2009-07-06

Numerical Solutions for Stochastic Differential Equations and Some Examples

Yi Luo
Brigham Young University - Provo

NUMERICAL SOLUTIONS FOR STOCHASTIC DIFFERENTIAL EQUATIONS AND SOME EXAMPLES

by

Yi Luo

A thesis submitted to the faculty of

Brigham Young University

in partial fulfillment of the requirements for the degree of

Master of Science

Department of Mathematics

Brigham Young University

August 2009

BRIGHAM YOUNG UNIVERSITY

GRADUATE COMMITTEE APPROVAL

of a thesis submitted by

Yi Luo

This thesis has been read by each of member of the following graduate committee and by majority vote has been found to be satisfactory.

Date	Kening Lu, Chair
Date	Kenneth L. Kuttler
Date	Tiancheng Ouyang

BRIGHAM YOUNG UNIVERSITY

As chair of the candidate's graduate committee, I have read the thesis of Yi Luo in its final form and have found that (1) its format, citations, and bibliographical style are consistent and acceptable and fulfill university and department style requirements; (2) its illustrative materials including figures, tables, and charts are in place; (3) the final manuscript is satisfactory to the graduate committee and is ready for submission to the university library.

Date

Kening Lu
Chair, Graduate Committee

Accepted for the Department

Tyler J. Jarvis
Dapartment Chair

Accepted for the College

Thomas W. Sederberg, Associate Dean
College of Physical and Mathematical Sciences

Abstract

NUMERICAL SOLUTIONS FOR STOCHASTIC DIFFERENTIAL EQUATIONS AND SOME EXAMPLES

Yi Luo

Department of Mathematics

Master of Science

In this thesis, I will study the qualitative properties of solutions of stochastic differential equations arising in applications by using the numerical methods. It contains two parts. In the first part, I will first review some of the basic theory of the stochastic calculus and the Ito-Taylor expansion for stochastic differential equations (SDEs). Then I will discuss some numerical schemes that come from the Ito-Taylor expansion including their order of convergence. In the second part, I will use some schemes to solve the stochastic Duffing equation, the stochastic Lorenz equation, the stochastic pendulum equation, and the stochastic equations which model the spread options.

Acknowledgments

I would like to express my deep appreciation to my advisor and mentor Kening Lu for the time, guidance, encouragement he has given me, for his patience and for the opportunities he has provided for me to obtain a Master degree at BYU.

I would like to thank my grandparents Zhensheng Luo, and Yunxiang Li for teaching me the value of work, and my parents Yongjiang Luo and Meiju He, for their encouragement, concern and support throughout college and graduate school, I also thank my aunt Jianping Luo and uncle Yafei He for their help to make my learning experience at BYU possible. Lastly, I also thank my wife Sarah, for her support and encouragement for my research and patience while I was busy working.

I would also like to extend my appreciation to my committee: Professor Ouyang Tiancheng, Kenneth Kuttler, and Todd Fisher for their willingness to give of their time and provide counsel. I thank the graduate committee for admitting me and the financial support they provided.

Contents

List of Figures

Chapter 1. Introduction

Stochastic differential equations arise in the modelling of many phenomena in physics, biology, climatology, economics, etc., when uncertainties or random influences (called *noises*), are taken into account. These random effects are not only introduced to compensate for the defects in some deterministic models, but also are often rather intrinsic phenomena. In finance, the Black-Scholes-Merton stochastic equations are used to model the option price. In climatology, the stochastic Lorenz system is used to study the flow of the atmosphere. The stochastic lattice differential equations are used to model systems such as cellular neural networks with applications to image processing, pattern recognition, and brain science.

A way to deal with these random effects is the Ito stochastic analysis based on the Brownian motion. Due to the irregularity of the Brownian motion, one can only interpret the stochastic differential equations in terms of the stochastic integral equations. A major difference from the deterministic differential equations is the chain rule for the "differentials". This is the so-called Ito formula. The numerical approaches I used here is based on the Ito-Taylor expansion for stochastic differential equations, which is much more complicated than the Taylor expansion in the deterministic case.

The earliest work on SDEs was the description on the Brownian motion done in Einstein's paper *On the Motion Required by the Molecular Kinetic Theory of Heat of Small Particles Suspended in a Stationary Liquid*. The Brownian motion or Wiener process was discovered to have exceptionally complex mathematical properties. The Wiener process is nowhere-differentiable, thus we cannot apply the calculus rules or methods to it. Ito calculus, named after Kiyoshi Ito, extends the rules and methods from calculus to stochastic processes such as Wiener process.

In 1945, after Ito received his doctoral degree, he published several papers on stochastic calculus. Among them were *On a Stochastic Integral Equation* (1946), *On the Stochastic Integral* (1948), *Stochastic Differential Equations in a Differentiable Manifold* (1950), and *Brownian Motions in a Lie Group* (1950). Ito gave the famous Ito's lemma in the paper *On Stochastic Differential Equations* (1951). In 1952, Ito published his famous book *Probability Theory*. In this book, he

used terms and tools from measure theory to develop the theory on a probability space. During his stay at Princeton University, he published another famous book, *Stochastic Process*. This book studied processes with independent increments, stationary process, Markov process, and some theory of diffusion process.

Because of the "randomness" of the SDEs and the irregularity of the Brownian motion , it is difficult to solve them analytically. Very few SDEs have explicit solutions. For this reason, "for the past decade there has been an accelerating interest in the development of numerical methods for SDEs"[1]. I shall mention several different approaches that have been proposed for the numerical solutions of SDEs. In 1978, in the paper *Approximate Solution of Random Ordinary Differential Equations*, Boyce proposed to use Monte Carlo methods to find numerical solutions for SDEs. In the paper *Numerical Methods for Stochastic Control Problems in Continuous Time*, Kushner and Dupuis suggested the discretization of both time and space variables. Some higher order approximations by Markov chains are proposed in 1992 by Platen.

In the first part of this thesis, I will review some of the basic theory of the stochastic calculus and the Ito-Taylor expansion for stochastic differential equations. Then I will discuss some numerical schemes that come from the Ito-Taylor expansion including their order of convergence.

In the second part, I will study the qualitative properties of solutions of stochastic differential equations arising in applications by using the numerical methods developed in [1]. I will apply these methods to the stochastic Duffing equation, stochastic Lorenz equations, and the stochastic pendulum equation. You will see plots of solutions with different parameters and strength of random noise. I will also compare the numerical results with the deterministic cases, and comment on the differences. I will also study the SDE's for the spread option pricing model.

Chapter 2. Background on Brownian Motion and Ito Calculus

In this chapter, I will review basic concepts and results on the Brownian motion and Ito stochastic calculus. For the sake completeness, I will also include the outlines of proofs for some results. I refer readers to [2] for the details of the proofs and some of definitions.

2.1 The Brownian Motion

First I will define what is a stochastic process:

Definition 2.1. A stochastic process is a parameterized collection of random variables

$$\{X_t\}_{t\in[0,T]}$$

defined on a probability space $(\Omega, \mathcal{F}, P)$.

Note that for each fixed $t \in [0,T]$ we have a random variable $X_t(\omega)$, and for a fixed $\omega \in \Omega$ we have a function that maps t to $X_t(\omega)$ which is called a path of X_t. We can also regard the process as a function of two variables $(t,\omega) \to X(t,\omega)$. A good example of a stochastic process is the Brownian motion.

Definition 2.2. A real-valued stochastic process $W(t), t \in [0,T]$ is called a Brownian motion if

- the process has independent increments for $0 \le t_0 \le t_1 \le \cdots \le t_n \le T$
- for all $t \ge 0, h > 0$, $W(t+h) - W(t)$ is normally distributed with mean 0 and variance h,
- the function $t \to W(t)$ is continuous a.s.

We also call a Brownian motion a Wiener processes. Here are some pathwise properties (for a fixed ω) of a Brownian motion

1. W_t is continuous in t a.s.

2. For any interval $[a, b] \subset [0, \infty)$, W_t is not monotone.

3. W_t is not differentiable at any point.

2.2 Ito Integral

Suppose we have a model with some random noise

$$\begin{cases} \dfrac{dX_t}{dt} = a(t, X_t) + b(t, X_t) \cdot \xi \\ X(0) = x_0 \end{cases} \tag{2.1}$$

Ito considered the case where the noise term $\xi = \Delta W_t$, and W_t is Brownian motion. Then we have the stochastic differential equation

$$dX_t = a(t, X_t)dt + b(t, X_t)dW_t, \tag{2.2}$$

a stochastic process X_t is a solution of (2.2) if

$$X_t = x_0 + \int_0^t a(s, X_s)ds + \int_0^t b(s, X_s)dW_s \tag{2.3}$$

Note that a Brownian motion is nowhere differentiable, so we need to define the term

$$\int_0^t b(s, X_s)dW_s.$$

Definition 2.3. Suppose $0 \leq S \leq T$, let $\mathcal{D} = \mathcal{D}(S, T)$ be the class of functions that

$$f(t, \omega) : \ [0, \infty] \times \Omega \to \mathbb{R}^n, \tag{2.4}$$

satisfy

1. The function $(t, \omega) \to f(t, \omega)$ is $\mathcal{B} \times \mathcal{F}$ measurable, where $\mathcal{B}$ is the Borel algebra.

2. f is adapted to $\mathcal{F}_t$.

3. $E\left[\int_S^T f(t, \omega)^2 dt\right] < \infty$.

A function $\phi \in \mathcal{D}$ is called an elementary function if

$$\phi(t,\omega) = \sum_{j=1}^{k-1} e_j(\omega)\chi_{[t_j,t_{j-1})}(t) \tag{2.5}$$

where $e_j(\omega)$ is $\mathcal{F}_{t_j}$ measurable, and $\chi_{[t_j,t_{j-1})}(t)$ is the indicator function, (in the measure theory χ is called the characteristic function). Then the **Ito integral of elementary functions** is defined as following:

Definition 2.4. Let $\{S = t_1 < t_1 < \cdots < t_k = T\}$ is a partition of the interval $[S,T]$, then

$$\int_S^T \phi(s,\omega)dW_t(\omega) = \sum_{j=1}^{k-1} e_j(\omega)(W_{t_{j+1}}(\omega) - W_{t_j}(\omega)) \tag{2.6}$$

Now I introduce a important result for elementary functions:

Lemma 2.5. ***(The Ito isometry for elementary functions)*** *If $\phi(t,\omega)$ is bounded and elementary then*

$$E\left[\left(\int_S^T \phi(t,\omega)dW_t(\omega)\right)^2\right] = E\left[\int_S^T \phi^2(t,\omega)dt\right] \tag{2.7}$$

Proof. Let $\Delta W_j = W_{t_{j+1}} - W_{t_j}$, and note that

$$E(e_i e_j \Delta W_j \Delta W_i) = \begin{cases} 0 & i \neq j \\ E(e_i^2)(t_{j+1} - t_j) & i = j \end{cases} \tag{2.8}$$

thus we have

$$\begin{aligned} E\left[\left(\int_S^T \phi(t,\omega)dW_t(\omega)\right)^2\right] &= E\left(\sum_{j\geq 0} e_j \Delta W_j\right)^2 \\ &= \sum_{i,j} E[e_j e_i \Delta W_i \Delta W_j] \\ &= \sum_j E[e_j^2](t_{j+1} - t_j) \\ &= E\left[\int_S^T \phi^2 dt\right] \end{aligned} \tag{2.9}$$

□

Now we can extend this result from elementary functions to all functions in $\mathcal{D}$. It can be done in the following three steps, for more details see [2]:

Step 1. Let $g \in \mathcal{D}$ be bounded and continuous in t. Define

$$\phi_n = \sum_j g(t_j, \omega)\chi_{[t_j, t_{j+1})}(t) \tag{2.10}$$

Then ϕ_n are elementary functions since $g \in \mathcal{D}$, and is also bounded. Because g is continuous in t, then it is uniformly continuous in $[S, T]$, thus for a given ϵ, there exists δ such that when $|t_j - t_i| \leq \delta$ we have

$$|g(t_i, \omega) - g(t_j, \omega)| \leq \epsilon. \tag{2.11}$$

Now we make the interval $[t_j, t_{j+1})$ have length less than δ. Then for all $t \in [t_j, t_{j+1})$ we have

$$\int_S^T (g - \phi_n)^2 dt = \sum_k \int_{t_j}^{t_{j+1}} (g - \phi_n)^2 dt = \epsilon^2 (T - S). \tag{2.12}$$

This implies that for each ω we have $\int_S^T (g - \phi_n)^2 dt \to 0$ as $n \to \infty$, thus by the bounded convergence theorem we have

$$E\left[\int_S^T (g - \phi_n)^2 dt\right] \to 0, \quad \text{as } n \to \infty$$

Step 2. Let $h \in \mathcal{D}$ be bounded, $|h| \leq M$. Let ψ_n be a continuous function in t, such that

(1) $\psi_n(t) \geq 0$

(2) $\psi_n(t) = 0 \quad \text{for } t \leq -\frac{1}{n}, \text{ or } t \geq 0$

(3) $\int_{-\infty}^{\infty} \psi_n(t) dt = 1.$

Now define

$$g_n(t, \omega) = \int_0^t \psi_n(s - t) h(s, \omega) ds \tag{2.13}$$

then g_n is continuous in t and bounded. With the same reasoning as in step 1 we have

$\int_S^T (h - g_n)^2 dt \to 0$ as $n \to \infty$, thus by the bounded convergence theorem we have

$$E\left[\int_S^T (h - g_n)^2 dt\right] \to 0, \quad \text{as } n \to \infty$$

Step 3. Let $f \in \mathcal{D}$, then define

$$h_n(t,\omega) = \begin{cases} -n & f(t,\omega) < -n \\ f(t,\omega) & -n \le f(t,\omega) \le n \\ n & f(t,\omega) > n \end{cases} \tag{2.14}$$

Then $h_n \in \mathcal{V}$, and is bounded by $|f|$, then by the Dominated Convergence Theorem we have

$$E\left[\int_S^T (f - h_n)^2 dt\right] \to 0 \quad \text{as } n \to \infty \tag{2.15}$$

Now we are ready for the definition of the Ito integral.

