B - (\partial B)^T = (\nabla b - (\nabla b)^T) + 2 d \mathcal{F} \partial_{xx}$$

$$(\partial B)^T (B) = (\nabla b)^T B + \kappa b_{xx} + (\beta^2 + d^2) \varphi_{xxxx}$$

with

$$\nabla b - (\nabla b)^T = \begin{pmatrix} b_{1v} - b_{2u} \\ b_{2u} - b_{1v} \end{pmatrix} \quad \text{and} \quad (\nabla b)^T B = \begin{pmatrix} b_{1u}B_1 & b_{2u}B_2 \\ b_{1v}B_1 & b_{2v}B_2 \end{pmatrix}$$

So if we write everything out as two coupled equations for $u$ and $v$, we get:
\[ \dot{u} = \lambda^2 u'' - \lambda (b_1 v - b_2 u - 2dv'_{xx}) - (\beta^2 + d^2)u_{xxxx} - [b_1 u B_1 + b_2 u B_2 + \beta b_{1xx} + db_{2xx}] \]

\[ \dot{v} = \lambda^2 v'' - \lambda (b_2 u - b_1 v + 2du'_{xx}) - (\beta^2 + d^2)v_{xxxx} - [b_1 v B_1 + b_2 v B_2 + \beta b_{2xx} - db_{1xx}] \]

This is the prescription for the relaxation trajectory according to gMAM. Let us look in detail at how one can carry out this process.

%Computational Parameters
N = 1024; % discretization in \alpha
nP = 1024; % discretization in x
d = 2*nP; % dimension of the system
dt = 0.00005; % time step for relaxation
MaxIter = 4000; % max number of iterations

% Physical Parameters :
beta = 1.0; %beta
d = 1.0; %d
a = 1.0; % a
% grid in x
xMin = -25; xMax = 25; xD = xMax-xMin;

x = linspace(xMin,xMax-xD/nP,nP); dx = x(2)-x(1);
alpha_eq = linspace(0,1,N);

The code snippet above indicates the chosen parameters for the simulation and gives an idea of the computational load associated with simulating this problem. Figure 11 shows the results of one such simulation. The function \( u(\alpha,x) \) has a straight forward behavior, taking a near-exponential path between the two initial conditions. The behavior in \( v(\alpha,x) \) is more interesting, although the magnitude of the effect is small. We can check the result by comparing the instanton solution to the gMAM solution for a particular set of parameters. Figure 12 shows one such comparison. It depicts a slice of the two solutions at position corresponding to \( \alpha = .78 \). The two solutions show very close agreement, the
small difference that exists goes away if we increase the resolution of the simulation. The agreement between these two results is very significant and further suggests that we can be confident in the effectiveness of gMAM as a predictive tool.

5 Applications

5.1 Maier-Stein

The Maier-Stein system is a two-dimensional SDE that can be, depending on the choice of parameter, either gradient or non-gradient. The Maier-Stein SDE is a relatively well known system in the field of escape phenomena. This section explores the solution of Maier-Stein for both SDE and PDE cases using the string method and the gMAM. Also, we note that some of these cases have been studied before such that we can compare our calculation to the previously obtained results.

5.1.1 Maier-Stein: String Method

Recall the general form of the Maier-Stein drift vector:

\[
\begin{bmatrix}
    b_1(u,v) \\
    b_2(u,v)
\end{bmatrix}
= \begin{bmatrix}
    u - u^3 - \beta uv^2 \\
    - (1 + u^2)v
\end{bmatrix}.
\] (115)

As noted above, this drift vector can be written as the gradient of a potential \( V \) for certain choice of parameters. If \( \beta = 1 \) we can write \( V \):

\[
V = \frac{u^2}{2} \left( \frac{u^2}{2} - 1 \right) - \frac{v^2}{2} (u^2 + 1).
\] (116)

This allows us to apply the string method. Remember, the purpose of the string method is to find the trajectory that a particle is most likely to take given an initial state and a final
George Poppe

Applications of gMAM

state. The kinds of initial and final states that we are interested in are stable fixed points. So, before we can apply the string method to solve this problem, we need to locate the stable fixed points. As discussed in section 2.1, the stable fixed points occur at positions $(0,0)$, $(0,0.2)$, $(0,0.4)$, $(0,0.6)$, $(0,0.8)$, and $(0,1)$.

Figure 13: Initial conditions for Maier-Stein string method. The initial condition is a curve that connects the fixed points $(u,v) = (-1,0)$ and $(u,v) = (1,0)$.

But let us find them analytically rather than merely by inspection, as we did in 2.1. To find the fixed points, we need to set the drift vector equal to zero:

\[ 0 = u - u^3 - \beta uv^2, \quad 0 = -(1 + u^2)v. \]  

(117)

This forces the time derivative of $u$ and $v$ to be equal to the noise at that point. This condition has 2 real solutions:

\[ x_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad x_2 = \begin{bmatrix} -1 \\ 0 \end{bmatrix} \]  

(118)

We also need to determine if these points are stable. To do that we use the usual stability analysis techniques of taking the derivatives with respect to $u$ and $v$ and ensure that the
forcing tends to push the state back towards the fixed point. We find that indeed these fixed points are stable.

Now that we know what our fixed points are, we want to create a discretized space where those two endpoints are included in the set of positions. Then with the space defined, we can now assume an initial condition for the trajectory of our system. This trajectory is the "string" that gets pushed by the potential until the system reaches a minimum. For our purposes let us suppose that we choose our initial trajectory to be as follows:

\[ u = -1 + 2\alpha + \sin^2(\pi\alpha), \quad v = \sin^2(\pi\alpha). \]

Figure 13 shows what the initial conditions look like by plotting both \( u \) and \( v \) as a function of \( \alpha \). It also displays the final result which is found by integrating the string evolution equation given by equation (85):

\[ \dot{\phi} = -\nabla V(\phi) + \lambda \dot{t} \]
Where $-\nabla V$ is equivalent to the drift vector $b$ from equation eq. (115). If we examine the potential function given by the drift vector from (115). In this particular case, a simple Euler-scheme given by

\[ u_{\text{new}} = u + b_1(u,v) d\tau, \quad v_{\text{new}} = v + b_2(u,v) d\tau \]

will lead to sufficiently fast convergence. After each time step $d\tau$ the string is reparametrized. The string method is very easy to implement relative to the gMAM, which is covered in the following section.

### 5.1.2 Maier-Stein(SDE): gMAM

Consider the following coupled system of SDEs:

\[
du = (u - u^3 - \beta uv^2)dt + \sqrt{\varepsilon}dW_u(t), \quad dv = -(1 + u^2)vdt + \sqrt{\varepsilon}dW_v(t)
\]  

(119)

This is known as the Maier-Stein system. Here, as in other cases, the $dW$ terms represent Wiener increments with the subscripts $u$ and $v$ to indicate that they are independent from one another. The parameter $\beta$ (as above) is a constant that we can set. The string method could only guarantee us the minimizing trajectory for the case that $\beta = 1$ because it was only in that case that we could write the drift vector as the derivative of a potential $V$. When applying gMAM, however, we are interested in cases where $\beta > 0$ because it is in those cases that we will certainly have two stable fixed points to transition between. Putting this in the form of Equation (11) implies that our drift vector $b$ must be:

\[
b(u,v) = \begin{pmatrix}
    u - u^3 - \beta uv^2 \\
    -(1 + u^2)v
\end{pmatrix}
\]  

(120)
When $\beta$ is set to 1, the drift vector $b$ reduces to the gradient of some potential. This fact, among others, makes Maier-Stein a good test case for gMAM.

We can use the same fixed points that we found previously in Section 5.1.1 given by Equation (118). Namely:

$$x_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad x_1 = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

These fixed points will characterize our boundary conditions. Let us look at how to actually implement the algorithm that we defined in the Section 4.3. Our governing equation describes the motion of the "string" (corresponding to the trajectory) that connects our initial point $x_1$ and our final point $x_2$. It is important to note that this transformation requires that we parameterize $\varphi$ as a function of $\alpha$ and that, during the simulation, we maintain that parameterization. For our case, we choose $|\varphi'| = \text{const}$. In other words that the positions "$\alpha$" are akin to markings on a string so that the length of string between them never changes). To start, we need to pick an initial state for the string in our space. The simplest choice is to assume that the particle moves in a straight line between the two endpoints. This can be justified as a starting assumption given what we learned in section 2.1, where we examined the qualitative character of the Maier-Stein drift vector when $\beta$ was set equal to 1 (See figure (4)). Equation (118) tells us that a straight line initial condition gives us a line that lies along the $u$-axis starting from $u = -1$ and ending at $u = 1$.

Once we have chosen an initial condition, we can proceed to applying our equation of motion corresponding to the relaxation of the trajectory, see eq.(100):

$$\frac{\Phi - \Phi}{\Delta \tau} = (\lambda_i)^2 \frac{\Phi_{i+1} - 2\Phi_i + \Phi_{i-1}}{1/N^x} - \lambda_i^k H_{\theta_x} \varphi'_i + H_{\theta_\theta} H_x + \lambda_i \lambda_i^k \varphi'_i$$
Here, $H$ is the Hamiltonian according to the Freidlin-Wentzell theory (eq. (68)) and $\lambda$ is given by $\lambda = \frac{|h(\phi)|}{|\phi'|}$. This problem is indeed an ODE with a diffusion tensor equal to the identity, so we can use the result derived in section 4.3. Therefore, we can write the entire relaxation equation in terms of $\phi, \phi'$, and $\lambda$ according to equation (102):

$$\dot{\phi} = \lambda^2 \phi'' - \lambda\left(\nabla b - (\nabla b)^T\right)\phi' - (\nabla b)^T b + \lambda \lambda' \phi'.$$

In the case of Maier-Stein system, we can write $\nabla b$ as a 2x2 matrix:

$$\begin{pmatrix}
1 & -3 u^2 - \beta v^2 \\
-2uv & -(1+u^2)
\end{pmatrix}$$

(121)

So that when we compute $\nabla b(\phi)\phi'$ we are left again with a two-dimensional vector which is the dimension of $\phi$. Now that all the terms in equation (102) are accounted for, it needs to be integrated in $\tau$ so that the system can relax from the chosen initial condition into the minimizing trajectory.

Finally, after each step of the integration, it is important to make sure that the parameterization $|\phi'| = const.$ still holds. If it does not, the string must be reparametrized to obey
this rule.

The results are shown in figure (15). In the case where $\beta = 1$ (the gradient case) the minimizing path is given by a straight line between the fixed points. In the case where $\beta = 10$ the minimizing path follows an arc in the first half of the journey and then for the second half it follows a straight line from the critical point at $u = 0$ to $u = 1$. This result is the same as was found in [20].