Definition 2.6. (**The Ito integral**) Let $f \in \mathcal{D}(S,T)$, Then the Ito integral of f is defined by

$$\int_S^T f(t,\omega) dW_t(\omega) = \lim_{n\to\infty} \int_S^T \phi_n(t,\omega) dW_t(\omega) \tag{2.16}$$

where $\{\phi_n\}$ is a sequence of elementary functions such that

$$E\left[\int_S^T (f - \phi)^2 dt\right] \to 0 \quad \text{as. } n \to \infty. \tag{2.17}$$

We can choose such a sequence by steps 1-3. We can also have the **Ito isometry** for all functions in $\mathcal{D}$ as:

Theorem 2.7. ***(The Ito isometry)*** *For each $f \in \mathcal{V}$ we have*

$$E\left[\left(\int_S^T f(t,\omega) dW_t(\omega)\right)^2\right] = E\left[\int_S^T f^2(t,\omega) dt\right] \tag{2.18}$$

Let us look at an example of an Ito integral

Example 2.8. Assume $W_0 = 0$, show

$$\int_0^t W_s dW_s = \frac{1}{2}W_t^2 - \frac{1}{2}t \tag{2.19}$$

Proof. In order to use the Ito integral, we need to find a sequence of elementary functions that converge to W_t in $\mathcal{L}^2$. Let

$$\phi_n = \sum W_j \chi_{[t_j, t_{j+1})}, \tag{2.20}$$

where $W_j = W_{t_j}$, Then

$$\begin{aligned} E\left[\int_0^t (\phi_n - W_s)^2 ds\right] &= E\left[\sum_j \int_{t_j}^{t_{j+1}} (W_j - W_s)^2 ds\right] \\ &= \sum_j \int_{t_j}^{t_{j+1}} E\left[(W_j - W_s)^2 ds\right] \\ &= \sum_j \int_{t_j}^{t_{j+1}} (s - t_j) ds \\ &= \sum_j \frac{1}{2}(t_{j+1} - t_j)^2 \\ &\leq \frac{\Delta t_n}{2} \sum_j (t_{j+1} - t_j) \\ &= \frac{t \Delta t_n}{2} \to 0 \quad \text{as } n \to \infty \end{aligned} \tag{2.21}$$

where $\Delta t_n = \max_j (t_{j+1} - t_j)$, thus

$$\int_0^t W_s dW_s = \lim_{n \to \infty} \int_0^t \phi_n(s, \omega) dW_s. \tag{2.22}$$

Since $W_0 = 0$, we have

$$\begin{aligned} W_t^2 &= \sum_j \Delta W_j^2 \\ &= \sum_j (W_{j+1}^2 - W_j^2) \\ &= \sum_j [(W_{j+1} - W_j)^2 + 2W_j(W_{j+1} - W_j)] \\ &= \sum_j (\Delta W_j)^2 + 2W_j \Delta W_j \\ &= \sum_j (\Delta W_j)^2 + 2\sum_j W_j \Delta W_j \end{aligned} \tag{2.23}$$

therefore

$$\int_0^t \phi dW_s = \sum_j W_j \Delta W_j = \frac{1}{2}W_t^2 - \frac{1}{2}\sum_j (\Delta W_j)^2 \to \frac{1}{2}W_t^2 - \frac{1}{2}t \tag{2.24}$$

as $n \to \infty$. □

2.3 Ito Formula

From example (2.8), we can see that the basic definition of the Ito integral is not very handy when evaluating a given stochastic integral. However thanks to Ito, we have a formula called the Ito formula which makes evaluating stochastic integrals easier.

Theorem 2.9. ***(The 1-dimensional Ito formula)*** *Let X_t be a stochastic process given by*

$$dX_t = u(t, X_t)dt + v(t, X_t)dW_t \tag{2.25}$$

Let $g(t, x) \in C^2([0, \infty] \times \mathbb{R})$, then

$$Y_t = g(t, X_t) \tag{2.26}$$

is a stochastic process and

$$dY_t = \frac{\partial g}{\partial t}(t, X_t)dt + \frac{\partial g}{\partial x}(t, X_t)dX_t + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t, X_t) \cdot (dX_t)^2 \tag{2.27}$$

where $(dX_t)^2 = (dX_t) \cdot (dX_t)$ *is computed according to the rules*

$$dt \cdot dt = dt \cdot dW_t = dW_t \cdot dt = 0, \quad dW_t \cdot dW_t = dt \tag{2.28}$$

Please read [2] for a sketch of the proof of the Ito formula.

Now let us redo example (2.8) using the Ito formula. Choose $X_t = W_t$, and $g(t,x) = \frac{1}{2}x^2$. Then $Y_t = \frac{1}{2}W_t^2$, and by the Ito formula

$$d(\frac{1}{2}X_t^2) = dY_t = W_t dW_t + \frac{1}{2}(dW_t)^2 = W_t dW_t + \frac{1}{2}W_t. \tag{2.29}$$

Thus

$$\frac{1}{2}W_t^2 = \int_0^t W_s dW_s + \frac{1}{2}t. \tag{2.30}$$

The Ito formula also gives:

Theorem 2.10. ***(Integration by parts)*** *Suppose* $f(t,\omega)$ *is continuous and of bounded variation with respect to* $s \in [0,t]$ *for a.a.*ω*. Then*

$$\int_0^t f(s)dW_s = f(t)W_t - \int_0^t W_s df_s. \tag{2.31}$$

We also have the Ito product formula

Theorem 2.11. ***(Ito product formula)*** *If* X_t, Y_t *are Ito processes, then*

$$d(X_t Y_t) = X_t dY_t + Y_t dX_t + dX_t dY_t \tag{2.32}$$

Proof. Let $g(x,y) = x \cdot y$, then $X_t Y_t = g(X_t, Y_t)$, by the Ito formula we have

$$\begin{aligned} d(X_t Y_t) &= \frac{\partial g}{\partial X_t}dX_t + \frac{\partial g}{\partial Y_t}dY_t + \frac{1}{2}\frac{\partial^2 g}{\partial X_t^2}(dX_t)^2 + \frac{1}{2}\frac{\partial^2 g}{\partial Y_t^2}(dY_t)^2 + \frac{\partial^2 g}{\partial X_t \partial Y_t}dX_t dY_t \\ &= Y_t dX_t + X_t dY_t + 0 + 0 + dX_t dY_t. \end{aligned} \tag{2.33}$$

Therefore

$$X_t dY_t = X_t Y_t + Y_t dX_t + dX_t dY_t$$

□

I close this chapter with the multi-dimensional Ito formula

Theorem 2.12. ***(Multi-dimensional Ito formula)*** *Let*

$$dX_t = u(t, X_t)dt + v(t, X_t)dW_t \tag{2.34}$$

be an n-dimensional Ito process as above. Let $g(t, x) = (g_1, \cdots, g_n)$ *be a* C^2 *map from* $[0, \infty) \times \mathbb{R}^n$ *in to* $\mathbb{R}^m$*, Then* $Y(t, \omega) = g(t, X_t)$ *is an stochastic process and for* $k = 1, \cdots, n$

$$dY_k = \frac{\partial g_k}{\partial t}(t, X_t)dt + \frac{\partial g_k}{\partial x_i}(t, X_t)dX_i + \frac{1}{2}\sum_{i,j}\frac{\partial^2 g_k}{\partial x_j \partial x_j}(t, X_t) \cdot dX_i dX_j \tag{2.35}$$

Chapter 3. Ito-Taylor Expansion and Numerical Schemes

In this chapter, I introduce the Ito-Taylor expansion and some numerical schemes derived from it, which are taken from [1].

3.1 Some Notations

3.1.1 Introduction. Suppose we have the stochastic process

$$dX_t = a(t, X_t)dt + b(t, X_t)dW_t. \tag{3.1}$$

We want to find its Ito-Taylor expansion center at the initial time t_0. To to this we need

- Iterated application of Ito's formula.
- Some notations: multiple Ito integrals and coefficient functions

The integral form of (3.1) is

$$X_t = X_0 + \int_{t_0}^{t} a(s, X_s)ds + \int_{t_0}^{t} b(s, X_s)dW_s. \tag{3.2}$$

Applying the Ito formula to $a(t, X_t)$ we have

$$da(t, X_t) = (a_t + aa_{X_t} + \frac{1}{2}a_{X_tX_t}b(t, X_t)^2)dt + a_{X_t}b(t, X_t)dW_t. \tag{3.3}$$

We now integrate both sides and get

$$a(t, X_t) = a(t_0, X_0) + \int_{t_0}^{t} a_s + aa_{X_s} + \frac{1}{2}a_{X_sX_s}b(s, X_s)^2ds + \int_{t_0}^{t} a_{X_s}b(s, X_s)dW_s. \tag{3.4}$$

Doing the same thing for $b(t, X_t)$ we have

$$b(t, X_t) = b(t_0, X_0) + \int_{t_0}^{t} b_s + b_{X_s}a(s, X_s) + \frac{1}{2}b(s, X_s)^2b_{X_sX_s}ds + \int_{t_0}^{t} b_{X_s}b(s, X_s)dW_s. \tag{3.5}$$

Now plug (3.4) and (3.5) into (3.2) to get the linear term of the Ito-Taylor expansion

$$X_t = X_0 + a(t_0, X_0)(t - t_0) + b(t_0, X_0)(W_t - W_{t_0}) + R \tag{3.6}$$

where R is the term with integrals. To simplify, [1] introduces the multiple Ito integral, coefficient functions, hierarchical and remainder sets. I will give the a brief introduction on these concepts, please read [1] for more details.

3.1.2 Multiple Ito Integral.

Definition 3.1. Suppose k be a nonnegative integer, for $m = 1, 2, 3, \ldots$, the row vector

$$\alpha = (j_1, j_2, \cdots, j_k) \tag{3.7}$$

is called a multi-index of length k, where

$$j_i \in \{0, 1, \cdots, m\}, \quad i = 1, 2, \cdots, k. \tag{3.8}$$

Here m denote the dimension of the Brownian motion.

Definition 3.2. Let $l(\cdot)$ be the length function for a multi-index α, and $n(\cdot)$ be the number of components of a multi-index α which are equals 0.

For example $\alpha = (2, 4, 0, 2)$, then $l(\alpha) = 4,\ n(\alpha) = 1$.

For completeness we denote by $\hbar$ the multi-index such that $l(\hbar) = 0$, and $\mathcal{M}$ denotes the set of all multi-indices.

Given $\alpha \in \mathcal{M}$ with $l(\alpha) \geq 1$, we write $-\alpha$ and $\alpha-$ for the multi-index in $\mathcal{M}$ obtained by deleting the first and last component, respectively, of α. Thus

$$-(1, 5, 4, 9) = (5, 4, 9) \quad (1, 5, 4, 9)- = (1, 5, 4). \tag{3.9}$$

Finally, we define the *concatenation operator* $*$ for any two multi-indices $\alpha = (j_1, \cdots, j_n)$, and $\beta = (b_1, \cdots, b_k)$ by

$$\alpha * \beta = (j_1, \cdots, j_n, b_1, \cdots, b_k) \tag{3.10}$$

Now we define three sets of adapted, right continuous stochastic process f with left hand limits.

Definition 3.3. 1. $\mathcal{H}_v$ is the set of all adapted, bounded, right continuous stochastic processes f with left hand limits, i.e. if $f \in \mathcal{H}_v$, then $|f(t,\omega)| \leq \infty$.

2. $\mathcal{H}_{(0)}$ contains all right continuous stochastic processes f with left hand limits with

$$\int_0^t |f(s,\omega)|ds \leq \infty. \tag{3.11}$$

3. $\mathcal{H}_{(1)}$ is the set of all right continuous stochastic processes f with left hand limits with

$$\int_0^t |f(s,\omega)|^2 ds \leq \infty. \tag{3.12}$$

In addition we write

$$\mathcal{H}_{(j)} = \mathcal{H}_{(1)} \tag{3.13}$$

for each $j \geq 2$.

Definition 3.4. (Multiple Ito Integral) Let ρ, τ be two stopping times with $0 \leq \rho(\omega) \leq \tau(\omega) \leq T$. Then for a multi-index $\alpha = (j_1, \cdots, j_k) \in \mathcal{M}$ and a process $f \in \mathcal{H}_\alpha$ we define the multiple Ito integral recursively by

$$I_\alpha[f(\cdot)]_{\rho,\tau} := \begin{cases} f(\tau) & l(\alpha) = 0 \\ \int_\rho^\tau I_{\alpha-}[f(\cdot)]_{\rho,s} ds & l(\alpha) = 1 \text{ and } j_k = 0 \\ \int_\rho^\tau I_{\alpha-}[f(\cdot)]_{\rho,s} dW_s^{j_k} & l(\alpha) = 1 \text{ and } j_k \geq 1 \end{cases} \tag{3.14}$$

where $\mathcal{H}_\alpha$ is the set of adapted, right continuous processes f with left hand limits such that the integral process $\{I_{\alpha-}[f(\cdot)]_{\rho,t}, t \geq 0\}$ considered as a function of t satisfies

$$I_{\alpha-}[f(\cdot)]_{\rho,\cdot} \in \mathcal{H}_{(j_k)} \tag{3.15}$$

Let's look at some examples:

$$
\begin{aligned}
&I_\hbar[f(\cdot)]_{0,t} = f(t) \\
&I_{(0)}[f(\cdot)]_{t_i,t_{i+1}} = \int_{t_i}^{t_{i+1}} I_\hbar[f(\cdot)]_{t_i,t_{i+1}} ds = \int_{t_i}^{t_{i+1}} f(s) ds \\
&I_{(1)}[f(\cdot)]_{t_1,t_2} = \int_{t_1}^{t_2} I_\hbar[f(\cdot)]_{t_1,t_2} dW_s^1 = \int_{t_1}^{t_2} f(s) dW_s^1 \\
&I_{(0,1)}[f(\cdot)]_{0,t} = \int_0^t I_{(0)}[f(\cdot)]_{0,s_2} dW_{s_1}^1 = \int_0^t \int_0^{s_1} f(s_2) ds_2 dW_{s_1}^1 \\
&I_{(0,1,2)}[f(\cdot)]_{0,t} = \int_0^t \int_0^{s_2} \int_0^{s_3} f(s_1) ds_1 dW_{s_2}^1 dW_{s_3}^2
\end{aligned}
\tag{3.16}
$$

3.1.3 Coefficient Functions. First define two operators:

$$
L^0 = \frac{\partial}{\partial t} + \sum_{k=1}^{d} a_k \frac{\partial}{\partial x_k} + \frac{1}{2} \sum_{k,l=1}^{d} \sum_{j=1}^{m} b_{k,j} b_{l,j} \frac{\partial^2}{\partial x_k \partial x_l} \tag{3.17}
$$

and for $j = 1, \ldots, m$

$$
L^j = \sum_{k=1}^{d} b_{k,j} \frac{\partial}{\partial x_k}. \tag{3.18}
$$

For each α and function f we defined recursively the **Ito coefficient function**

$$
f_\alpha = \begin{cases} f & l(\alpha) = 0 \\ L^{j_1} f_{-\alpha} & l(\alpha) \geq 1 \end{cases} \tag{3.19}
$$

For example, in the 1-dimensional case $d = m = 1$ for $f(t, x) = x$ we have

$$
\begin{aligned}
&f_\hbar = X_t, \quad f_{(0)} = L^0 X_t = a(t, X_t), f_{(1)} = b(t, X_t) \\
&f_{(1,1)} = L^1 f_{(1)} = b b_{X_t}, \quad f_{(0,1)} = L^0 f_{(1)} = b_t + a b_{X_t} \frac{1}{2} b^2 b_{X_t X_t}
\end{aligned}
\tag{3.20}
$$

3.1.4 Hierarchical and remainder sets. This concept will tell us up to what degree we are going to expand the Ito-Taylor expansion.

Definition 3.5. We call a subset $\mathcal{A} \subset \mathcal{M}$ an hierarchical set if

- $\mathcal{A} \neq \emptyset$

- $\sup_{\alpha \in \mathcal{A}} l(\alpha) < \infty$
- $-\alpha \in \mathcal{A} \quad \text{for each} \quad \alpha \in \mathcal{A} \backslash \{\hbar\}$

For example, the sets

$$\{\hbar\}, \quad \{\hbar, (0), (1)\}, \quad \{\hbar, (0), (1), (1,1)\} \tag{3.21}$$

are hierarchical sets.

Definition 3.6. Define a set $\mathcal{B}(\mathcal{A})$ to be a remained set of $\mathcal{A}$ if

$$\mathcal{B}(\mathcal{A}) = \{\alpha \in \backslash \mathcal{A} : -\alpha \in \mathcal{A}\}. \tag{3.22}$$

Basically, $\mathcal{B}(\mathcal{A})$ contains all the multi-indices that are not in $\mathcal{A}$.

3.2 Ito-Taylor Expansions

Now here is the Ito-Taylor expansion for a d-dimensional Ito process

$$\begin{cases} X_t = a(t, X_t)dt + b(t, X_t)dW_t \\ X_{t_0} = X_0 \end{cases} \tag{3.23}$$

where $t \in [t_0, T]$.

Theorem 3.7. *Let ρ and τ be two stopping times with $t_0 \leq \rho \leq \tau \leq T$. Let $\mathcal{A} \subset \mathcal{M}$ be an hierarchical set; and let $f : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}$. Then the Ito-Taylor expansion*

$$f(\tau, X_\tau) = \sum_{\alpha \in \mathcal{A}} I_\alpha[f_\alpha(\rho, X_\rho)]_{\rho,\tau} + \sum_{\alpha \in \mathcal{B}(\mathcal{A})} I_\alpha[f_\alpha(\cdot, X_\cdot)]_{\rho,\tau}, \tag{3.24}$$

holds, provide all of the derivative of f, a and b and all of the multiple Ito integrals exist.

Let us look at an example. In the case $d = m = 1$ for $f(t, x) = x,\ \rho = t_0,\ \tau = t \in [t_0, T]$ and

the hierarchical set

$$\mathcal{A} = \{\alpha \in \mathcal{M} : l(\alpha) \leq 2\} = \{\hbar, (0), (1), (0,0), (0,1), (1,0), (1,1)\} \tag{3.25}$$

then we have following coefficient functions

$$\begin{aligned}
&f_{\hbar}(t_0, X_0) = X_0 \\
&f_{(0)}(t_0, X_0) = a(t_0, X_0) \\
&f_{(1)}(t_0, X_0) = b(t_0, X_0) \\
&f_{(1,0)} = b(t_0, X_0) a_{X_t}(t_0, X_0) \\
&f_{(1,1)} = b(t_0, X_0) b_{X_t}(t_0, X_0) \\
&f_{(0,0)}(t_0, X_0) = a_t(t_0, X_0) + a(t_0, X_0) a_{X_t}(t_0, X_0) + \frac{1}{2} b(t_0, X_0)^2 a_{X_t X_t}(t_0, X_0) \\
&f_{(0,1)}(t_0, X_0) = b_t(t_0, X_0) + a(t_0, X_0) b_{X_t}(t_0, X_0) + \frac{1}{2} b(t_0, X_0)^2 b_{X_t X_t}(t_0, X_0)
\end{aligned} \tag{3.26}$$

For simplification, we shall let $a = a(t_0, X_0)$ and $b = b(t_0, X_0)$. Then we have the Ito-Taylor expansion,

$$\begin{aligned}
X_t = X_0 + a \int_{t_0}^{t} ds + b \int_{t_0}^{t} dW_s + b b_{X_t} I_{(1,1)} + b a_{X_t} I_{(1,0)} + (a_t + a a_{X_t} + \frac{1}{2} b^2 a_{X_t X_t}) I_{(0,0)} \\
+ (b_t + a b_{X_t} + \frac{1}{2} b^2 b_{X_t X_t}) I_{(0,1)} + R
\end{aligned} \tag{3.27}$$

where

$$\begin{aligned}
I_{(1,1)} &= \int_{t_0}^{t} \int_{t_0}^{s} dW_r dW_s = \frac{1}{2}((W_t - W_{t_0})^2 - (t - t_0)) \\
I_{(0,0)} &= \int_{t_0}^{t} \int_{t_0}^{s} dr ds = \frac{(t_0 - t)^2}{2}.
\end{aligned} \tag{3.28}$$

For $I_{(0,1)}, I_{(1,0)}$ we do not have a form without an integral, thus

$$\begin{aligned}
X_t = X_0 + a \cdot (t - t_0) + b \cdot (t - t_0) + \frac{1}{2} b b_{X_t} ((W_t - W_{t_0})^2 - (t - t_0)) \\
+ (a_t + a a_{X_t} + \frac{1}{2} b^2 a_{X_t X_t}) \frac{(t_0 - t)^2}{2} + b a_{X_t} I_{(1,0)} \\
+ (b_t + a b_{X_t} \frac{1}{2} b^2 b_{X_t X_t}) I_{(0,1)} + R
\end{aligned} \tag{3.29}$$

For the case of $l(\alpha) \leq 3$, we add following terms

$$\begin{aligned} f_{(0,0,0)} =& a_{tt} + a_t a_{X_t} + a a_{X_t} a_t + b b_t a_{X_t X_t} + \frac{1}{2} b^2 a_{X_t X_t t} \\ &+ a \left(a_{t X_t} + a_{X_t}^2 + a a_{X_t X_t} + b b_{X_t} a_{X_t X_t} + \frac{1}{2} b^2 \frac{\partial^3 a}{\partial X_t^3} \right) \\ &+ \frac{1}{2} b^2 \left(a_{t X_t X_t} + 3 a_{X_t} a_{X_t X_t} + a \frac{\partial^3 a}{\partial X_t^3} + b_{X_t}^2 a_{X_t X_t} + b b_{X_t X_t} \frac{\partial^3 a}{\partial X_t^3} + b b_{X_t} \frac{\partial^3 a}{\partial X_t^3} + \frac{1}{2} b^2 \frac{\partial^4 a}{\partial X_t^4} \right) \end{aligned} \tag{3.30}$$

$$\begin{aligned} f_{(0,0,1)} =& b_{tt} + 2 a_t b_{X_t} + b b_t b_{X_t X_t} + \frac{1}{2} b^2 b_{X_t X_t t} \\ &+ a \left(b_{t X_t} + a_{X_t} b_{X_t} + a b_{X_t X_t} + b b_{X_t} b_{X_t X_t} \frac{1}{2} b^2 \frac{\partial^3 b}{\partial X_t^3} \right) \\ &+ \frac{1}{2} b^2 \left(b_{t X_t X_t} a_{X_t X_t} b_{X_t} + 2 a_{X_t} b_{X_t X_t} + a \frac{\partial^3 b}{\partial X_t^3} + b_{X_t}^2 \frac{\partial^3 b}{\partial X_t^3} + b b_{X_t X_t}^2 + 2 b b_{X_t} \frac{\partial^3 b}{\partial X_t^3} + \frac{1}{2} b^2 b \frac{\partial^4 b}{\partial X_t^4} \right) \end{aligned} \tag{3.31}$$

$$\begin{aligned} f_{(0,1,0)} &= b_t a_{X_t} b a_{X_t X_t} + a(b_{X_t} a_{X_t} + b \frac{\partial^2 a}{\partial X_t^2}) + \frac{1}{2} b^2 (a_{X_t} \frac{\partial^2 b}{\partial X_t^2} + 2 b_{X_t} \frac{\partial^2 a}{\partial X_t^2} + b \frac{\partial^3 a}{\partial X_t^3}) \\ f_{(0,1,1)} &= b_t b_{X_t} + b b_{X_t} + a(b_{X_t}^2 + b \frac{\partial^2 b}{\partial X_t^2}) + \frac{1}{2} b^2 (3 b_{X_t} \frac{\partial^2 b}{\partial X_t^2} + b \frac{\partial^3 b}{\partial X_t^3}) \\ f_{(1,0,0)} &= b(a_{X_t} + a_{X_t}^2 + a \frac{\partial^2 a}{\partial X_t^2} + b b_{X_t} \frac{\partial^2 b}{\partial X_t^2} + \frac{1}{2} b^2 \frac{\partial^3 a}{\partial X_t^3}) \\ f_{(1,0,1)} &= b(b_{t X_t} a_{X_t} b_{X_t} + a \frac{\partial^2 b}{\partial X_t^2} + b b_{X_t} \frac{\partial^2 b}{\partial X_t^2} + \frac{1}{2} b^2 \frac{\partial^3 b}{\partial X_t^3}) \\ f_{(1,1,0)} &= b(b_{X_t} a_{X_t} + b \frac{\partial^2 a}{\partial X_t^2}) \\ f_{(1,1,1)} &= b(b_{X_t} \frac{\partial^2 b}{\partial X_t^2} + b \frac{\partial^2 b}{\partial X_t^2}) \end{aligned} \tag{3.32}$$

as the coefficients of corresponding Ito integrals and we have

$$\begin{aligned} I_{(0,0,0)} &= \int_{t_0}^{t} \int_{t_0}^{s} \int_{t_0}^{r} dz dr ds = \frac{(t_0 - t)^3}{3!} \\ I_{(1,1,1)} &= \int_{t_0}^{t} \int_{t_0}^{s} \int_{t_0}^{r} dW_z dW_r dW_s = \frac{(W_t - W_{t_0})^3}{6} - \frac{(W_t - W_{t_0})(t - t_0)}{2}. \end{aligned} \tag{3.33}$$

For the rest of multiply index integral we do not have a form without integral.