5.1.3 Maier-Stein (PDE): gMAM

Now we can expand the problem covered in Section 5.1.2 to include a partial derivative on $u$ and $v$. The theory on using gMAM to solve PDEs is covered in Section (4.4). Consider the following set of equations:

$$
\begin{align*}
    u_t &= \kappa u_{xx} + u - u^3 - \beta uv^2 + \epsilon \eta_u, \\
    v_t &= \kappa v_{xx} - (1 + u^2)v + \epsilon \eta_v
\end{align*}
$$

Here we have two new terms, $\kappa u_{xx}$ and $\kappa v_{xx}$, affecting the evolution of the system. The presence of these terms means that this set of equations now qualifies as a set of non-linear second-order stochastic PDEs. Note that we assume that the noise terms $\eta_u$ and $\eta_v$ are uncorrelated.

The gMAM is employed to describe the most likely transition path between two stable fixed points. As explained in section (4.4) the technique employed to allow gMAM to deal with PDE’s is referred to as an infinite dimensional generalization. The idea is that in order to numerically compute the value of the partial derivative, it is necessary to simulate strings that are adjacent and coupled to one another in the $x$-direction so that we can use those other trajectories to find out how $u$ is likely to change with a change in $x$. This means that the simulation requires that we have a discretized space representing the $x$-dimension. Therefore, rather than having an initial condition be given by two points in space with a line connecting them (as is done in the gMAM ODE formulation), the initial
condition is given by two boundary lines that are connected by a surface representing adjacent trajectories. We choose that \( x \) be a variable that can take the values from 0 to 1, just like \( \alpha \) with the discretization in \( x \) having \( N_x = 128 \) points. The number of points in the \( \alpha \) dimension is \( N_{\alpha} = 100 \), therefore \( u \) and \( v \) each can be represented by a 100x128 grid of values.

The stable fixed points still exist at \((u,v) = (-1,0)\) and \((u,v) = (1,0)\). To proceed, we are forced to choose an initial condition. The initial condition recommended in [1] is given by the following equation:

\[
\begin{bmatrix}
\varphi(\alpha,x,\tau = 0) \\
u(\alpha,x,\tau = 0) \\
v(\alpha,x,\tau = 0)
\end{bmatrix} = \begin{bmatrix}-1 + 2\alpha + 2\sin^2(\pi\alpha)\sin^2(\pi x) \\
0
\end{bmatrix}
\]

This is a good choice because it smoothly varies with \( \alpha \) and \( x \), and has all the properties we want our solution to have. Namely that its \( u \) component is equal to -1 when \( \alpha = 0 \), equal to 1 when \( \alpha = 1 \), and that it is periodic across the \( x \) boundary. We want our \( x \) dimension to be periodic at the boundary because that allows us to employ spectral methods in that direction.

Now, since the partial derivative term in equation (122) is of second order in \( x \), and diffusion tensor is equal to unity, the analysis in section 4.4 gives us the explicit equation (Eq. (108)) for \( \varphi \) that we need to integrate to find the minimizer, :

\[
\dot{\varphi} = \lambda^2 \varphi'' - \lambda (\nabla b - (\nabla b)^T) \varphi' - (\nabla b)^T b - \kappa(\nabla b + (\nabla b)^T) \varphi_{xx} - \kappa^2 \varphi_{xxxx} - \kappa \left( \langle \varphi_x, \nabla \nabla b_1 \varphi_x \rangle \right) - \lambda' \varphi'.
\]
We construct the operator $B$ according to:

$$
\phi = B(u,v) + \varepsilon \eta
$$

which implies:

$$
B(u) = \begin{bmatrix}
    B_1(u,v) \\
    B_2(u,v)
\end{bmatrix} \equiv \begin{bmatrix}
    \kappa u_{xx} + b_1(u,v) \\
    \kappa v_{xx} + b_2(u,v)
\end{bmatrix}
$$

$$
b(u) = \begin{bmatrix}
    b_1(u,v) \\
    b_2(u,v)
\end{bmatrix} = \begin{bmatrix}
    u - u^3 - \beta uv^2 \\
    -(1 + u^2)v
\end{bmatrix}
$$

where $\nabla b$ is given in the Section 5.1.2 by equation (121), and $\nabla \nabla b_1$ and $\nabla \nabla b_2$ are the Hessian of $b_1$ and $b_2$ and can be written explicitly as:

$$
\nabla \nabla b_1 = \begin{pmatrix}
    -6u & -2\beta v \\
    -2\beta v & -2\beta u
\end{pmatrix}
$$
\[ \nabla \nabla b_2 = \begin{pmatrix} -2v & -2u \\ -2u & 0 \end{pmatrix} \]

Thus with these we can write all of the terms in equation (108) explicitly in terms of \( u, v, \varphi, \lambda, \) and \( \varphi' \):

\[
-\lambda (\nabla b - (\nabla b)^T) \varphi' = -\lambda \begin{bmatrix} 2uv(1-\beta) \varphi_2' \\ -2uv(1-\beta) \varphi_1' \end{bmatrix}
\]

\[
-(\nabla b)^T b = - \begin{bmatrix} (1 - 3u^2 - \beta v^2)(u - u^3 - \beta uv^2) + 2\beta uv(1 + u^2)v \\ -2uv(u - u^3 - \beta uv^2) + (1 + u^2)^2v \end{bmatrix}
\]

\[
-\kappa (\nabla b + (\nabla b)^T) \varphi_{xx} = -\kappa \begin{bmatrix} 2(1 - 3u^2 - \beta v^2) \varphi_{1xx} - 2uv(\beta + 1) \varphi_{2xx} \\ -2uv(\beta + 1) \varphi_{1xx} - 2(1 + u^2) \varphi_{2xx} \end{bmatrix}
\]

As section 4.4 points out, the \( \varphi_{xxxx} \) term is best handled by using a split step algorithm to stabilize the integration process. Finally, as always, parameterization (\( |\varphi'| = constant \)) must be enforced at the end of every integration step.

The result of following this procedure can be seen in figure 16. The figure shows \( u(\alpha, x) \) where each line in the figure represents a line of constant \( \alpha \). The trajectory proceeds from the bottom of the image at \( u(0, x) = -1 \) towards the top of the image \( u(1, x) = 1 \). This result is in good agreement with the result produced in [1].

Another result is given in figure (17) which shows that there is a saddle point that exists for \( \alpha = .53 \). This is yet another confirmation that we have indeed found the minimizing path.
Figure 17: Displays $\lambda$ versus $\alpha$ corresponding to the minimizing solution. The root at $\alpha = 0.5$ signals the existence of a saddle point and therefore suggests that our solution has captured the minimizer.

5.2 Real Ginzburg Landau Equation

The Real Ginzburg-Landau Equation (RGLE) appears in many areas of physics. Physical situations include: Rayleigh-Benard convection, Taylor-Couette flow, or flames which are stabilized on a burner. The equation is given by:

$$\dot{u} = u - u^3 + \kappa u_{xx} + \sigma f.$$  (123)

One of the interesting aspects of this problem is that without noise, the solutions always tend toward the stationary fixed points. That is to say the system eventually falls into one of the states $u = 1$ or $u = -1$ and will afterward remain in that state. If we incorporate noise, noise-induced transitions between the states become possible. We are interested in finding the minimizer of the associated Freidlin-Wentzell functional.

This equation represents a one-dimensional PDE that also happens to be symmetric and, therefore, it is a gradient system. This is fortunate for us, because it means we can use this problem as an example to demonstrate the functionality of both the string method
and the gMAM for PDEs.

5.2.1 Applying the String Method (RGLE)

The theory of the string method is outlined in section 4.2. As mentioned before, the string method is employed to describe the most likely transition path between two states, often the transition between two stable fixed points. The first step in applying the string method to a problem is to decide what fixed points are of interest in the given case. Generally, the procedure is that one picks the two end points, assumes a transition path to start the iteration, and then allows the numerical algorithm to relax the trajectory towards the minimizing path. The RGLE is a partial differential equation due to the appearance of the diffusive term $u_{xx}$. Just as we saw in section 4.4 and 5.1.3, we again have to extend our simulation into the x-dimension to calculate the $u_{xx}$ vector. The stable fixed points can be found by analyzing equation (123) as has been described earlier in the paper. In this case we find that there exist stable fixed points at $u = -1$, and $u = 1$. In the string method, we replace the variable $u(t,x)$ with a string $\varphi(\alpha,x,\tau)$ where $\tau$ represents an artificial time

![Figure 18: Displays the initial condition for the string method. We attempt to find the transition between $u = -1$ and $u = 1$ by creating a line of fixed endpoints corresponding to those two values at the start and the end of the transition.](image)
that corresponds to relaxation. Suppose that we wish to find the transition between those
two fixed points. As mentioned above, we need to choose a surface that connects the
two lines given by $\phi_0 = \phi(\alpha_1, x) = -1$ and $\phi_N = \phi(\alpha_N, x) = 1$. One possible choice for the
initial condition is:

$$u(\alpha, x, \tau = 0) = -1 + 2\alpha + 2\sin^2(\pi\alpha)\sin^2(\pi x)$$

It is the same initial condition that was suggested by Heymann and Vanden-Eijnden in [20]
for the Maier-Stein problem. It turns out to be applicable to many PDE problems
including the RGLE. Figure (18) shows what this initial condition looks like. The parameter
$\alpha$ has a value between 0 and 1 and represents the distance along the string measure in
arc length parametrization. In this case it was divided into $N_\alpha = 100$ discrete points.
Similarly, in this case $x$ also takes values between 0 and 1 and is divided into $N_x = 128$
points. Additionally, there is a periodic boundary condition on the $x$ dimension such that
$\phi(\alpha_n, x_1) \approx \phi(\alpha_n, x_N)$. This periodic boundary is chosen so that a Fast Fourier Transform
can be applied to numerically calculate the $u_{xx}$. Following this, the procedure is very
simple. The deterministic part of equation (123) represents the gradient of a potential
function, therefore the relaxation of the string is given by:

$$-\nabla V = u - u^3 - \kappa u_{xx}$$

In order to carry out this scheme, we must construct the $u_{xx}$ vector. There are a number
of ways to find the partial derivative, $u_{xx}$. As mentioned, spectral methods work well in
this case. Whatever method is used to extract $u_{xx}$, the next step is to integrate $\phi$ along
the artificial time $\tau$ until the system relaxes into the minimizing state.

$$\phi(\alpha, x, \tau_n) = -\nabla V \quad (124)$$
One particularly simple choice for integrating this equation is a forward Euler-scheme, which works in many cases as long as \( d\tau \) is sufficiently small:

\[
\phi(\alpha, x, \tau_{n+1}) = \phi(\alpha, x, \tau_n) - \nabla V \cdot d\tau
\]

Finally, there is a condition that relates the local value of alpha to the distance in u that one travels along the string. Whatever parameterization is chosen, there needs to be a condition in the code such that the parameterization of \( \alpha \) will be maintained after the string \( \phi \) is sufficiently deformed by the force \(-\nabla V\). For the solution corresponding to figure (19), we choose the parameterization such that \(|d\phi/d\alpha| = constant\). Figure (19) shows the minimizer that gives the most likely transition path between the two points \( u = -1 \) and \( u = 1 \).