For the proof of the Ito-Taylor expansion, I refer the readers to [1]. Now let's look at some numerical schemes that are derived from the Ito-Taylor expansion.

3.3 Strong Taylor Approximations

We shall look at some numerical schemes derived from the Ito-Taylor expansion. I omit the details of the calculations and proof. For the details, please read [1].

3.3.1 The Euler Approximation. The Euler approximation is one of the simplest time discrete approximations of an Ito process. Let $\{X_t\}$ be an Ito process on $t \in [t_0, T]$ satisfying the stochastic differential equation

$$\begin{cases} X_t = a(t, X_t)dt + b(t, X_t)dW_t \\ X_{t_0} = X_0 \end{cases} \tag{3.34}$$

For a given discretization $t_0 = \tau_0 < \tau_1 < \cdots < \tau_N = T$, an Euler approximation is a continuous time stochastic process $\{Y_t\}$ satisfying the iterative scheme

$$Y_{n+1} = Y_n + a(\tau_n, Y_n)(\tau_{n+1} - \tau_n) + b(\tau_n, Y_n)(W_{\tau_{n+1}} - W_\tau) \tag{3.35}$$

with initial value

$$Y_0 = X_0 \tag{3.36}$$

We write

$$\Delta_n = \tau_{n+1} - \tau_n \tag{3.37}$$

for the nth time increment and call

$$\delta = \max_n \Delta_n \tag{3.38}$$

the *maximum* time step. We often use a constant time increment $\delta = (T - t_0)/N$ for some integer N large enough such that $\delta \in (0, 1)$.

In Matlab we can use a built-in pseudo-random number generator *randn*() to generate the Brownian motion.

3.3.2 Absolute Error. Usually we don't know know the solution of a stochastic differential equation explicitly, which is why we do simulation. But if we do happen know the solution explicitly,

we can use the absolute error criterion to calculate the error. It is simply the expectation of the absolute value of the difference between the approximation and the Ito process at time T,

$$\epsilon = E(|X_T - Y_T|) \tag{3.39}$$

We repeat N different simulations of sample paths of the Ito process and their Euler approximation corresponding to the same sample paths of the Brownian motion and estimate the absolute error by

$$\tilde{\epsilon} = \frac{1}{N}\sum_{k=1}^{N} |X_{T,k} - Y_{T,k}| \tag{3.40}$$

3.3.3 Strong Convergence.

Definition 3.8. We say that a time-discrete approximation Y^δ with maximun step size δ *converges strongly* to X at time T if

$$\lim_{\delta \to 0} E(|X_T - Y_T^\delta|) = 0, \tag{3.41}$$

and we shall say that Y^δ *converges strongly with order* $\gamma > 0$ at time T if there exists a positive constant C, which does not depend on δ, and a $\delta_0 > 0$ such that

$$\epsilon(\delta) = E(|X_T - Y_T^\delta|) \leq C\delta^\gamma \tag{3.42}$$

for each $\delta \in (0, \delta_0)$.

3.4 The Euler Scheme

Let $\{X_t\}$ be an Ito process on $t \in [t_0, T]$ satisfying the stochastic differential equation

$$\begin{cases} X_t = a(t, X_t)dt + b(t, X_t)dW_t \\ X_{t_0} = X_0 \end{cases} \tag{3.43}$$

For a given discretization $t_0 = \tau_0 < \tau_1 < \cdots < \tau_N = T$, an Euler approximation is a continuous time stochastic process $\{Y_t\}$ satisfying the iterative scheme

$$Y_{n+1} = Y_n + a(\tau_n, Y_n)(\tau_{n+1} - \tau_n) + b(\tau_n, Y_n)(W_{\tau_{n+1}} - W_\tau) \tag{3.44}$$

with initial value

$$Y_0 = X_0 \tag{3.45}$$

Following theorem [1] states it is an *order 0.5 strong Ito-Taylor approximation.*

Theorem 3.9. *Suppose the following conditions for the Ito process (79):*

- $E(|X_0|^2) \leq \infty$
- $E(|X_0 - Y_0^\delta|^2)^{1/2} \leq C_1\delta^{1/2}$
- $|a(t,x) - a(t,y)| + |b(t,x) - b(t,y)| \leq C_2|x - y|$
- $|a(t,x)| + |b(t,x)| \leq C_3(1 + |x|)$
- $|a(t,x) - a(s,x)| + |b(t,x) - b(s,x)| \leq C_3(1 + |x|)|s - t|^{1/2}$

for all $t, s \in [0,T]$ and $x, y \in \mathbb{R}^d$, where the constants do not depend on δ. Then for the Euler approximation Y^δ the estimate

$$E(|X_T - Y_T^\delta|) \leq C\delta^{1/2} \tag{3.46}$$

holds,where C does not depend on δ.

Please read [1] for the proof.

3.5 The Milstein Scheme

We shall now look at a scheme proposed by Milstein, which turns out to be an *order 1.0 strong Taylor scheme.* In the 1-dimensional case with $d = m = 1$, we add to the Euler scheme the term

$$\frac{1}{2}bb_{X_t}[(dW)^2 - dt] \tag{3.47}$$

from the Ito-Taylor expansion, then we obtain the *Milstein scheme*

$$Y_{n+1} = Y_n + a(t_n, Y_n)dt + b(t_n, Y_n)dW + \frac{1}{2}b(t_n, Y_n)b_{X_t}(t_n, Y_n)[(dW)^2 - dt]. \tag{3.48}$$

In the multi-dimensional case with $m = 1$ and $d \geq 1$ the kth component of Milstein Scheme is

$$Y_{n+1}^k = Y_n^k + a^k(t_n, Y_n)dt + b^k(t_n, Y_n)dW + \frac{1}{2}\sum_{i=1}^{d} b^i(t_n, Y_n)\frac{\partial b^k}{\partial x^i}(t_n, Y_n)[(dW)^2 - dt]. \tag{3.49}$$

In the general multi-dimensional case with $d,\ m = 1, 2, \ldots$ the kth component of the Milstein scheme is

$$Y_{n+1}^k = Y_n^k + a^k dt + \sum_{j=1}^{m} b^{k,j}\Delta W^j + \sum_{j_1,j_2=1}^{m} L^{j_1} b^{k,j_2} I_{(j_1,j_2)} \tag{3.50}$$

When $j_1 = j_2$ we have

$$I_{(j_1,j_2)} = \int_{\tau_n}^{\tau_{n+1}} \int_{\tau_n}^{s_1} dW_{s_1}^{j_1} dW_{s_1}^{j_2} \tag{3.51}$$

For $j_1 \neq j_2$, the multiple stochastic integrals $I_{(j_1,j_2)}$ cannot be so easily expressed in terms of the increments $\Delta W^{j_1},\ \Delta W^{j_2}$ of the components of the Wiener process as in the case $j_1 = j_2$. But it can be approximated by following way.

Let p be a positive integer, and $r = 1, 2, \ldots, p$, then for $j_1 \neq j_2$ with $j_1, j_2 = 1, \ldots, m$ we can approximate the multiple stochastic integrals $I_{(j_1,j_2)}$ as

$$I_{(j_1,j_2)}^p = \frac{1}{2}dt\xi_{j_1}\xi_{j_2} - \frac{1}{2}\sqrt{dt}(a_{j_2,0}\xi_{j_1} - a_{j_1,0}\xi_{j_2}) + dtA^p \tag{3.52}$$

with

$$A^p = \frac{1}{2\pi}\sum_{r=1}^{p}\frac{1}{r}(\zeta_{j_1,r}\eta_{j_2,r} - \zeta_{j_2,r}\eta_{j_1,r}) \tag{3.53}$$

where $\xi_j, \zeta_{j,r},\ \eta_{j,r}$ are independent standard Gaussian random variables that are determined by dt and W^j.

3.6 The Order 1.5 Strong Taylor Scheme

By adding more terms from the Ito-Taylor expansion to the Milstein scheme, in the 1-dimensional case $d = m = 1$, *the order 1.5 strong Ito-Taylor scheme* is as follows:

$$\begin{aligned} Y_{n+1} = Y_n + adt + bdW + \frac{1}{2}bb_{X_t}(\Delta W^2 - dt) + (a_t + aa_{X_t} + \frac{1}{2}b\frac{\partial^2 a}{\partial X_t^2})\frac{dt^2}{2} \\ + ba_{X_t}\Delta Z + (b_t + ab_{X_t} + \frac{1}{2}b^2\frac{\partial^2 b}{\partial X_t^2})I_{(0,1)} + f_{(1,1,1)}I_{(1,1,1)} \end{aligned} \tag{3.54}$$

where

$$\begin{aligned} \Delta Z = I_{(1,0)} &= \int_{t_n}^{t_{n+1}} \int_{t_n}^{s} dW_r ds \\ f_{(1,1,1)} &= b(b_{X_t}\frac{\partial^2 b}{\partial X_t^2} + b\frac{\partial^2 b}{\partial X_t^2}) \\ I_{(1,1,1)} &= \frac{\Delta W^3}{6} - \frac{\Delta W dt}{2} \end{aligned} \tag{3.55}$$

we cannot simplify $I_{(0,1)}$ into a form without an integral, but we can represent it in term of ΔZ and ΔW. Note that

$$\begin{aligned} I_{(0,1)} &= \int_t^s \int_t^{s_1} ds_2 dW_{s_1} \\ &= \int_t^s s_1 dW s_1 - \int_t^s t dW s_1 \\ &= sW_s - tW_t - \int_s^t W_{s_1} ds_1 - t(W_s - W_t) \\ &= W_s dt - \int_s^t W_{s_1} ds_1 \end{aligned} \tag{3.56}$$

and

$$\Delta Z = \int_t^s (W_{s_1} - W_t) ds_1 = \int_s^t W_{s_1} ds_1 - W_t dt \tag{3.57}$$

thus we have

$$I_{(0,1)} + \Delta Z = \Delta W dt \Rightarrow I_{(0,1)} = \Delta W dt - \Delta Z \tag{3.58}$$

Now claim that

$$\Delta Z = I_{(1,0)} = \int_{t_n}^{t_{n+1}} \int_{t_n}^{s} dW_r ds \tag{3.59}$$

is normally distributed with mean 0, variance $dt^3/3$, and covariance $E(\Delta Z \Delta W) = dt^2/2$.

Proof. From equation (3.57) we have

$$E(\Delta Z) = E\left(\int_t^s W_{s_1} - W_t ds_1\right) = \int_t^s E(W_{s_1} - W_t) ds_1 = \int_t^s E(W_{s_1}) - E(W_t) ds_1 = 0, \quad (3.60)$$

and

$$\begin{aligned} E(\Delta Z \Delta W) &= E\left[\left(\int_t^s W_{s_1} ds_1 - W_t\right)(W_s - W_t)\right] \\ &= E\left(\int_t^s W_{s_1} W_s ds_1\right) - E\left(\int_t^s W_{s_1} W_t ds_1\right) - (s-t)(EW_s Wt - EW_t^2) \\ &= \int_t^s s_1 ds_1 - \int_t^s t ds_1 \\ &= \frac{s^2 - t^2}{2} - t(s-t) = \frac{dt^2}{2} \end{aligned} \quad (3.61)$$

which proves the result for the covariance. For the variance,by equation (3.57) we have

$$\begin{aligned} E(\Delta Z^2) &= E\left[\left(\int_t^s W_{s_1} ds_1 - W_t dt\right)^2\right] \\ &= E\left[\left(\int_t^s W_{s_1} ds_1\right)^2\right] - 2(s-t)E\left(\int_t^s W_t W_{s_1} ds_1\right) + (s-t)^2 EW_t^2 \\ &= A - 2(s-t)\int_t^s E(W_t W_{s_1}) ds_1 + t(s-t)^2 \\ &= A - 2(s-t)\int_t^s t ds_1 + t(s-t)^2 \\ &= A - 2t(s-t)^2 + t(s-t)^2 \\ &= A - t(s-t)^2 \end{aligned} \quad (3.62)$$

where

$$
\begin{aligned}
A &:= E\left[\left(\int_t^s W_{s_1} ds_1\right)^2\right] \\
&= E\left[\left(sW_s - tW_t - \int_t^s s_1 dW_{s_1}\right)^2\right] \\
&= E[(sW_s - tW_t)^2] - 2E\left[(sW_s - tW_t)\int_t^s s_1 dW_{s_1}\right] + E\left[\left(\int_t^s s_1 dW_{s_1}\right)^2\right] \\
&= E(s^2W_s^2 - 2stW_sW_t + t^2W_t^2) - 2E\left[sW_s\int_t^s s_1 dW_{s_1}\right] + E\left(\int_t^s s_1^2 ds_1\right) \\
&= s^3 - 2st^2 + t^3 - 2W + \frac{s^3 - t^3}{3}
\end{aligned} \tag{3.63}
$$

where

$$
B := E\left[sW_s\int_t^s s_1 dW_{s_1}\right]. \tag{3.64}
$$

Let $\{\phi_n\}$ be a sequence of elementary functions such that

$$
E\left[\int_S^T (s_1 - \phi)^2 dt\right] \to 0 \quad \text{as. } n \to \infty \tag{3.65}
$$

then

$$
\begin{aligned}
B &= sE\left((W_s \lim_{n\to\infty}\int_t^s \phi_n dW_{s_1})\right) \\
&= s\lim_{n\to\infty} E\left(\sum_{j\geq 1} e_j W_s(W_{j+1} - W_j)\right) \\
&= s\lim_{n\to\infty} E\left(\sum_{j\geq 1} e_j(W_s - W_{j+1} + W_{j+1})(W_{j+1} - W_j)\right) \\
&= s\lim_{n\to\infty} E\left(\sum_{j\geq 1} e_j(W_s - W_{j+1})(W_{j+1} - W_j) + e_j W_{j+1}(W_{j+1} - W_j)\right) \\
&= s\lim_{n\to\infty}\sum_{j\geq 1} E\left(e_j W_{j+1}(W_{j+1} - W_j)\right) \\
&= s\lim_{n\to\infty}\sum_{j\geq 1} e_j(t_{j+1} - t_j) \\
&= s\int_t^s s_1 ds_1 = \frac{s^3 - st^2}{2}
\end{aligned} \tag{3.66}
$$

putting all this together we get the variance. □

In addition, the pair of correlated normally distributed random variables $(\Delta W, \Delta Z)$ can be determined from two independent $N(0,1)$ distributed random variables $U_1,\ U_2$ with transformation

$$\Delta W = U_1\sqrt{dt} \quad \Delta Z = \frac{1}{2}dt^{3/2}\left(U_1 + \frac{1}{\sqrt{3}}U_2\right). \tag{3.67}$$

In the general multi-dimensional case with $d,\ m = 1, 2, \cdots$ the kth component of the order 1.5 Taylor scheme is

$$\begin{aligned} Y_{n+1}^k = Y_n^k &+ a^k dt + \frac{1}{2}L^0 a^k dt^2 \\ &+ \sum_{j=1}^{m}(b^{k,j}\Delta W^j + L^0 b^{k,j} I(0,j) + L^j a^k I_{(j,0)}) \\ &+ \sum_{j_1,j_2=1}^{m} L^{j_1} b^{k,j_1} I_{(j_1,j_2)} \\ &+ \sum_{j_1,j_2,j_3=1}^{m} L^{j_1}L^{j_2} b^{k,j_3} I_{(j_1,j_2,j_3)} \end{aligned} \tag{3.68}$$

3.7 Higher Order Ito-Taylor Approximation

By adding more terms from Ito-Taylor expansion, we can get higher order approximations. However some Multi-Ito integrals are not easy to calculate and they require more numerical approximation and more calculating time. Sometimes we have to use Stratonovich integral to approximate them. From my experience, if the SDE is simple, then the Milstein scheme is sufficient. I refer the readers to [1] for the details on how to get the higher order Ito-Taylor Approximation.

Chapter 4. Applications

An important application of numerical approximations is the simulation of stochastic dynamical systems. The examples in this chapter use the approximation schemes mentioned above to simulate some well-known deterministic ordinary differential equations with random noise. There will also be an example on how these schemes could be used to price spread options.

4.1 Stochastic Duffing equation

4.1.1 Deterministic Duffing equation. The Duffing equation is a non-linear second-order ODE. The equation is given by

$$\ddot{x} + \delta\dot{x} + \beta x + \alpha x^3 = \gamma\cos(\omega t + \phi). \tag{4.1}$$

This equation describes the motion of a damped oscillator with a potential. For example, it models a spring pendulum whose spring's stiffness does not exactly obey hooke's law.

For $\beta > 0$, the Duffing oscillator can be interpreted as a forced oscillator with a spring with restoring force $F = -\beta x - \alpha x^3$. When $\alpha > 0$, this spring is called a hardening spring, and, when $\alpha < 0$, it is called a softening spring.

For $\beta < 0$, the Duffing oscillator describes the dynamics of a point mass in a double well potential. It is known that chaotic motions can be observed in this case.

We should look at the simplified version of a Duffing equation with no forcing.

$$\ddot{x} + \delta\dot{x} - x + x^3 = 0. \tag{4.2}$$

If we let $x = x_1$ and $\dot{x_1} = x_2$, we obtain the following system of ODEs

$$\begin{cases} \dot{x_1} = x_2 \\ \dot{x_2} = -\delta x_2 + x_1 - x_1^3 \end{cases} \tag{4.3}$$

The fixed points of this system of ODEs are $(0,0)$, $(1,0)$, and $(-1,0)$. The stability of the fixed points can be determined by looking at the Jacobian matrix

$$\begin{bmatrix} 0 & 1 \\ 1-3x_1^2 & -\delta \end{bmatrix}. \tag{4.4}$$

For the point $(0,0)$, we find the eigenvalues to be

$$\lambda_{(0,0)} = \frac{-\delta \pm \sqrt{\delta^4+4}}{2}, \tag{4.5}$$

Since $\delta^2 > 0$, both eigenvalues are real. Also because $\sqrt{\delta^2+4} > |\delta|$, one eigenvalue is positive. So this point is unstable.

For the fixed points $(\pm 1,0)$, the eigenvalues are

$$\lambda_{(\pm 1,0)} = \frac{-\delta \pm \sqrt{\delta^2-8}}{2}, \tag{4.6}$$

If $\delta > 0$, then the real part of both eigenvalues are negative, so the points are asymptotically stable. If $\delta < 0$, then the eigenvalues have positive real part, so the fixed points are unstable. When $\delta = 0$, the eigenvalues have no real part, and there should be periodic solutions.

The following picture shows the trajectories when $\delta = 0$ and there is no damping.

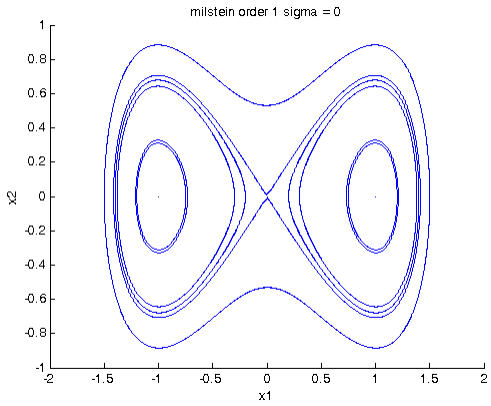

Figure 4.1: Trajectories of undamped deterministic Duffing equation

The following picture shows how the trajectories with $\delta = 1$ with same initial values as in the figure 4.1.

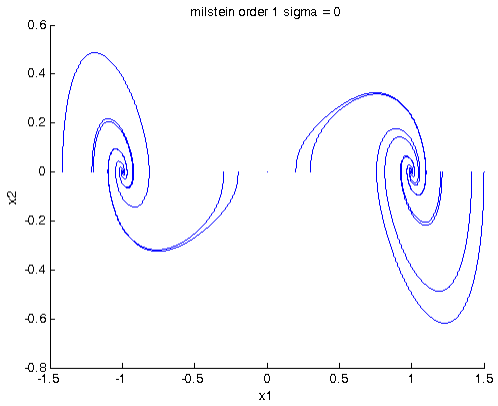

Figure 4.2: Trajectories of damped deterministic Duffing equation

4.1.2 Stochastic Duffing Equation. Now by adding some noise we get the Stochastic Duffing equation

$$\begin{cases} \dot{X}_1 = X_2 \\ \dot{X}_2 = -\delta X_2 + X_1 - X_1^3 + \sigma X_1 \dot{W}, \end{cases} \tag{4.7}$$

where σ is a positive constant and W is a one-dimensional Wiener process. The question is what changes in the trajectories of the solution will be observed with the addition of this multiplicative noise.

Suppose $\delta = 0$, the there is no damping. I used the Milstein scheme with step size $dt = 2^{-15}$ and $\sigma = 0.2$ to simulate the trajectories of the system (4.7). Each trajectory has different initial value but with same random path. The result is shown below

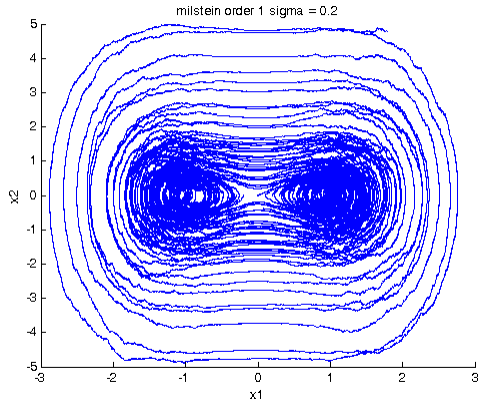

Figure 4.3: Trajectories of undamped stochastic Duffing equation with $\sigma = 0.2$

As shown in the figure 4.3, the solution is not periodic but is still has periodic-like behavior. With larger values of σ, it seems like more chaotic as shown in figures 4.4 and 4.5.

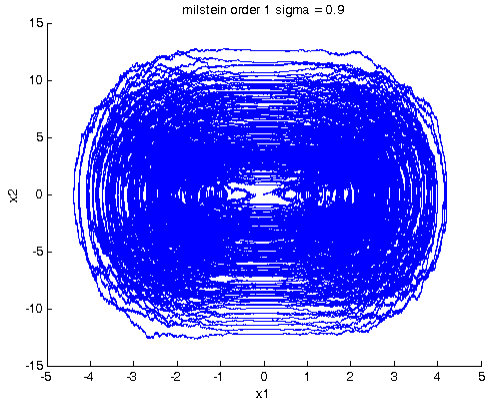

Figure 4.4: Trajectories of undamped stochastic Duffing equation with $\sigma = 0.9$

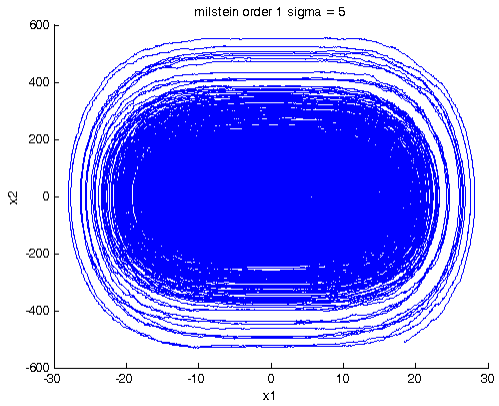

Figure 4.5: Trajectories of undamped stochastic Duffing equation with $\sigma = 5$

It is clearer if we only look at one solution. This is the case with no random noise.

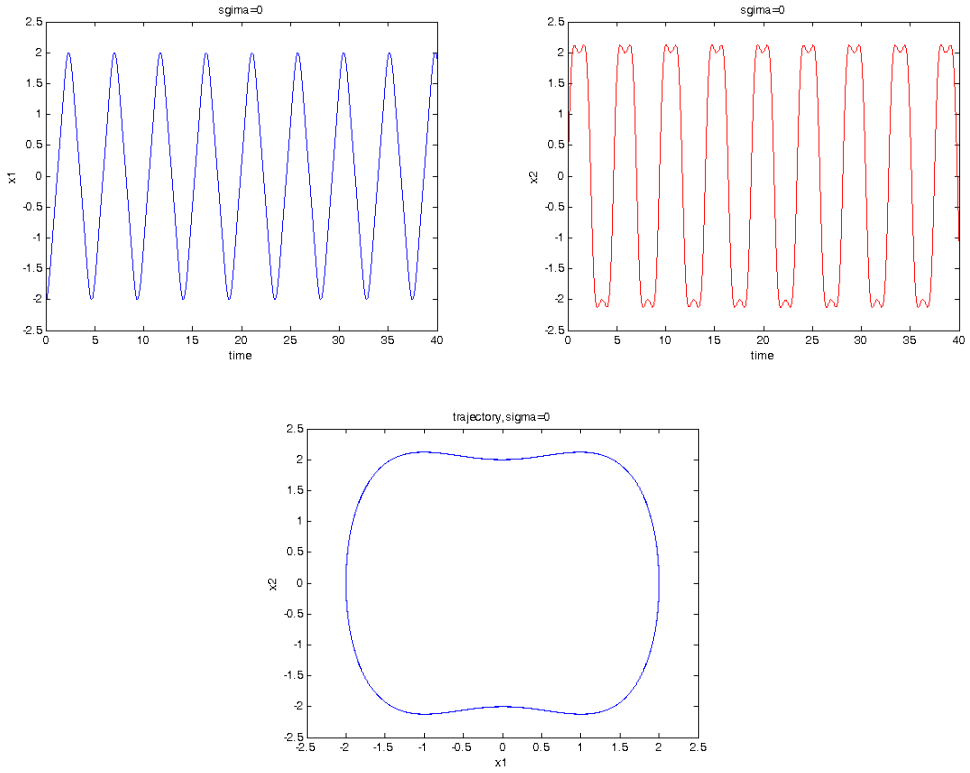

Figure 4.6: One trajectory of undamped deterministic Duffing equation.