### 5.2.2 Applying the gMAM (RGLE)

The set up for the gMAM is essentially very similar to the situation for the string method, although there is some degree of complication beyond that.
To restate the problem: We wish to find, for equation (123), the most likely transition path between stable fixed points \( u = -1 \) and \( u = 1 \). Just as in the string method, the gMAM seeks to construct an effective string \( \phi(\alpha, x, \tau) \) which, given some forcing terms, will eventually relax into the minimum path solution as the artificial time \( \tau \) progresses. The case of PDEs requires again that our initial condition be given by a surface that connects the lines: \( \phi_0 = \phi(0, x) \) and \( \phi_N = \phi(1, x) \), see Fig. (18). gMAM differs from the string method when it comes to finding the forcing term that will relax the system over time. Rather than simply using \(-\nabla V\) as we had in the previous section, gMAM prescribes as algorithm to compute the updated path \( \tilde{\phi} \) the following discretization:

\[
\frac{\Phi_i - \Phi_i}{\Delta \tau} = (\lambda_i)^2 \frac{\Phi_{i+1} - 2\Phi_i + \Phi_{i-1}}{1/N^2} - \lambda_i H_{\theta x} \phi'_i + H_{\theta \theta} H_x + \lambda_i \lambda'_i \phi'_i,
\]

where \( H \) is the Hamiltonian according to the Friedlin-Wentzell theory (see eq. (68)). All operators and their derivatives are evaluated using the current string \( \phi \). In the case of a stochastic PDE, we obtain the following expression:

\[
\phi = \lambda^2 \phi'' - \lambda (\partial B - (\partial B)^T) \phi' - (\partial B)^T B + \lambda \lambda' \phi'.
\]  

(126)

In order to carry out the integration, we have to construct the operator \( B \) according to:

\[
\dot{u} = B(u) + \sigma f \rightarrow B \equiv \kappa u_{xx} + b(u), \quad b(u) = u - u^3.
\]

Note that is useful to split up the operator \( B(u) \) into the ordinary part \( b(u) \) and the diffusive partial differential part \( u_{xx} \). As we will see below, this will make it clearer how to compute the partial derivative of the Hamiltonian \( H \) present in the gMAM iteration algorithm.

The \( \partial B \) terms in Equation (126) represent operators that act on \( \phi' \). The operator is
defined as:

\[ \partial B = \nabla b + \kappa \partial_{xx}. \]

With this construction we can now write out the terms given by equation (126). Since our problem is one-dimensional there are cancellations in various terms in the gMAM algorithm. In particular, we have \( \partial B = \partial B^* \equiv \partial B^T \) and therefore \( \lambda \left( \partial B - (\partial B)^T \right) \phi' = 0. \) We end up with a scheme that looks like this:

\[ \dot{\phi} = \lambda^2 \phi'' - \kappa^2 \phi_{xxxx} - (b(\phi)B(\phi) + \kappa b_{xx}(\phi)). \]

Once these terms are written out, all that remains is to integrate the equation and, as with the string method, to ensure that the parameterization of choice (in this case \( |\phi'| = \text{const} \)) is maintained. In our implementation in Matlab, we chose an implicit vectorized solver. The resulting linear system for the string update \( \tilde{\phi} \) can be solved very efficiently using the backslash operator.
Figure 21: Displays $\lambda$ vs $\alpha$ corresponding to the minimizing solution. The root at $\alpha = .5$ signals the existence of a saddle point and therefore suggests that our solution has captured the minimizer.

Just as we saw in the Maier-Stein problem, it is good to use a split-step to stabilize the integration process where in the first part of the step, we have

$$\dot{\phi} = \lambda^2 \phi'' - (b(\varphi)B(\varphi) + \kappa b_{xx}(\varphi))$$

and then in the second step we have only the 4th derivative term:

$$\dot{\phi} = -\kappa^2 \phi_{xxxx}$$

The solution that we arrive at can be seen in figure (20), which is essentially identical to the solution for the string method (Fig. (19)). For a simple case like this, the string method is more practical given that it is generally easier to implement.

Lastly, we can see that this is indeed the minimizing path if we look at the plot of $\lambda$ vs $\alpha$ in figure (21). We can see that half way through the trajectory, that $\lambda = 0$. This signals to us that there is a saddle point. For the case of the RGLE, it can be shown that the
minimizing trajectory will go through this saddle point.

6 Instantons of the cubic nonlinear Schrödinger equation

6.1 Mode-Locked Laser PDE/Introduction

One of the very important applications of gMAM is in the field of optics. A lot of work has been done in the goal of understanding the stable solutions of the nonlinear Schrödinger equation in the context of fiber optics and laser systems ([62], [63], [64],[22, 65-67]). Here we present the application of gMAM to a PDE corresponding to a phenomenological model of an ultrashort pulse mode-locked laser (MLL):

\[ iA_z + A_{tt} + \frac{1}{2}|A|^2A = -b\cos(\omega t)A - ic_1A \]
\[ + ic_2A_{tt} + id_1|A|^2A - id_2|A|^4A + i\epsilon f(t,z) . \]

This SPDE is very complex. Let us look at simpler cases before tackling this challenging problem using gMAM.

6.2 NLSE and Solitons

In Section 1.2 it was mentioned that for optical systems in a medium that has a nonlinear index of refraction, the propagation of the electric field is described by the nonlinear Schrödinger equation (NLSE) [68, 69]. It was also mentioned that this fact implies that there exist soliton solutions corresponding to the optical signal in the medium [70]. In this section we introduce some of the particulars of those kinds of solutions. Consider the simplest form of the NLSE:

\[ iu_z + u_{xx} + 2|u|^2u = 0. \]
Where \( u = u(z, x) \) with \( z \) being the evolution variable, and \( x \) being the transverse variable. There are known closed-form solutions to the NLSE. A particularly simple solution to this equation is a single soliton. A soliton is loosely defined as a "solitary wave" and exists due to a balance of dispersion (the \( u_{xx} \) term) and nonlinearity (here the cubic nonlinear term \( 2|u|^2u \)). For a single soliton, there is a localized peak in the solution profile, and outside the solitary peak, there are no others; the magnitude of the envelope of the electric field decreases as a function of distance from the central peak. The particular soliton that corresponds to eq. (128) is given by the following:

\[
E[\phi^2] = 0
\]

\[
E[\phi^2] = (.002)^2
\]

Figure 22: Soliton solutions for NLSE for \( E[\phi^2] = 0 \) and \( E[\phi^2] = .002^2 \) case. The small noise in the second case has the effect of destabilizing the soliton and displacing it from its natural equilibrium position.

\[
u(x, z) = a \text{sech} (ax)e^{ia^2z}
\]  

We can check that this indeed satisfies eq. (128) by computing the appropriate derivatives of the assumed solution. Since we are interested in modeling real optical systems, we need to consider the effects that noise in a system will have on our soliton. As we touched on earlier in Section 1.2, laser systems are inherently statistical, so there is always going
George Poppe  

Applications of gMAM

to be some noise in a real optical application. To begin studying the effects of noise, we can slightly modify eq. (128) to include a noise term $\phi$ as additive noise on the right-hand side.

$$iu_z + u_{xx} + 2|u|^2 u = \phi$$

Consider the case that $\phi$ represents a white noise, that is $\delta$-correlated in time and space. In order to get an idea of how this will affect our system, we can turn to numerical methods to simulate the NLSE with such noise. For the particular case of figure 22 a split-step algorithm was used.

### 6.3 Chernykh-Stepanov Step applied towards the cubic diffusive NLSE

We want to apply the gMAM to the NLSE. In order to do that effectively, we need a method to test how accurate the gMAM solution is. One such method is to write down the instanton equations associated with the system and integrate using a Chernykh-Stepanov step in a simulation.

To see how to approach this, we will take the cubic diffusive NLSE as an example problem.

$$A_t = -\alpha A + \kappa A_{xx} + i\epsilon A|A|^2 A + i\epsilon f(x,t)$$ (130)

Firstly we will quickly review how to apply the gMAM to this type of PDE. The field $A(x,t)$ is complex, so we need to split it into real and imaginary parts as we had done in 4.5.2.

$$A = u + iv$$
this substitution results the B vector:

\[
B_1(u,v) = -\alpha u + \kappa u_{xx} - dv_{xx} - c(u^2 + v^2)v \\
B_2(u,v) = -\alpha v + \kappa v_{xx} + du_{xx} + c(u^2 + v^2)u
\]

Now we must select the boundary conditions. Suppose that want to study the transition from a small amplitude soliton to a larger amplitude soliton. In that case, we have the initial condition:

\[
\begin{bmatrix}
u_0 \\
v_0
\end{bmatrix} = \begin{bmatrix}
\frac{\lambda_0}{\cosh(\lambda_0 x)} \\
0
\end{bmatrix}
\]

And the final condition:

\[
\begin{bmatrix}
u_1 \\
v_1
\end{bmatrix} = \begin{bmatrix}
\frac{\lambda_1}{\cosh(\lambda_1 x)} \\
0
\end{bmatrix}
\]

Where \(\lambda_0\) is small relative to \(\lambda_1\), say \(\lambda_0 = .2\) and \(\lambda_1 = 2.5\). Now, with the problem appropriately defined, we can apply gMAM by integrating the equation of relaxation given by Equation (106). This operation is very similar to that performed in Section 4.5.2. Writing the gradient of \(b\) gives:

\[
\nabla b = \begin{pmatrix}
b_{1u} & b_{1v} \\
b_{2u} & b_{2v}
\end{pmatrix} = \begin{pmatrix}
-\alpha - 2cvu & -cu^2 - 3cv^2 \\
cv^2 + 3cu^2 & -\alpha + 2cvu
\end{pmatrix}
\]

The Hessian of the two components \(b_1\) and \(b_2\) can be written:

\[
\nabla \nabla b_1 = \begin{pmatrix}
b_{1uu} & b_{1uv} \\
b_{1uv} & b_{1vv}
\end{pmatrix} = \begin{pmatrix}
-2cv & -2cu \\
-2cu & -6cv
\end{pmatrix}
\]

\[
\nabla \nabla b_2 = \begin{pmatrix}
b_{2uu} & b_{2uv} \\
b_{2uv} & b_{2vv}
\end{pmatrix} = \begin{pmatrix}
6cu & 2cv \\
2cv & 2cu
\end{pmatrix}
\]
As in other cases, we integrate the $\phi_{xxxx}$ term separately from the rest of the equation so that in the second integration step, we have:

$$\phi = -(\kappa^2 + d^2)\phi_{xxxx}$$

Figure 23 shows a result of applying the gMAM to find this transition.