Here is what happened after we add the some random noise. As you can see, with random noise, the solution is no longer periodic.

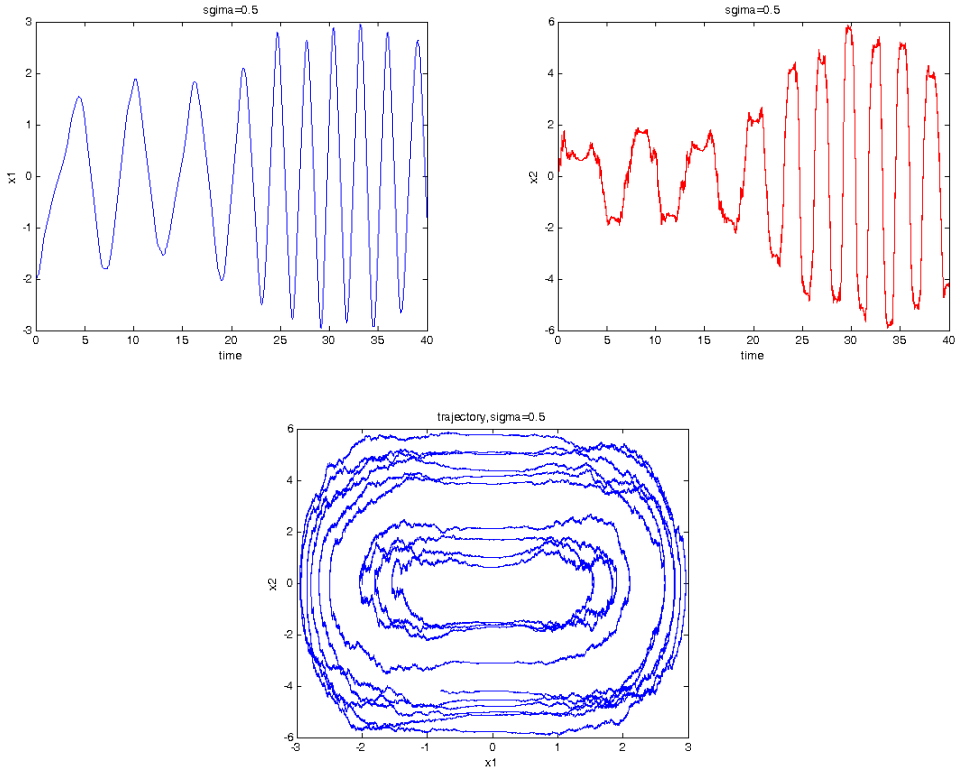

Figure 4.7: One trajectory of undamped stochastic Duffing equation with $\sigma = 0.5$.

Suppose there is periodic forcing $F(t) = \cos t$ and no random noise. Using the same initial values as in figure 4.1, the trajectories are like following:

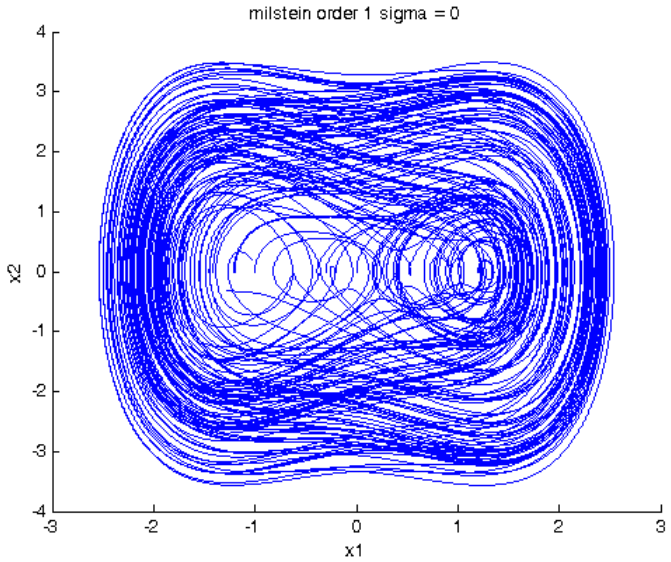

Figure 4.8: Trajectories of undamped deterministic Duffing equation with force $F(t) = \cos t$.

By adding more and more noise, the trajectories' behavior gets more and more chaotic as shown below:

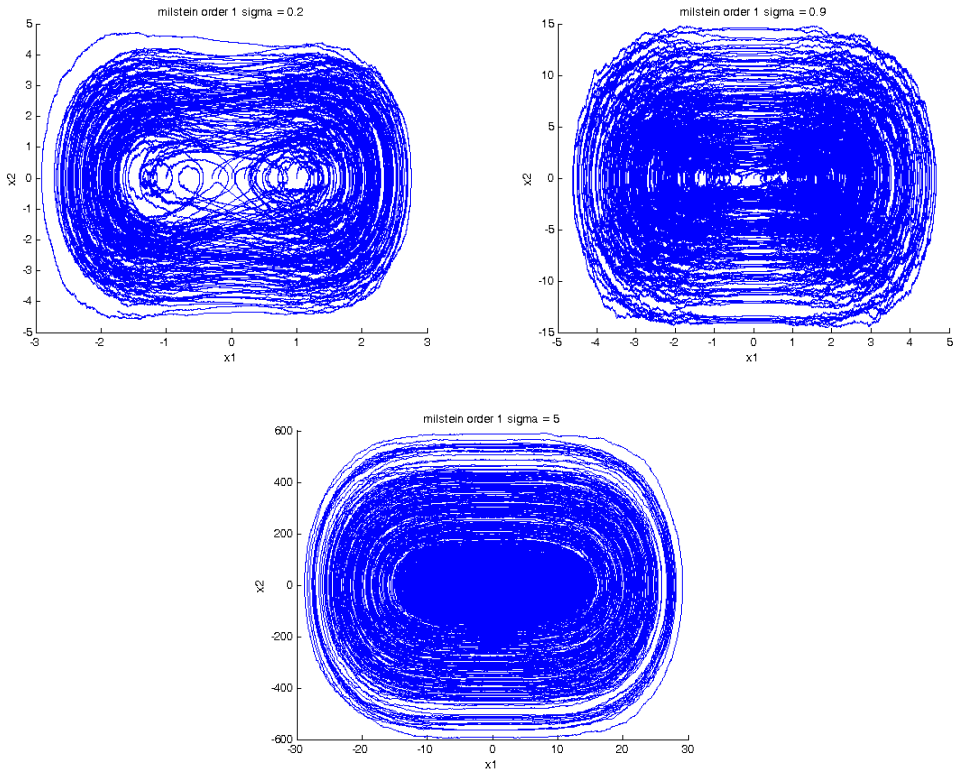

Figure 4.9: Trajectories of undamped stochastic Duffing equation with and forcing $F(t) = \cos t$.

By only looking at one trajectory, it is clear that when there is no random noise, the solution is still periodic with a longer period and the magnitude changes as well, as shown below:

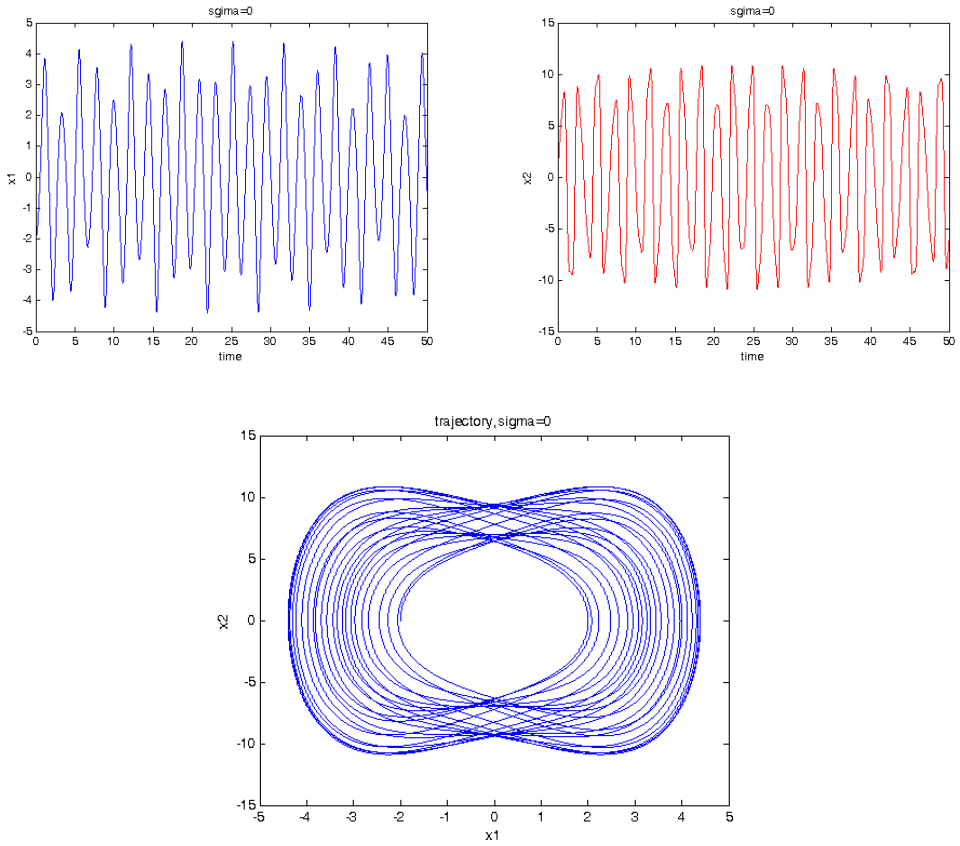

Figure 4.10: One trajectory of undamped deterministic Duffing equation with forcing $F(t) = \cos t$.

However if random noise is added then the solution is not periodic and its values oscillate.

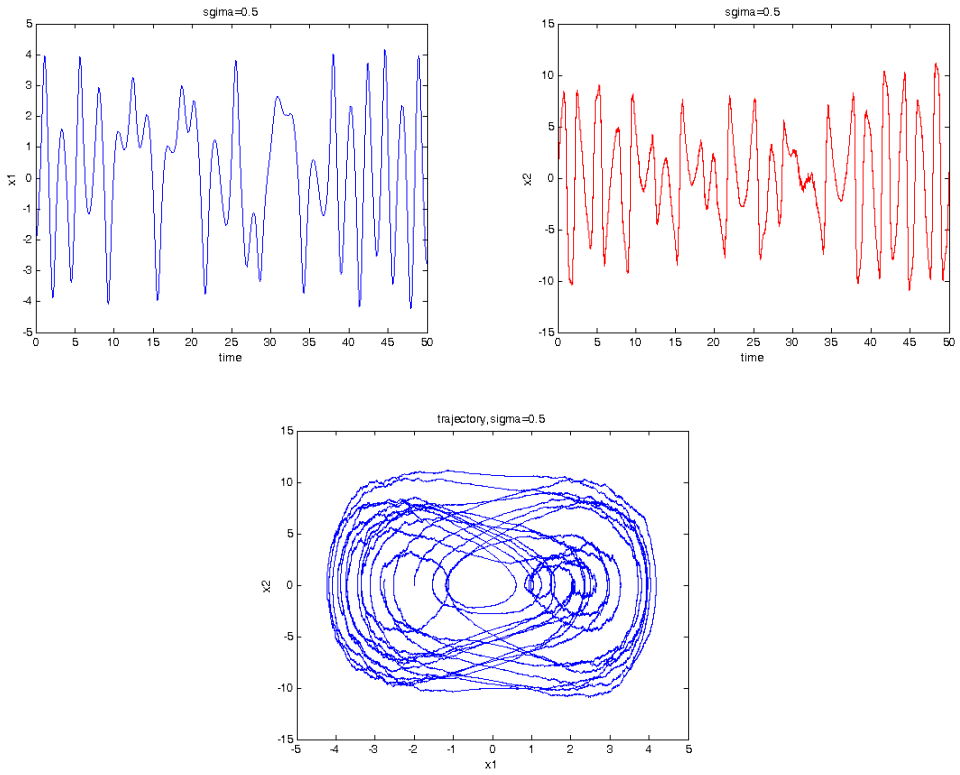

Figure 4.11: One trajectory of undamped stochastic Duffing equation with $\sigma = 0.5$ and forcing $F(t) = \cos t$.