We will now examine one technique to verify the validity of the minimizer acquired via gMAM. The Chernykh-Stepanov (CS) step is a two step method of integration where in the first step, the position fields ($u$ and $v$) are integrated forward in time, and in the second step, the momentum fields ($\theta_u$ and $\theta_v$) are integrated backwards in time [71].

Both integration steps are carried out using an exponential integrator algorithm. Let us briefly review how that works. Suppose we write the expression:

$$u_t = Lu + R(u)$$  \hspace{1cm} (134)

Where $u_t$ is the time derivative of $u$, $L$ is a linear operator and $R(u)$ is a nonlinear function.
of \( u(t) \). Then this implies:

\[
    u_t - Lu = R(u) \tag{135}
\]

\[
    e^{-Lt}u_t - e^{-Lt}Lu = \frac{d}{dt}(e^{-Lt}u) = e^{-Lt}R(u) \tag{136}
\]

if we take the integral on both sides we obtain

\[
    e^{-Lt}u(t) - u(0) = \int_0^t e^{-Ls}R(u(s))ds \tag{137}
\]

\[
    e^{-Lt}u(t) = u(0) + \int_0^t e^{-Ls}R(u(s))ds \tag{138}
\]

\[
    u(t) = e^{Lt}u(0) + e^{Lt} \int_0^t e^{-Ls}R(u(s))ds \tag{139}
\]

\[
    u(t) \approx e^{Lt}u(0) - e^{Lt}e^{-Lt} - \frac{1}{L}R(u(0)) \tag{140}
\]

Where in the last step we assume that for small times \( t \), the nonlinear function, \( R(u) \), does not change much. If we define new terms \( M_1 \) and \( M_2 \) in the following way:

\[
    M_1 \equiv e^{Lt}, \quad M_2 \equiv \frac{1 - e^{L\Delta t}}{L} \tag{141}
\]

\[
    \text{then our integrating step can be written:}
\]

\[
    u_{\text{new}} = M_1u_{\text{old}} - M_2R(u_{\text{old}}) \tag{143}
\]

Where, in the first step of the integration, we take \( M_2 = -\Delta t \). Recall from Section 3.2 that the instanton equations of motion are given by equation (74), where we had the the position field, \( u \), and the momentum field, \( \theta \). On the other hand, when we carry out the gMAM in the case where the field is complex, the result gives us four fields, which we
want to call $u,v$, $\theta_u$, and $\theta_v$. In this case we use the complex variables $A$ to denote the

$$A = u + iv, \quad \beta = \theta_u + i\theta_v$$

(144)

We will take the boundary condition on the position variable $A$ is to be: $A(-\infty) = 0$ and $A(0,x) = ae^{-|x|\sqrt{\kappa}}$.

Now we will apply the CS step to the diffusive NLSE with a cubic nonlinearity. In that case, our instanton equations of motion can be written:

$$A_t = -\alpha A + \kappa A_{xx} + idA_{xx} + ic|A|^2A + \beta$$

(145)

$$\beta_t = \alpha \beta - \kappa \beta_{xx} + id\beta_{xx} + 2ic|A|^2\beta - icA^2\beta^*$$

(146)

Now, as described above, we need to integrate the field $A$ forward in time from $t = -\infty$ to
Figure 25: Here we see a plot of the difference between the two solutions of the momentum fields corresponding to the diffusive cubic NLSE PDE (One solution resulting from the gMAM, the other resulting from exponential integration according to Chernykh-Stepanov step). For the momentum fields, the CS step prescribes that we integrate backwards in time towards $t = -\infty$. The agreement between the two solutions is not as strong as we saw in figure 24. (peak height of $\theta_u \approx .6$)

$t = 0$ and verify that the simulation corresponds to the result that gMAM gives us. Figure 24 shows the comparison of the results of these simulations. The difference between the two solutions is very small which supports the case for using the gMAM to model these kinds of PDE's.

For the momentum field we integrate time in the other direction so that we start at time $t = 0$ and finish at time $t = -\infty$. The results of carrying out this simulation and comparing it to our gMAM solution is shown in figure 25. Unlike in the case of the the position fields, we see only moderate agreement between the results of these two methods for the momentum fields $\theta_u$ and $\theta_v$.

6.4 Nonlinear OU comparison using Instantons

In exploring the numerical landscape of the perturbative nonlinear PDEs, and as we saw in the previous section, it was found that small nonlinearity strength resulted in relatively small differences in the values of the position fields when comparing the linear result to
the nonlinear, however the difference between the momentum fields of the linear and nonlinear cases showed larger differences in general. In an attempt to understand this observation, we can try investigating a similar analytical system. Recall that in section 3 we introduced the technique of using instanton calculations to find analytical soliton solutions for SDE systems. Take the simple case of the nonlinear Ornstein-Uhlenbeck instanton problem given by the following equations:

\[ x = -kx - cx^3 + p, \quad p = kp + 3cx^2p \]

This implies our drift vector \( b \) is given by:

\[ b(x) = -kx - cx^3 \quad (147) \]

To solve this system, we can write the Hamiltonian associated with these equations:

\[ H = \frac{p^2}{2} + bp = 0 \quad (148) \]

This equation is true for two values of \( p \):

\[ p = 0 \quad \text{and} \quad p = -2b = 2k + 2cx^3 \]

Ignoring the zero solution for now, we begin by writing the expression for \( \dot{x} \):

\[ \dot{x} = b + p = -b = kx + cx^3 \quad (149) \]

The solution to this differential equation is:

\[ x(t) = \frac{a\sqrt{a^2ce^{-2kt} - ca^2} + ke^{-2kt}}{a^2ce^{-2kt} - ca^2 + e^{-2kt}k} \quad (150) \]
For the solution to the linear case, we can set $c=0$:

$$\bar{x}(t) = \frac{kae^{-kt}}{ke^{-2kt}} = ae^{kt} \quad (151)$$

With these, we can get a value for the difference between the linear case and the nonlinear case.

$$\tilde{x}(t) \equiv x(t) - \bar{x} = \left( \sqrt{\frac{k}{a^2c(1-e^{2kt})}} - 1 \right) ae^{kt} \equiv \omega(t) ae^{kt} \quad (152)$$

If $a = 1$, $k = 1$, and $c = 2$ (typical values for soliton solutions) we see that $\omega \approx 1$. This means that the difference $\tilde{x}(t)$ is going to be small (keeping in mind that in our interval $t = -\infty \to 0$, time is always negative). This result corroborates our experience that the difference in the position fields that result from the linear and nonlinear case is not that large.

If we look at the momentum fields we can write the same difference equation:

$$\tilde{p} = p - \bar{p}$$

with

$$p(t) = 2kx + 2cx^3, \quad \bar{p}(t) = 2k\bar{x}$$

From these we can write the difference which gives:

$$\tilde{p}(t) = p(t) - \bar{p}(t) = 2k(x - \bar{x}) + 2cx^3$$

$$\approx 2k\bar{x} + 2cx \approx 2cae^{kt}$$
This shows that $\tilde{p}$ is about four times as large as $\tilde{x}$, which explains why we see such a large difference when comparing the momentum fields derived using the CS step as compared to the momentum fields derived from the gMAM (Section 6.3).

6.5 Kerr Lens Mode-Locked Laser SPDE

Having seen a number of simpler cases of NLSE systems, we are now ready to apply gMAM to the SPDE presented in equation (127). As mentioned at the beginning of the section, this SPDE describes a pulse signal associated with a Kerr lens MLL with feedback. The SPDE of interest is given by:

$$iA_z + \frac{c_3}{2} A_{tt} + c_4 |A|^2 A = -b \cos(\omega t) A - ic_1 A + ic_2 A_{tt} + id_1 |A|^2 A - id_2 |A|^4 A + i\epsilon f(t, z).$$

Here we have rewritten Eq: (127) with the addition of two new parameters ($c_3$ and $c_4$). These were added so that we can work with a larger parameter space. The parameters $b, \omega, c_1, c_2, d_1, d_2, c_3$ and $c_4$ are constants and $f(t, z)$ represents the noise in the system. One can see that this has the form of a nonlinear Schrödinger equation with an $|A|^2$ nonlinearity as well as an $|A|^4$ nonlinearity. One of the issues we need to tackle is the fact that, in this case, $A$ is a complex number. We saw before with the linear Schrödinger equation that we could pull apart the real and imaginary components by setting $\Psi = u + iv$. So, we will follow our previous examples and write $A$ as a two component vector where the components are the real part ($u$) and imaginary part ($v$) respectively. If we carry out this operation we get a pair of coupled SPDE’s:

$$u_z = c_2 u_{tt} - \frac{c_3}{2} v_{tt} + \left[ -b \cos(\omega t) - c_4 (u^2 + v^2) \right] v - \left[ c_1 - d_1 (u^2 + v^2) + d_2 (u^2 + v^2)^2 \right] u$$
George Poppe

Applications of gMAM

\[ v_z = \frac{c_3}{2} u_{tt} + c_2 v_{tt} - \left[ -b \cos(\omega t) - c_4 (u^2 + v^2) \right] u - \left[ c_1 - d_1 (u^2 + v^2) + d_2 (u^2 + v^2)^2 \right] v \]

Once we have it in this form, we can see the similarity to other problems we have looked at earlier. (Maier-Stein, linear Schrödinger).

Next we want to figure out what are the boundary conditions that correspond to a transition between two stable fixed points. For this case, it turns out to be a non-trivial exercise, however, one possibility is to find the fixed point numerically. We can locate the fixed point of interest by simulating a soliton that has an initial position nearby that fixed point and allowing the deterministic version of the system to relax into the stable state via integration of the equations of motion. Since we are taking the deterministic case, there is no random component, and therefore the soliton will be slowly drawn into an existing nearby stable state. We know that when the soliton stops moving that we have found the fixed point. The value of the fixed point is obviously going to vary depending on the chosen parameters of the system so we have to carry out this search each time we run the simulation. Following this logic, we use an RK4 algorithm to integrate the PDE directly in order to resolve the boundary conditions.

Figure 26: Displays the solution of the u and v fields generated via gMAM for the PDE laser model. This figure was generated using parameters \( c_1 = 0.01, c_2 = 0.1, c_3 = 1, c_4 = 1, d_1 = 0.3, d_2 = 0.1, b = 0.85, \omega = 1.5 \).

With the boundary conditions resolved, we need to get our PDE into the proper form so
that we can apply the gMAM. We need to express the PDE according to:

\[ A_z = B(u, v) + \sigma(u, v) \xi \]

Giving \( B(u, v) \) the same form as before:

\[ B(u, v) = \kappa \begin{pmatrix} u_{xx} \\ v_{xx} \end{pmatrix} + b(u, v); \quad (154) \]

Which in this case implies

\[ \kappa = \begin{pmatrix} c_2 & -c_1 \\ -c_1 & c_2 \end{pmatrix} \quad (155) \]

\[ b(u, v) = \begin{pmatrix} \left[ -b \cos(\omega t) - c_4(u^2 + v^2) \right] v - \left[ c_1 - d_1(u^2 + v^2) + d_2(u^2 + v^2)^2 \right] u \\ \left[ -b \cos(\omega t) - c_4(u^2 + v^2) \right] u - \left[ c_1 - d_1(u^2 + v^2) + d_2(u^2 + v^2)^2 \right] v \end{pmatrix} \quad (156) \]

Now that we all we have to do is apply Equation (100) which reduces to Equation (108). Where we, again use a split step to handle the \( \phi_{xxxx} \) term.

Figure 26 displays the results of the computation.

### 7 ODE reduction and comparison to PDE model

#### 7.1 NLSE with Cosine Potential

When dealing with numerical simulations, one of the main challenges is to assess whether a numerical result is valid. There is a paradoxical nature such that we wish to use numerical methods to solve problems that cannot be solved analytically, however if we cannot solve it analytically, it is often difficult to be confident about the accuracy of the numerical
result. It is therefore important to experiment with simple examples that can be checked analytically. To this end, another useful introductory case to consider is the problem of the deterministic (no noise) NLSE with a perturbing cosine potential.

\[ iu_t + u_{xx} + 2|u|^2u = -\varepsilon u \cos(kx) \] (157)

If we assume that the cos potential is slowly varying in space, then Scharf and Bishop showed how one can analytically treat this problem in their 1992 paper [72]. The basic idea is that one can propose ansatz solution of the form:

\[ u(x,t) = 2\eta e^{i\frac{q}{2} x - i\Phi(t)} \cosh \left[ 2\eta(x-q) \right] \]

The parameter \( q \) in this case represents the canonical position of the soliton. By making this ansatz, we can transform the problem into an effective classical Hamiltonian system with the soliton acting as a particle in a cosine potential well. For a single soliton, the effective Hamiltonian becomes:

\[ H_1 = \frac{p^2}{2\eta} - \frac{\varepsilon k\pi}{2\sinh(k\pi/4\eta)} \cos(kq) \]

From this we see that the \( \eta \), which modulated the amplitude and width of the wave \( u(x,t) \), acts as a mass in this Hamiltonian approximation. Notice also that the depth of the well in the Hamiltonian picture is no longer proportional to \( \eta \), but depends on the ratio of the spatial frequency of the well to the width of the soliton \( k/\eta \).

If we take this Hamiltonian at face value, we can construct the equations of motion that give the trajectory of the soliton.

\[ \frac{\partial H}{\partial q} = -\dot{p} = \frac{k^2\varepsilon \pi}{2\sinh(\frac{\pi k}{4\eta})} \sin(kq) \]

\[ \frac{\partial H}{\partial p} = \dot{q} = \frac{p}{\eta} \]
George Poppe

Applications of gMAM

Figure 27: Comparison of simulated solutions of the deterministic NLSE with a cosine potential vs the simulated solution of the ODE that results from the effective Hamiltonian. The cos potential is shown in red; the PDE solution is in blue with "x" markers; the ODE solution is in green. The agreement between the two is very good. For this particular case, the soliton is in a bound state such that it bounces back and forth between the peaks of the cosine potential.

This results in a second order nonlinear ODE, similar to the equation of the pendulum:

\[ \ddot{q} = -k^2 \left( \frac{\varepsilon \pi}{2\eta \sinh\left(\frac{\pi k}{4\eta}\right)} \right) \sin(kq) \]

We can integrate this equation numerically alongside the PDE in Equation (157) as a test of its validity. Figure 27 shows the results of one such simulation. The cos potential is shown in red; the PDE solution is in blue with "x" markers; the ODE solution is in green. The agreement between the two is very good to the point where it is difficult to tell the two solutions apart. For this particular case, the soliton is in a bound state such that it bounces back and forth between the peaks of the cosine potential. One interesting idea that comes out of this is that, if we add a noise term and tune it just right, we may be able to construct a quasi-bound state such that the soliton will be trapped in the potential well most of the time, but every now and then it will tunnel to one of the adjacent wells in the periodic potential. Altering the problem to include additive noise, we consider now:
Figure 28: Simulated solution to the NLSE cosine potential with additive Gaussian white noise. The parameters of the problem are tuned so that the destabilizing effects of the noise are nearly in balance with the stabilizing effects of the potential well. This creates a tunneling phenomenon with a metastable bound state. The top image displays the soliton at time \( t = 2 \), while the bottom shows the position at time \( t = 40 \). The soliton has "tunneled" from one metastable bound state into a nearby one.

\[
i u_x + u_{xx} + 2 |u|^2 u = -\varepsilon u \cos(kx) + \sigma \phi
\]  

(158)

\( \phi \) in this case is Gaussian white noise. This was simulated using a split step algorithm with good success. For the correctly tuned parameters, the soliton solution will show quasi-stable behavior such that it will usually say bound in the potential well with some low probability of tunneling into a nearby well.
7.2 Moore’s Model

Let us look at another model that we can use to understand gMAM as well as gaining insight into the relationship between an SODE and a corresponding SPDE. Earlier we discussed the physical effects of passing light through a Kerr lens.

In the current work, Moore showed that there is a trade-off between line width and loss-of-lock frequency in such laser systems.

A phenomenological model of a Kerr lens MLL is given by Equation (153).

\[
\begin{align*}
    iu_z + \frac{c_3}{2} u_{tt} + c_4 |u|^2 u &= -b \cos(\omega t) u - ic_1 u \\
    &+ ic_2 u_{tt} + id_1 |u|^2 u - id_2 |u|^4 u + i \varepsilon f(t, z) \tag{159}
\end{align*}
\]

Where we have rewritten the SPDE given in equation (153) such that \( u(z, t) \) has replaced \( A(z, t) \) as the variable denoting the complex field. The reader may notice that, similar to an example we looked at earlier, the cosine term represents a potential well whose valleys, for certain choices of the parameters, are going to have bound states for the solitons we are looking for. The noise term \( f(t, z) \) is again Gaussian white noise. We have already seen how to apply gMAM to solve the transitions of the SPDE in section 6.5. Let us now examine a corresponding SODE. In 2014, Richard Moore discussed an ODE approximation to this model with the assumption that the solution will take the form:

\[
    u_s(t, z) = A(z) \text{sech}[A(z)(t - T(z)) + i\phi(z) + i\Omega(z)] \tag{160}
\]

This form assumes that the amplitude is in the shape of a \( \text{sech} \)-pulse. Here the parameters \( A, T, \Omega, \) and \( \phi \) are functions only of \( z \). In order to resolve what the functional behavior of these variables ought to be, we have to solve a modified version of the Euler-Lagrange equations associated the SPDE ([2],[73]). The specifics of this calculation are detailed in
the Appendix, section 9.1 of the thesis, however the general approach and results of [2] are summarized as follows. The Lagrangian density is given by:

\[
L[t, z, u(t, z), u_t(t, z), u_z(t, z)] = \text{Im}(u^* u) - \frac{c_3}{2} |u_t|^2 + \frac{c_4}{2} |u|^4 + b \cos(\omega t) |u|^2
\]  

(161)

If we integrate the Lagrangian density with respect to t, we arrive at the Lagrangian, \( L \), from which we can get the Euler-Lagrange equation:

\[
\frac{\partial L}{\partial q_j} - \frac{d}{dz} \frac{\partial L}{\partial q_{jz}} = 2 \text{Re} \left[ \int i \left[ -c_1 u + c_2 u_t + d_1 |u|^2 u - d_2 |u|^4 u + \varepsilon f(x, t) \right] \frac{\partial u^*}{\partial q_j} dt \right]
\]  

(162)

Here, \( q_j \) are the variables that control the dynamics of the system. With \( j = 1, 2, 3, 4 \) corresponding to \( [A, \Omega, T, \phi] \) respectively. In other words, this gives us 4 equations. Our goal is to write these equations as an SODE, so we are interested in solving for the derivative of each of the variables with respect to \( z \). Our SODE should take the following form:

\[
dU = F(U)dz + \varepsilon \sigma(U)dW
\]  

(163)

Now, a difficulty arises for those not familiar with stochastic problems. In (162) we are taking an integral over the random variable, \( f(t, z) \).

\[
\Xi_q \equiv 2 \text{Re} \left[ \int i \varepsilon f(t, z) \frac{\partial u^*}{\partial q} dt \right]
\]  

(164)

At a glance, it may not be clear how one can carry out the integration because \( f \) is a random variable. The critical idea is that we are only interested in the distribution that this integral represents. Looking at (163) we can see that we are interested in representing this integral equation as \( \varepsilon \) (some small constant) times \( \sigma \) (some matrix multiplying white noise \( dW \)). To this end, we can find the variance (\( \sigma^2 \)) by finding \( \Xi^2 \). Remember that
\[ \sigma_2^2 = \langle \Xi^2 \rangle - \langle \Xi \rangle^2 \] (165)

and since in our case the randomness is centered about zero, the distribution is going to be symmetric about the y axis.

By carrying out this method, one arrives at the solution:

\[
F(U) = \begin{bmatrix}
-2c_1U_1 + \left( \frac{4}{3}d_1 - \frac{2}{3}c_2 \right)U_1^3 - \frac{16}{15}d_2U_1^5 - 2c_2U_1U_2^2 \\
-\frac{4}{3}c_2U_1^2U_2 - \frac{\pi b \omega^2}{2U_1} \text{csch} \left( \frac{\pi \omega}{2U_1} \right) \sin(\omega U_3) \\
c_3U_2
\end{bmatrix}
\] (166)

\[
\sigma(U) = \begin{bmatrix}
\sqrt{U_1} & 0 & 0 \\
0 & \sqrt{\frac{U_1}{3}} & 0 \\
-\frac{U_3}{U_1} & 0 & \sqrt{\frac{\pi^2}{12U_1^2} + \frac{U_2^2}{U_1}}
\end{bmatrix}
\]

Where here \((U_1, U_2, U_3) \equiv (A, \Omega, T)\). As can be seen the SDE that describes the dynamics has only three vector components. This is because the fourth vector component \(\phi\) is actually not independent from the other three. \(\phi\)'s value is entirely determined by the value of the other three variables according to:

\[
d\phi = \left[ \frac{1}{2}((2c_4 - c_3)A^2 - c_3\Omega^2) + \frac{4}{3}c_2A^2\Omega T + \frac{\pi b \omega^2}{2A} \left( \frac{T \sin(\omega T)}{\sinh \left( \frac{\pi \omega}{2A} \right)} \right) - \frac{\pi^2 \omega^2 b \cos(\omega T) \coth(\frac{\pi \omega}{2A})}{4A^2 \sinh \left( \frac{\pi \omega}{2A} \right)} \right] dz
\]

\[+ \varepsilon \left( \frac{\sqrt{12 + \pi^2}}{6\sqrt{A}} \right) dW_4 + \varepsilon \left( T \sqrt{\frac{2A}{3}} \right) dW_2
\]

### 7.3 Deterministic Comparison of Laser PDE vs ODE reduction

Having constructed the ODE reduction for the MLL PDE in the previous section (Section 7.2), we can now follow the same methodology that we applied to the cosine potential in
section 7.1 in order to compare the deterministic model of the PDE to that of the ODE. To start, we need to find the corresponding fixed points in both models. We can locate the fixed point of interest by using the technique that we used to locate fixed points for the PDE in section 6.5. In particular, by simulating the equations of motion for an initial state given by a soliton near a chosen fixed point, we will find that the soliton will slowly move towards the fixed point until reaching the fixed point where it will come to a stop.