Now suppose the force is $F(t) = \cos(t) + \sin(\sqrt{t})$, with two frequencies. Assume there is not random noise, and same initial values as before. Then the trajectories are as the following.

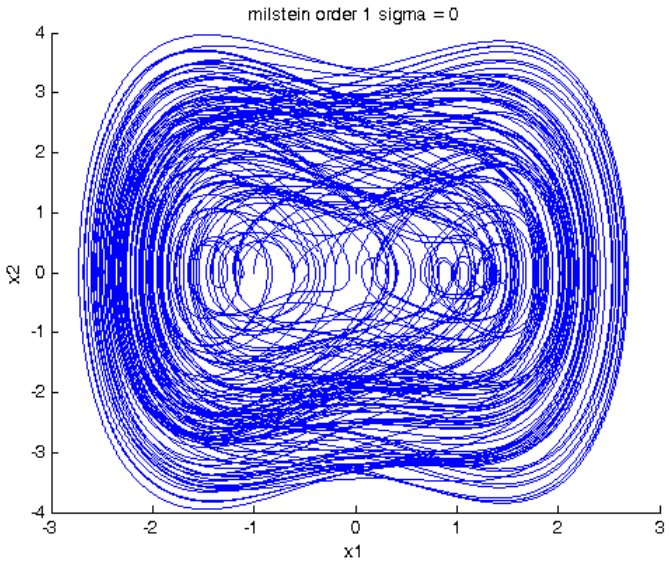

Figure 4.12: Trajectories of undamped deterministic Duffing equation with force $F(t) = \cos(t) + \sin(\sqrt{t})$.

The plot is more complicated than with single frequency due to the double frequencies. By adding random noise, the result is the same as pervious examples: the trajectories' behavior gets more chaotic.

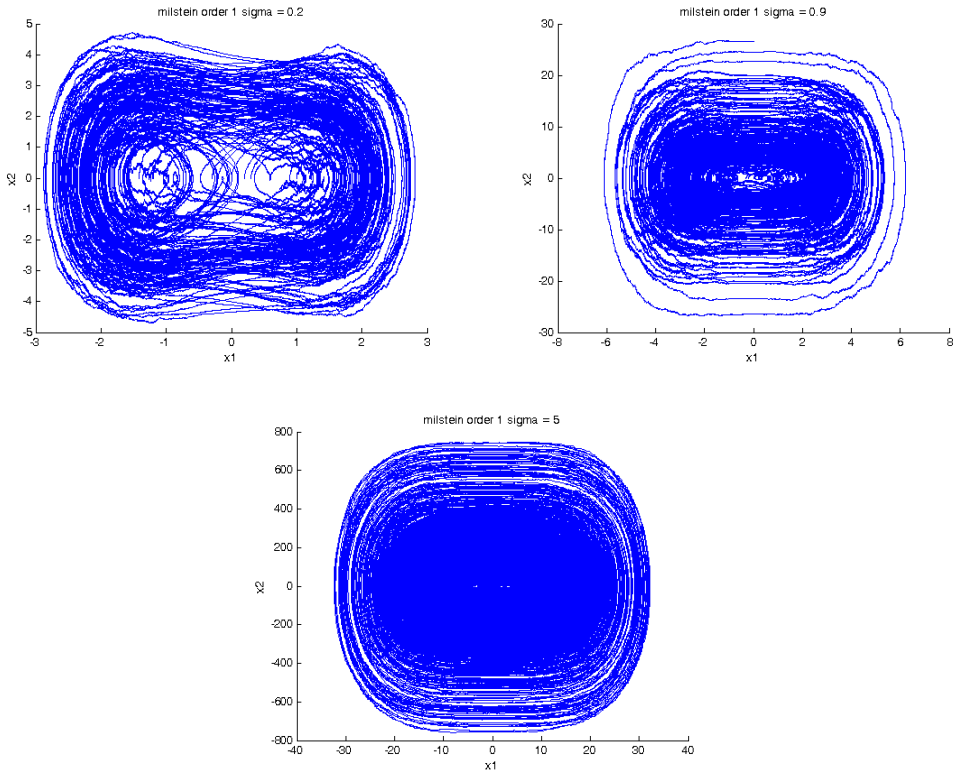

Figure 4.13: Trajectories of undamped stochastic Duffing equation with force $F(t) = \cos(t) + \sin(\sqrt{t})$.

Figure blow shows the plots for only one trajectory. The plots on the left side are deterministic and on the right side are stochastic.

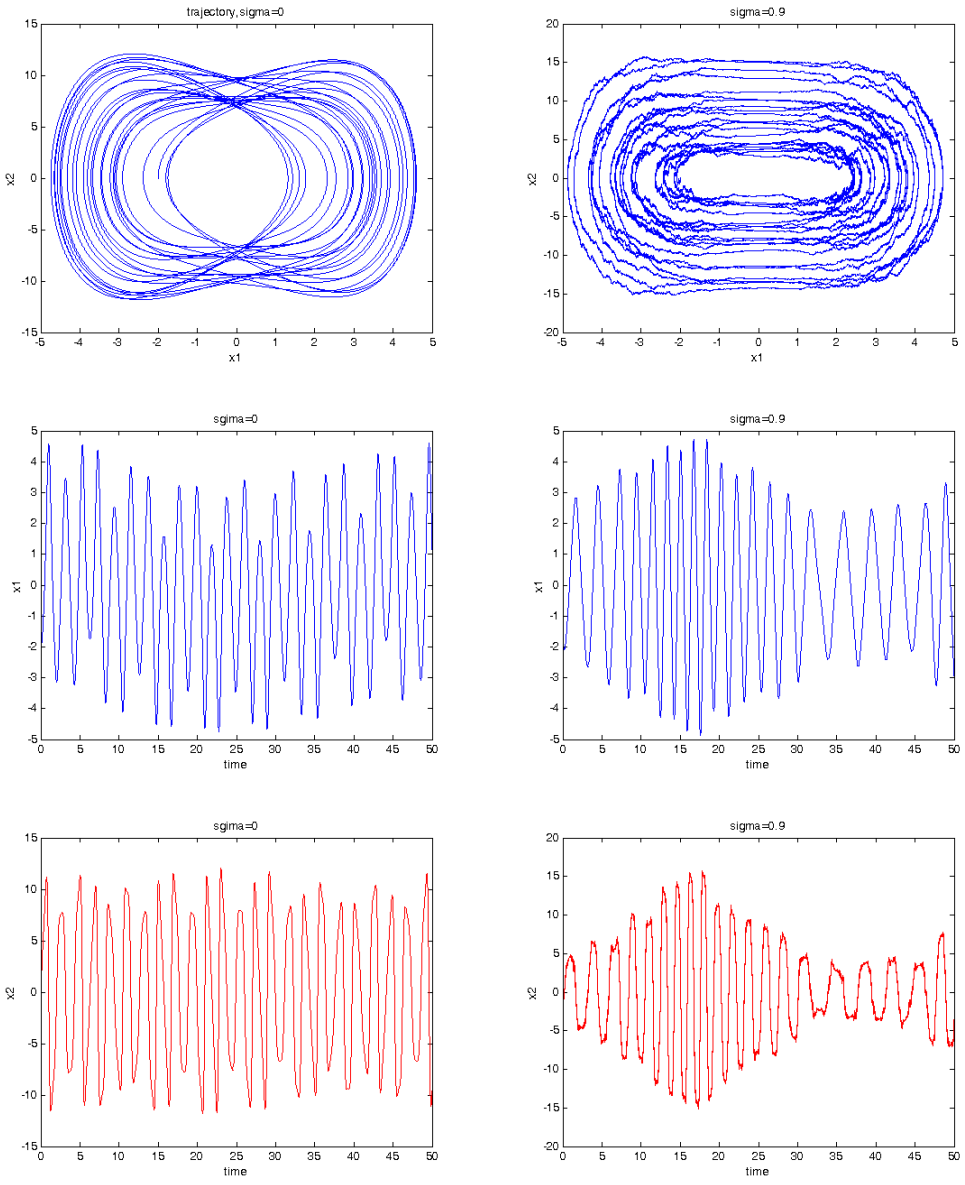

Figure 4.14: Trajectories of undamped deterministic and stochastic Duffing equation with force $F(t) = \cos(t) + \sin(\sqrt{t})$.

Let us consider another case by adding a non periodic force, $F(t) = t^2 + 1$. The simulated results are shown in figure 4.15.

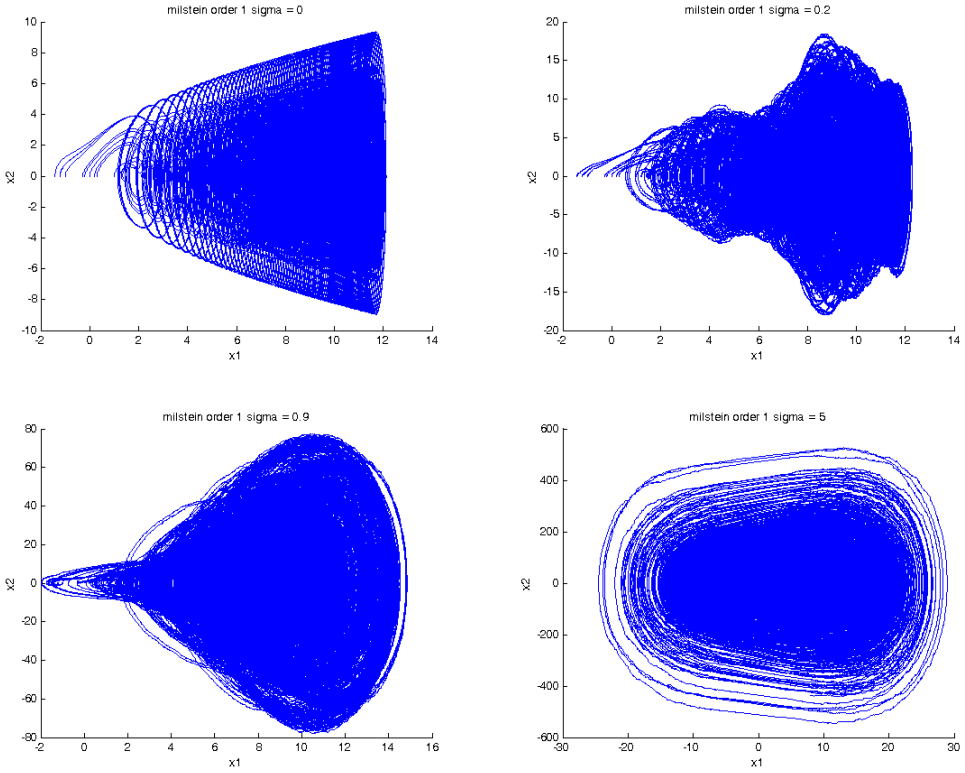

Figure 4.15: Trajectories of undamped deterministic and stochastic Duffing equation with force $F(t) = t^2 + 1$.

Plots of one trajectory can be found in the **Appendix A**.

Suppose there is damping, taking $\delta = 1$ and no forcing. I used the Milstein scheme again with step size $dt = 2^{-15}$ and $\sigma = 0.2$ to simulate the trajectories of the system (4.7). Each trajectory has a different initial value but with the same random path. My results are in figure 4.16:

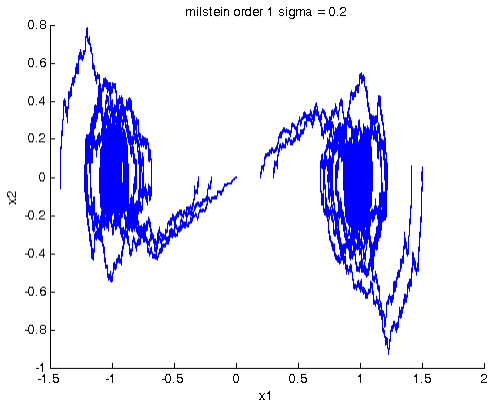

Figure 4.16: Trajectories of damped and unforced stochastic Duffing equation with $\sigma = 0.2$.

As you can see, the trajectories converge to a neighborhood of the three fixed points $(\pm 1, 0),\ (0, 0)$. But there are some trajectories go through the points (0,0) and (1,0).

I did another simulation with same path but a larger values of σ, the results are showed in figure 4.17 and 4.18.

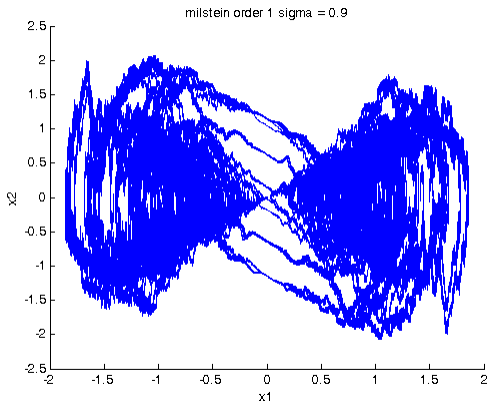

Figure 4.17: Trajectories of damped and unforced stochastic Duffing equation with $\sigma = 0.9$.