Figure 29: Comparison of simulated solutions of the deterministic phenomenological MLL PDE with vs the simulated solution of the ODE that results from the reduction carried out in [2]. The cosine potential is shown in red; the PDE solution is in blue with "x" markers; the ODE solution is in green. Just as with the NLSE with a cosine potential, there is good agreement between the solutions of the two cases.

We can carry this methodology out for both the PDE case and the ODE case and make a comparison of the resulting stable fixed points. If the ODE reduction is valid, there should be a correspondence in the solutions of the two systems just as there was in the case of the cosine potential in section 7.2. We can use the equation that relates the PDE variables to the ODE variables (Equation (160)) to find what the corresponding initial conditions for the two systems ought to be. We construct an integration scheme for the
ODE using the initial condition:

\[ U_s = \frac{A_0}{\cosh(A_0(t - t_{off}))} \]  \hspace{1cm} (167)

This implies that we assume that the phase \( \phi = 0 \) and \( \Omega = 0 \) to start. We then integrate along the z direction. The hope is that we will find a correspondence between the two frames and thereby validate the approximation. In fact, there is good agreement between the two in the deterministic case (this can be seen in figure (29)). Given this agreement, we should be able to find corresponding transitions between fixed points for the gMAM solution of the ODE and the gMAM solution of the PDE.

7.4 gMAM for Moore’s Model (SODE)

Now that we have the SDE (reproduced below), we want to use gMAM to find the transition path between two stable fixed points. Note: this section follows the paper by R. Moore [2].

\[ dU = F(U)dz + \epsilon \sigma(U)dW \]

For starters, we need to find the stable fixed points. This is a familiar exercise. We first set our drift vector from Eq(166) equal to zero. This implies that the rate of change of \( U \) is zero, and thus identifies fixed points.

\[
0 = \begin{bmatrix}
-2c_1U_1 + \left( \frac{4}{3}d_1 - \frac{2}{3}c_2 \right)U_1^3 - \frac{16}{15}d_2U_1^5 - 2c_2U_1U_2^2 \\
\frac{4}{3}c_2U_1^2U_2 - \frac{\pi b \omega^2}{2U_1} \text{csch}(\frac{\pi \omega}{2U_1}) \sin(\omega U_3) \\
c_3U_2
\end{bmatrix}
\]  \hspace{1cm} (168)
This gives us the conditions on \((U_1, U_2, U_3)\):

\[
\begin{pmatrix}
U_1 \\
U_2 \\
U_3
\end{pmatrix} = \begin{pmatrix}
\pm \sqrt{\frac{5}{16 d_2} \left[ 2d_1 - c_2 + \sqrt{(2d_1 - c_2)^2 - \frac{96}{5} c_1 d_2} \right]} \\
0 \\
n\pi/w
\end{pmatrix}
\] (169)

Where \((2d_1 - c_2)^2 > \frac{96}{5} c_1 d_2\) and where \(n\) is an integer. We can use these stable fixed points as initial conditions in our application of the gMAM.

Let us go through the implementation in close detail. To start, remember that in gMAM we are conceiving of the transition path as a kind of string that evolves in time and is given by \(\gamma : \gamma = \varphi(\alpha) : \alpha \in [0, 1]\) with a parameterization \(|\varphi'| = \text{const}\. The algorithm we want to numerically implement in gMAM was given in equation (100) by:

\[
\begin{cases}
\frac{\varphi - \tilde{\varphi}}{\Delta t} = (\lambda_i)^2 \frac{\varphi_{i+1} - 2\varphi_i + \varphi_{i-1}}{N^2} - \lambda_i H_{\theta x} \varphi_i' + H_{\theta b} H_{x} + \lambda_i \lambda_i' \varphi_i', \\
\tilde{\varphi}_0 = x_1, \quad \tilde{\varphi}_N = x_2
\end{cases}
\] (170)

\(\varphi\) represents the state of the system at current time \(t\) and \(\tilde{\varphi}\) represents the updated state at time \(t + \Delta t\). Remember that in our case, the \(x\) in the above equation corresponds to the vector \((u_1, u_2, u_3)\) such that:

\[
\varphi(\alpha, t) = \begin{pmatrix}
x_1 \\
x_2 \\
x_3
\end{pmatrix} = \begin{pmatrix}
U_1 \\
U_2 \\
U_3
\end{pmatrix}
\] (171)

Remember that in the Maier Stein ODE problem, we found that Equation (170) can be written:

\[
\varphi = \lambda^2 \varphi'' - \lambda (\nabla b - (\nabla b)^T) \varphi' - (\nabla b)^T b + \lambda \lambda' \varphi' + \mu \varphi'
\] (172)
Figure 30: Displays the minimizer produced by the gMAM 3D ODE model for the transition from
the point $U_1 = 1.5 \quad U_2 = 0 \quad U_3 = 0$ to the point $U_1 = 1.5 \quad U_2 = 0 \quad U_3 = 2$ with parameters: $c_1 = 1 \quad c_2 = 1 \quad d_1 = 3 \quad d_2 = 1 \quad b = .85 \quad \omega = 1.5$

However, in this case, our $\sigma$ matrix (and therefore our diffusion tensor $a$) is a function
of $\varphi$. This is referred to as multiplicative noise. This means that the derivatives on the
Hamiltonian $H_{\theta\theta}$, $H_x$, and $H_{\theta x}$ are going to be different than what we got in the Maier-Stein
case. Our Hamiltonian and its derivatives are given by:

$$H = \langle b, \theta \rangle + \frac{1}{2} \langle \theta, a\theta \rangle$$  (173)

$$H_{\theta\theta} = a, \quad H_{\theta x} = b_x + (a\theta)_x$$  (174)

$$H_x = (\nabla b)^T \varphi + \nabla (\varphi^T a\varphi)$$  (175)

In particular for Moore’s model:

$$H_{\theta\theta} = a = \begin{bmatrix}
U_1 & 0 & -U_3 \\
0 & U_1/3 & 0 \\
-U_3 & 0 & \pi^2/12U_1^3 + U_3^2/U_1^3
\end{bmatrix}$$  (176)
George Poppe

Applications of gMAM

Another detail we need to consider before we can proceed is the $\vartheta$. Our algorithm requires that we write $\vartheta$ in terms of $\phi$. For the Maier-Stein, we made the substitution:

$$\vartheta = \lambda \phi' - b$$

However recall that for diffusion processes, the Legendre transformation required that:

$$\vartheta(x,y) = a^{-1}(x) \left( \frac{b(x) |a_y y - b(x)|}{|y|_a} \right)$$

Equation (96) tells us that in we need to find the inverse of the diffusion tensor $a$ if we want to know the value of $\vartheta$. Numerically speaking it may be better to compute the inverse algebraically, rather than inverting it through brute force in each step.

Figure 30 shows the resulting gMAM transition with parameters: $c_1 = 1$, $c_2 = 1$, $d_1 = 3$, $d_2 = 1$, $b = .85$, $\omega = 1.5$. We obtain good agreement with R. Moore’s result [2].

7.5 Comparing the PDE to the ODE reduction (Simple Case: Diffusive NLSE)

The question that remains is whether or to what degree the Moore ODE reduction characterizes or relates to the full PDE solution. In order to understand how to compare PDE’s to ODE’s we should begin by examining a simpler model than the full MLL PDE. We can look at a dissipative NLSE and compare the resulting ODE and PDE transitions.

$$iA_z + \frac{1}{2} A_{tt} + |A|^2 A = -c_1 A + i c_2 A_{tt} + f(z,t)$$

Starting with the PDE, we can construct the gMAM as a special case of the analysis
conducted in section 6.5

\[
b(u, v) = \begin{pmatrix}
-(u^2 + v^2)v - c_1 u \\
+(u^2 + v^2)u - c_1 v
\end{pmatrix}
\]  \hspace{1cm} (179)

The \( \kappa \) matrix is given by:

\[
\kappa = \begin{pmatrix}
c_2 & -\frac{1}{2} \\
\frac{1}{2} & c_2
\end{pmatrix}
\]  \hspace{1cm} (180)

Let us turn now to the ODE reduction: If we follow the methodology presented in the Appendix 9.1, we can reduce this PDE to an ODE with the form of equation (11).

In this case, the 4 element drift vector \( b(u) \) can be written:

\[
\begin{bmatrix}
b_1(u) \\
b_2(u) \\
b_3(u) \\
b_4(u)
\end{bmatrix}
= \begin{bmatrix}
-2c_1 u_1 + -\frac{2}{3}c_2 u_1^3 - 2c_2 u_1 u_2^2 \\
\frac{-4}{3}c_2 u_1^2 u_2 \\
u_2 \\
\frac{1}{2}u_1^2 - \frac{1}{2}u_2^2 + \frac{4}{3}c_2 u_1^2 u_2 u_3
\end{bmatrix}
\]  \hspace{1cm} (181)
Figure 32: Displays comparison of the heat map of $v$ for the PDE gMAM solution (left) against that of the ODE gMAM solution (right) with parameters: $c_1 = .1$, $c_2 = .1$. As in figure 31, we see partial agreement in the $v$ field as well.

The noise matrix $\sigma$ depends only on the noise term in the PDE, $f(z,t)$, and on the primary ansatz of the ODE reduction. We are taking the same ansatz as described in Moore’s ODE reduction:

$$A(t,z) = u_1(z) \text{sech} \left( u_1(z)(t - u_3(z)) \right) e^{i(u_4(z) + u_2(z)t)}$$  \hspace{1cm} (182)

Therefore $\sigma$ has the same form as it did in the case of the MLL PDE:

$$\sigma = \sqrt{2} \begin{bmatrix} \sqrt{u_1} & 0 & 0 & 0 \\ 0 & -\sqrt{\frac{u_1}{3}} & 0 & 0 \\ -\frac{u_3}{\sqrt{u_1}} & 0 & \sqrt{\frac{u_3^2}{u_1} + \frac{\pi^2}{12u_1}} & 0 \\ 0 & u_3\sqrt{\frac{u_1}{3}} & 0 & \sqrt{\frac{\pi^2 + 12}{36u_1}} \end{bmatrix}$$  \hspace{1cm} (183)

Having written the $b$ vector and $\sigma$ matrix, we can now choose a boundary condition. Let us take the case that the system goes from a small amplitude, flat soliton to a large
amplitude soliton (similarly to what we have done in section 6.3). Now we can produce minimizers from the gMAM simulation for both the ODE reduction and the PDE using the same methodologies outlined in earlier chapters. There is one issue that arises when doing this, however: The gMAM algorithm in the ODE case will, in general, parameterize the system differently than the gMAM in the PDE case. This may cause the result to be misleading. To ensure this issue doesn’t affect our interpretation, we can take the solution for the ODE, construct the soliton trajectory from equation (182), and finally, reparametrize the result along the minimizer so that it matches the parameterization of the PDE minimizer. Assuming that the ODE reduction represents an accurate approximation for a particular set of parameters, then when we compare both of these solutions, the ODE waveform should be a similar to the PDE waveform.

\[ |A| \]

Figure 33: Left panel: plot of the amplitude \( A(\alpha) \) of the ODE minimizer (red) in comparison to the amplitude of the PDE minimizer (blue) for the transition of a soliton with \( |A|_1 = 0.5 \) to \( |A|_2 = 2.5 \). Right panel: same comparison but for the exit from zero to the soliton with an amplitude \( |A|_2 = 2.5 \).

Figures 31 and 32 display one such comparison using the parameters: \( c_1 = .1, c_2 = .1 \). The ODE solution appears to capture many features of the solution to the PDE. Figure 33 shows a comparison of the amplitudes along the minimizing path for the ODE and PDE cases. Here, we see strong agreement between these two solutions.

This example demonstrates that for at least simple cases, an ODE reduction can be used
as a reasonable (though not perfect) approximation to a PDE system.

### 7.6 PDE-ODE comparison for the full MLL PDE model

Now that we understand how to go about comparing these two kinds of systems, we can finally attempt to apply the same methodology to the full MLL PDE.

\[
iA_t + \frac{c_3}{2} A_{tt} + c_4 |A|^2 A = -b \cos(\omega t) A - i c_1 A \\
+ ic_2 A_{tt} + i d_1 |A|^2 u - i d_2 |A|^4 A + i \epsilon f(t, z).
\]

The analysis for applying the gMAM to the PDE and ODE reduction has already been carried out in sections 6.5 and 7.4. The one distinction we make from Section 7.4 is that we will use the 4-dimensional drift vector \( b \) and noise matrix \( \sigma \) instead of the corresponding 3-dimensional quantities. The addition of this 4th dimension adds an extra degree of freedom to the simulation, thereby allowing the system to potentially relax into a smaller action than it could otherwise.

Figure 34: Displays heat maps of \( u \) for the ODE gMAM solution (left) vs the PDE gMAM solution (right) with parameters: \( c_1 = .1, c_2 = .1, d_1 = 0.1, d_2 = 0.1, b = 0.1, \omega = 1.0 \). There does appear to be broad correspondence between the two images.
Figure 35: Displays heat maps of $v$ for the ODE gMAM solution (left) vs the PDE gMAM solution (right) with parameters: $c_1 = .1, c_2 = .1, d_1 = 0.1, d_2 = 0.1, b = 0.1, \omega = 1.0$. We see agreement similar to that seen in Figure 34.

A detailed derivation of the drift vector and the noise matrix can be found in the Appendix (Equations (188) and (193)). We wish to analyze the phenomenon of soliton creation. That is, we want to find the transition from a state where there is a very small soliton (amplitude .5) to a state where a large soliton exists (amplitude 2.5).

As described in the previous section, before we can compare the two solutions, we need to reconstruct the soliton from the ODE transition according to equation (160) and then reparametrize according to the PDE reparametrization protocol.

Figures 34 and 35 show a comparison of the resulting pulse transitions. The visual agreement between these two methods supports the validity and merit of the ODE reduction presented in [2] as an approximation on the PDE system.

One interesting difference between the ODE and the PDE solutions is the shape of the phase. In the case of the PDE minimizer, we see that during the transition from a small soliton to a large soliton, the phase has a region where it looks parabolic. There is an open question about whether these solitons should be modeled as having a parabolic phase. More study of the phenomenon needs to be done, however we do see some evidence in Figure 36 to support the case that indeed it should be.
Figure 36: Example plots of phase vs time (in blue) plotted on the same axes with a magnified plot of the amplitude vs time (in red) for $\alpha = .180$ (left) and $s = .624$ (right). We can see that the phase shows parabolic behavior within the time frame where the amplitude is large which could justify a using a parabolic phase model to characterize this system. As the system evolves along the minimizing path ($\alpha$), we see that the parabolic phase model starts to break down as an approximation of the waveform.

8 Conclusion

This thesis represents an exploration into the numerical and analytical techniques for understanding stochastic linear and nonlinear ODEs and PDEs in the context of rare events. Primarily it extends the known functionality of gMAM to previously unsolved problems including the real Ginsburg Landau equation, the linear Schrödinger equation, and various forms of the nonlinear Schrödinger equation, culminating in an application towards a phenomenological model describing a mode-locked laser. Each of these problems carries unique challenges in implementation as well as insights into the behavior of rare events transitions.

Studying these problems also required the simultaneous development of the analytical toolbox associated with these problems. One major tool was the application of instanton methods towards various previously unsolved systems, in particular the linear Schrödinger equation, and the cubic diffusive NLSE. Numerically integrating these in-
stanton solutions showed good agreement with the results derived via gMAM. Finally, and perhaps most importantly, the tools above were used to pioneer study of the relationship between a PDE and its associated ODE reduction. We compare the waveforms of the associated transitions for the ODE and PDE systems. We find that there is moderate agreement between the two waveforms, particularly in the amplitude. To improve these results, the ODE model could be extended. For example, the Moore’s model [2] includes a factor $e^{i\Omega t}$, however there is no $t^2$ term in the exponent (corresponding to the chirp, which can be included in other ODE models [74-76]). This assumption may cause the ODE model to be unable to account for certain aspects of PDE solution. There is much more work to be done on the precise nature of this disagreement, however this thesis takes some of the first steps in understanding it.
9 Appendix

9.1 Mode-Locked Laser SPDE to ODE Reduction

In this section we go through the details of mathematics involved in approximating the MLL SPDE as an ODE that was first done by Moore in [2].

The starting PDE:

\[
\frac{i}{\partial z} \frac{\partial u}{\partial z} + \frac{c_3}{2} \frac{\partial^2 u}{\partial t^2} + c_4 |u|^2 u = -b \cos(\omega t) u - ic_1 u + ic_2 \frac{\partial^2 u}{\partial t^2} + id_1 |u|^2 u - id_2 |u|^4 + i \varepsilon f(t, z)
\]

In order to derive a low dimensional approximation, we begin with the ansatz that the solution is a sech function with a phase:

\[
u = A \text{sech}(A(t - T)) e^{i(\phi + \Omega t)}
\]

Notation:

\[
\begin{align*}
\rho & \equiv A(t - T) \\
\alpha & \equiv (\phi + \Omega t) \\
\kappa & \equiv \text{sech}(\rho) \\
\gamma & \equiv e^{i(\phi + \Omega t)}
\end{align*}
\]

therefore:

\[
u = A \kappa \gamma
\]
9.1.1 Writing the Lagrangian

Write the Lagrangian Density \( L \):

\[
L = \text{Im}(u^*_z u) - \frac{c_3}{2} |u_t|^2 + \frac{c_4}{2} |u|^4 + b \cos(\omega t) |u|^2
\]

Writing the terms explicitly:

\[
u_z = A\kappa' \left[ \frac{A_z}{A} + \tanh(\rho)(T_z A - A_z(t - T)) + i(\phi_z + \Omega_z t) \right]
\]

\[
\text{Im}(u^*_z u) = A^2 \kappa^2 [- (\phi_z + \Omega_z t)]
\]

\[
|u_t|^2 = A^2 \kappa^2 [A^2 \tanh^2(\rho) + \Omega^2]
\]

For Lagrangian density, \( L \), we get:

\[
L = A^2 \kappa^2 \left[ - (\phi_z + \Omega_z t) - \frac{c_3}{2} (A^2 \tanh^2(\rho) + \Omega^2) + \frac{c_4}{2} A^2 \kappa^2 + b \cos(\omega t) \right]
\]

In order to get the Lagrangian \( \mathcal{L} \) from the Lagrangian density \( L \), we need to integrate the spatial part, denoted by variable \( t \):

\[
\mathcal{L} = \int L \, dt
\]

This integral can be evaluated analytically. The result is:

\[
\mathcal{L} = -2A(\phi_z + \Omega_z T + \frac{c_3 \Omega^2}{2}) + (2c_4 - c_3) \frac{A^3}{3} + b \frac{\pi \omega \cos(\omega T)}{\sinh(\zeta)}
\]

with
\[ \zeta \equiv \frac{1}{2} \frac{\pi \omega}{A} \]

The next step is to find the equations of motion associated with the Lagrangian. These can be found using the formula:

\[
\frac{\partial \mathcal{L}}{\partial q} - \frac{d}{dz} \frac{\partial \mathcal{L}}{\partial q_z} = 2Re \left[ i \int \left[ -c_1 u + c_2 u_t + d_1 |u|^2 u - d_2 |u|^4 u + \varepsilon f(x,t) \right] \frac{\partial u^*}{\partial q} \, dt \right] \tag{184}
\]

Where the \( q \) vector has components \((A, \Omega, T, \phi)\). The function \( f(x,t) \) represents white noise. That part of the integral will be dealt with separately and will be denoted as follows:

\[ \Xi_q \equiv 2Re \left[ \int i \varepsilon f(x,t) \frac{\partial u^*}{\partial q} \, dt \right] \]

### 9.1.2 Finding the Equations of Motion

Computing the left hand side of equation (184) gives:

\[
\begin{align*}
\frac{\partial \mathcal{L}}{\partial A} - \frac{d}{dz} \frac{\partial \mathcal{L}}{\partial A_z} &= -2(\phi_z + \Omega_z T + \frac{c_3 \Omega^2}{2}) + (2c_4 - c_3)A^2 + b \frac{\pi^2 \omega^2 \coth(\zeta) \cos(wT)}{2 A^2 \sinh(\zeta)} \\
\frac{\partial \mathcal{L}}{\partial \Omega} - \frac{d}{dz} \frac{\partial \mathcal{L}}{\partial \Omega_z} &= -2c_3 A \Omega + 2(A_z T + AT_z) \\
\frac{\partial \mathcal{L}}{\partial T} - \frac{d}{dz} \frac{\partial \mathcal{L}}{\partial T_z} &= -2A \Omega_z - b \frac{\pi \omega^2 \sin(\omega T)}{\sinh(\zeta)} \\
\frac{\partial \mathcal{L}}{\partial \phi} - \frac{d}{dz} \frac{\partial \mathcal{L}}{\partial \phi_z} &= 2A_z 
\end{align*}
\]

On the right hand side, we have:

\[ u_{tt} = A \kappa \gamma \left[ (-A \tanh(\rho) + i\Omega)^2 - A^2 \right] \]

For \( q = A \):

\[ \frac{\partial u^*}{\partial A} = \kappa \gamma' \left( 1 - \rho \tanh(\rho) \right) \]
Therefore for the RHS of Equation (184) when \( q = A \), we only retain one component of the sum. This is because we are only taking the real part and there is an overall i factor. Thus:

\[
RHS_A = 2Re \left[ \int iA\kappa^2[-2ic_2 A\Omega \tanh(\rho)][1 - \rho \tanh(\rho)] \right] + \Xi_A \\
= 0 + \Xi_A
\]

For \( q = \Omega \):

\[
\frac{\partial u^*}{\partial \Omega} = (-i)A\kappa \gamma^* 
\]

So the RHS becomes:

\[
RHS_\Omega = 2 * Re \left[ i(-i) \int dt t A^2 \kappa^2\left[ -c_1 + c_2 [(A^2 \tanh(\rho))^2 - \Omega^2] - A^2 \text{sech}(\rho)] + d_1 A^2 \kappa^2 - d_2 A^4 \kappa^4 \right] + \Xi_\Omega 
\]

Integrating gives:

\[
RHS_\Omega = \left[ -4c_1 AT - 4Tc_2 \frac{A^3}{3} + \Omega^2 A + \frac{8A^3 T}{3} - \frac{32d_2 A^5 T}{15} \right] + \Xi_\Omega
\]

for \( q = T \)

\[
\frac{\partial u^*}{\partial T} = A\kappa \gamma^* [A \tanh(\rho)] 
\]

Thus:

\[
RHS_T = 2Re \left[ \int dt iA^2 \kappa^2\left[ -2iA\Omega \tanh(\rho) \right] (A \tanh(\rho)) \right] + \Xi_T \\
= \frac{8c_2 A^3 \Omega}{3} + \Xi_T
\]
For \( q = \phi \):

\[
\frac{\partial u^*}{\partial \phi} = (-i)A\kappa \phi^*
\]

So for \(\text{RHS}_{\phi}\) we get:

\[
\text{RHS}_{\phi} = 2\text{Re} \left[ i(-i) \int dt A^2 \kappa^2 [-c_1 + c_2 \left( A^2 \text{tanh}(\rho)^2 - \Omega^2 - A^2 \kappa^2 \right) + d_1 A^2 \kappa^2 - d_2 A^4 \kappa^4] + \Xi_{\phi} \right]
= \left[ -4Ac_1 - c_2 \left( \frac{4A^3}{3} + 4\Omega^2 A \right) + d_1 \frac{8A^3}{3} - d_2 \frac{32A^5}{15} \right] + \Xi_{\phi}
\]

Combining the LHS with the RHS gives us four equations of motion

\[
-2(\phi_z + \Omega_z T + \frac{c_3 \Omega^2}{2}) + (2c_4 - c_3)A^2 + b \frac{\pi^2 \omega^2 \coth(\zeta) \cos(\omega T)}{2A^2 \sinh(\zeta)} = \Xi_A
\]

\[
-2c_3 A \Omega + 2(A_z T + AT_z) = \left[ -4c_1 AT - 4T c_2 \left( \frac{A^3}{3} + \Omega^2 A \right) + \frac{8d_1 A^3 T}{3} - \frac{32d_2 A^5 T}{15} \right] + \Xi_{\Omega}
\]

\[
-2A \Omega_z - b \frac{\pi \omega^2 \sin(\omega T)}{\sinh(\zeta)} = \frac{8c_2 A^3 \Omega}{3} + \Xi_T
\]

\[
2A_z = \left[ -4Ac_1 - c_2 \left( \frac{4A^3}{3} + 4\Omega^2 A \right) + d_1 \frac{8A^3}{3} - d_2 \frac{32A^5}{15} \right] + \Xi_{\phi}
\]

We can solve these equations for the components of the \( q_z \) vector.

\[
A_z = \left[ -2Ac_1 - c_2 \left( \frac{2A^3}{3} + 2\Omega^2 A \right) + d_1 \frac{4A^3}{3} - d_2 \frac{16A^5}{15} \right] + \frac{1}{2} \Xi_{\phi}
\]

\[
\Omega_z = -\frac{4}{3} A^2 c_2 \Omega - b \frac{\pi \omega^2 \sin(\omega T)}{2A \sinh(\zeta)} - \frac{1}{2A} \Xi_T
\]

\[
T_z = c_3 \Omega - \frac{T}{2A} \Xi_{\phi} + \frac{1}{2A} \Xi_{\Omega}
\]
We can also write the evolution of $\phi_z$. It turns out that $\phi_z$ is not independent of the other three variables, so it is not necessary to determine the evolution of the soliton.

$$\phi_z = \frac{1}{2}((2c_4 - c_3)A^2 - c_3\Omega^2) + \frac{4}{3}c_2A^2\Omega T + \frac{\pi b \omega^2}{2A} \left( \frac{T \sin(\omega T)}{\sinh(\xi)} \right) \ldots$$

$$\ldots + \frac{\pi^2 \omega^2 b \cos(\omega T) \coth(\xi)}{4A^2} - \frac{1}{2} \Xi_A + \frac{T}{2A} \Xi_T$$

This equation can also be written

$$\phi_z = \frac{1}{2}((2c_4 - c_3)A^2 - c_3\Omega^2) + \frac{4}{3}c_2A^2\Omega T + \frac{\zeta b \cos(\omega T)}{\sinh \xi} \left[ \omega T \tan(\omega T) + \zeta \coth(\xi) \right] - \frac{1}{2} \Xi_A + \frac{T}{2A} \Xi_T$$

Using these equations, we can construct the four dimensional $b$ vector associated with the SDE (11). Using $U$ for our position variable (as we saw in Section 7.2), we get:

$$b(U) = \begin{bmatrix}
-2c_1U_1 + \left( \frac{4}{3}d_1 - \frac{2}{3}c_2 \right)U_1^3 - \frac{16}{15}d_2U_1^5 - 2c_2U_1U_2^2 \\
-4c_2U_1^2U_2 - \frac{\pi b \omega^2}{20U_1} \text{csch} \left( \frac{\pi \omega}{2U_1} \right) \sin(\omega U_3) \\
c_3U_2 \\
\frac{1}{2}((2c_4 - c_3)U_1^2 - c_3U_2^2) + \frac{4}{3}c_2U_1^2U_2U_3 + \frac{\zeta b \cos(\omega U_3)}{\sinh \xi} \left[ \omega U_3 \tan(\omega U_3) + \zeta \coth(\xi) \right]
\end{bmatrix}$$

(188)

### 9.1.3 Resolving $\Xi$

To resolve what the value of the $\Xi_q$’s are, we note that the function $f(x,t)$ has a real and an imaginary component. Then, in the 4 cases, since we are only taking the real part of the integral, we can write:

$$\Xi_A = 2\epsilon \int dt \kappa [1 - \rho \tanh(\rho)] (f_r \sin(\alpha) - f_i \cos(\alpha))$$

$$\Xi_\Omega = 2A\epsilon \int dt \kappa (f_r \cos(\alpha) + f_i \sin(\alpha))$$
From these equations we are interested in constructing the stochastic differential equation of the form:

\[ dU = F(U)dz + \varepsilon\sigma(U)dW \]

Where the function F corresponds to the drift vector and \( \sigma \) corresponds to standard deviation associated with the noise. In order to solve for \( \sigma \), we can find the expectation value for the square of the \( \Xi \)'s, \( \langle \Xi_q^2 \rangle \).

Because \( f_i(t) \) and \( f_i(t' \neq t) \) correspond to uncorrelated random numbers, if we take the product of these numbers, we are simply left with another random number that has an expectation value of 0. Using this fact to integrate the above 4 equations gives:

\[ \langle \Xi_A^2 \rangle = (2\varepsilon)^2 \int dt \kappa^2 [1 - \rho \tanh(\rho)]^2 \]

\[ \langle \Xi_A^2 \rangle = \frac{1}{9} \frac{\pi^2 + 12}{A} \varepsilon^2 \quad (189) \]

\[ \langle \Xi_A^2 \rangle = (2A\varepsilon)^2 \int dt \kappa^2 t^2 \]

\[ \langle \Xi_A^2 \rangle = \left( \frac{2}{3} \right) \frac{12A^2T^2 + \pi^2}{A} \varepsilon^2 \quad (190) \]

\[ \langle \Xi_T^2 \rangle = (2A^2\varepsilon)^2 \int dt \kappa^2 \tanh(\rho)^2 \]

\[ \langle \Xi_T^2 \rangle = \frac{8}{3} A^3 \varepsilon^2 \quad (191) \]
George Poppe

Applications of gMAM

\[
\langle \Xi^2_{\phi} \rangle = (2\epsilon A) \int dt \, k^2 \\
\langle \Xi^2_{\phi} \rangle = 8\epsilon^2 A \tag{192}
\]

Now that we’ve got the \( \langle \Xi^2_q \rangle \)'s worked out, we can figure out what the standard deviations must be by looking at equations (185), (186), and (187).

\[
\sigma = \begin{bmatrix}
\frac{1}{2} \sqrt{8A} & 0 & 0 \\
0 & -\frac{1}{2\sqrt{A}} \frac{8}{3} A^3 & 0 \\
-T \frac{\sqrt{8A}}{2} & 0 & \sqrt{\frac{12A^2 T^2 + \pi^2}{6A^3}}
\end{bmatrix}
\]

if we pull out a \( \sqrt{2} \)

\[
\sigma = \sqrt{2} \begin{bmatrix}
\sqrt{A} & 0 & 0 \\
0 & -\sqrt{\frac{A}{3}} & 0 \\
-T \frac{\sqrt{A}}{\sqrt{A}} & 0 & \sqrt{\frac{T^2}{A} + \frac{\pi^2}{12A^3}}
\end{bmatrix} \tag{193}
\]

This quantity is used to construct the diffusion tensor \( a \) such that \( a = \sigma \sigma^T \). Written here is the 3D \( \sigma \) matrix as was used in the Richard Moore paper [2]. We could also include the \( \phi \) term in our SDE. If we do so, \( \sigma \) becomes a 4x4 Matrix given by:

\[
\sigma = \sqrt{2} \begin{bmatrix}
\sqrt{A} & 0 & 0 & 0 \\
0 & -\sqrt{\frac{A}{3}} & 0 & 0 \\
-T \frac{\sqrt{A}}{\sqrt{A}} & 0 & \sqrt{\frac{T^2}{A} + \frac{\pi^2}{12A^3}} & 0 \\
0 & T \sqrt{\frac{A}{3}} & 0 & \sqrt{\frac{\pi^2 + 12}{36A}}
\end{bmatrix} \tag{194}
\]
10 References


33. R. Brown, “A brief account of microscopical observations made in the months of June, July and August, 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies,” E. Phil. pp. 358–371 (1828).


50. S. R. S. Varadhan, Large deviations and applications (SIAM, 1984), Vol. 46.