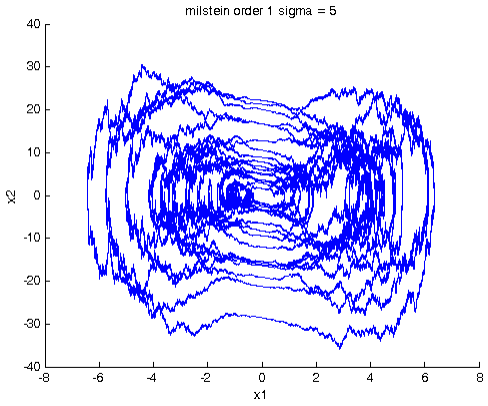

Figure 4.18: Trajectories of damped and unforced stochastic Duffing equation with $\sigma = 5$.

The trajectories go through all three stationary points and converge to a neighborhood of all three points. The result is more clear if we only look at one trajectory.

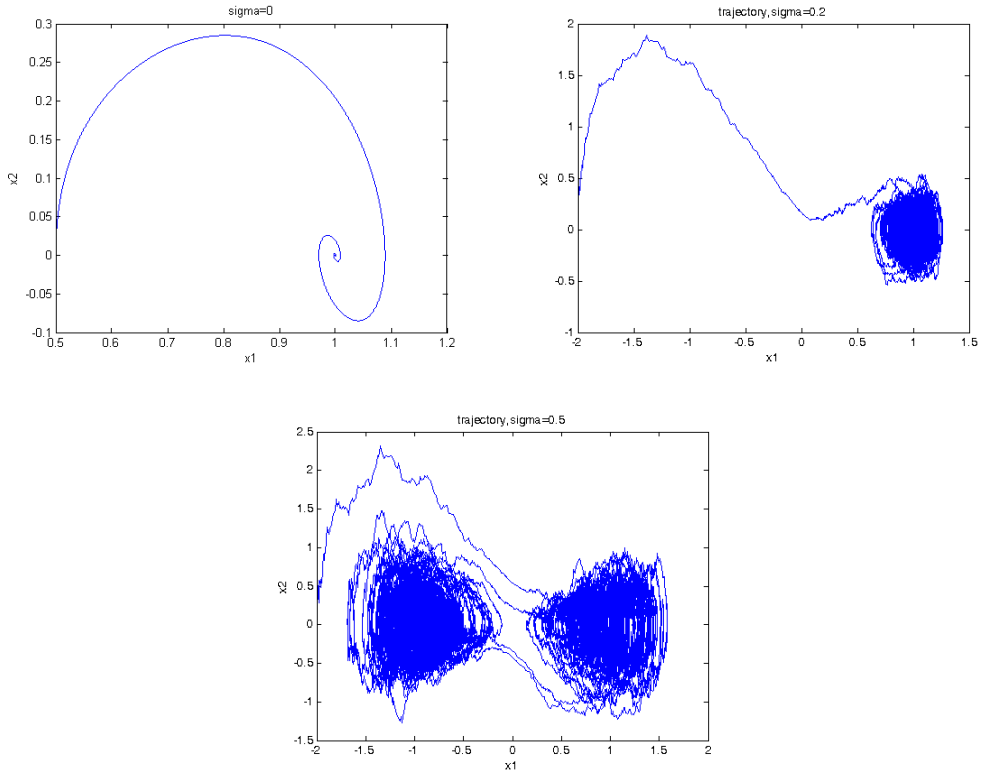

Figure 4.19: One trajectory of damped and unforced stochastic Duffing equation with $\sigma = 0.2$ and $\sigma = 0.5$.

Let us compare the flow of x_1 for both the deterministic case and the stochastic case with different values of σ.

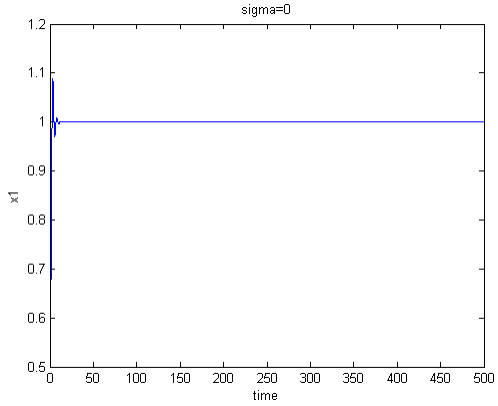

Figure 4.20: Plot of x_1 of deterministic Duffing equation.

Figure 4.20 shows the deterministic case. The flow goes to 1, and states there since $(1, 0)$ is stable.

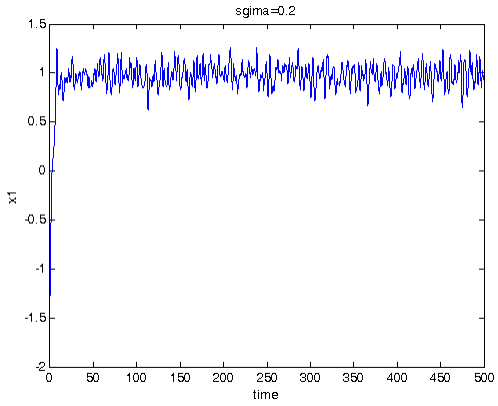

Figure 4.21: Plot of x_1 of stochastic Duffing equation with $\sigma = 0.2$.

With small noise, $\sigma = 0.2$, as show in figure 4.21, the flow oscillates in a neighborhood of 1.

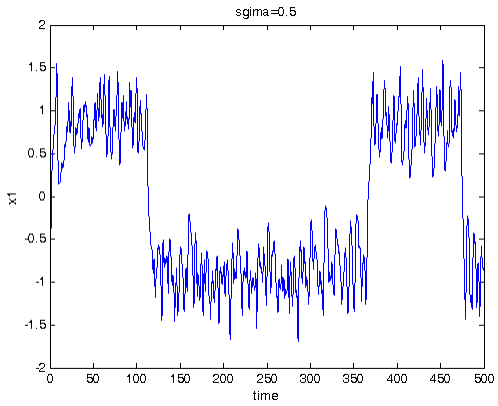

Figure 4.22: Plot of x_1 of stochastic Duffing equation with $\sigma = 0.5$.

In figure 4.22, $\sigma = 0.5$, we can see clearly that the flow goes to 1 and stays there for a certain amount of time, then jumps to another stable point $(-1, 0)$, stays for a certain amount of time, then jumps back to $(1, 0)$.

Now let us look at two differential flows with different initial values for x_1.

In the deterministic case, they convergent to -1 and 1 and stays there. In the stochastic case, they crosse each other couple times as showmen in the graph

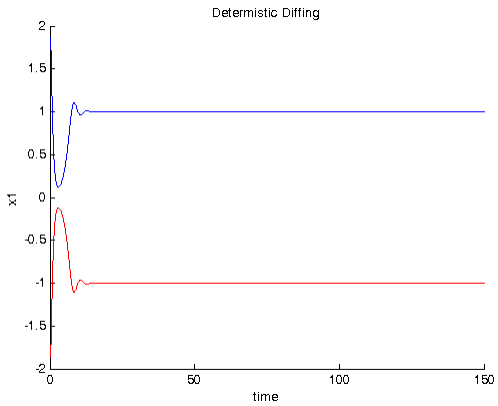

Figure 4.23: Plot of x_1 of deterministic Duffing equation with two initial values.

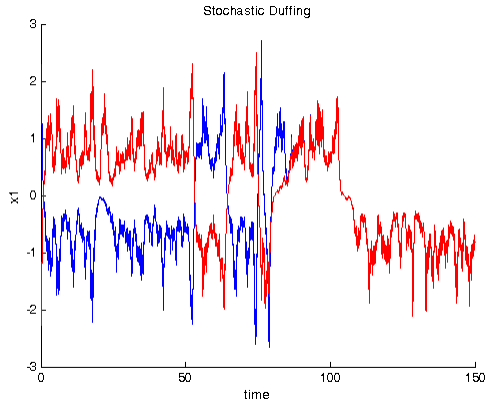

Figure 4.24: Plot of x_1 of stochastic Duffing equation with two initial values.

Is this strange behavior caused by the numerical scheme I used? In order to make sure it is not caused by the numerical scheme. I ran more simulations with smaller step sizes and some higher order schemes.

First, I used the Milstein scheme, but with different step sizes, namely 2^{-14}, 2^{-15}, 2^{-16}, and 2^{-17}. We will only look at two trajectories. The following four pictures show the result with one path for each step size with $\sigma = 0.5$.

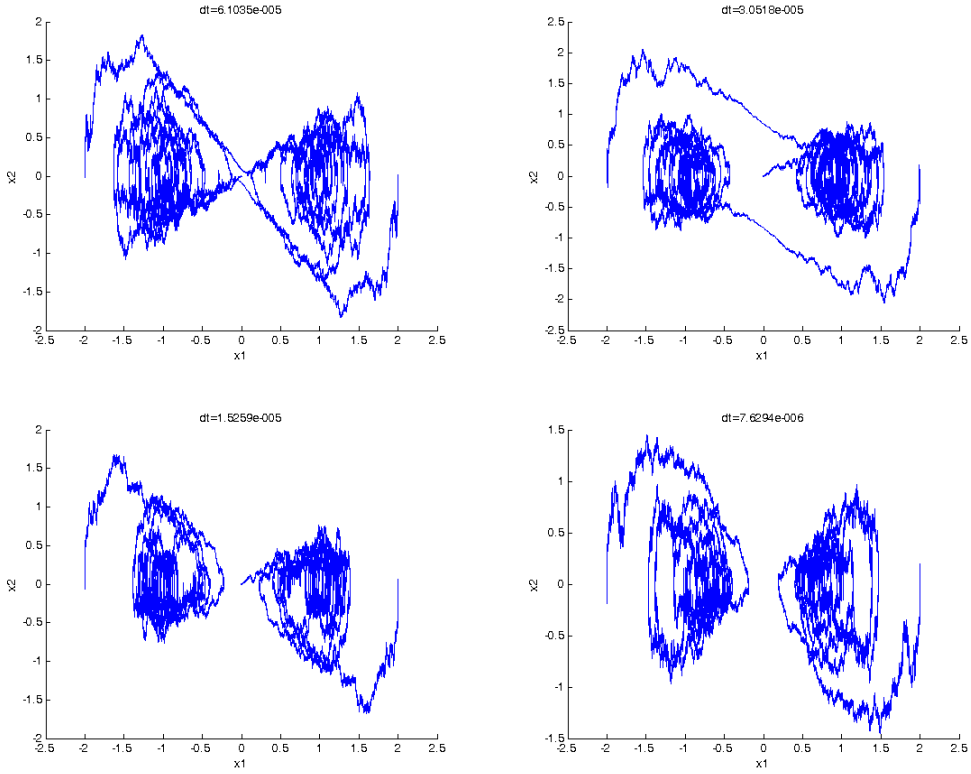

Figure 4.25: Comparison of different step sizes for stochastic Duffing equation I.

When the step size is smaller, the cross-over does not happen as often. However the figures only show the case with one path for the Brownian motion. To be more precise I took 10000 paths for each step size, I counted how many times it crossed 1. The result is shown in figure 4.26:

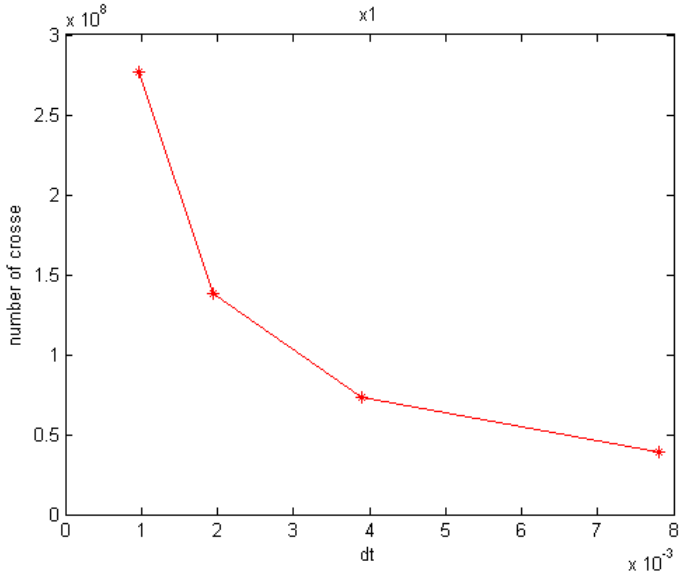

Figure 4.26: Comparison of different step sizes for stochastic Duffing equation II.

From the plot, we can see that as the step size decreases, the number of cross-overs between the two solutions decreases.

4.2 Stochastic Lorenz Equation

4.2.1 Lorenz equation. In 1963 Ed Lorenz derived a three-dimensional system from a simplified model of convection rolls in the atmosphere. The equations are:

$$\begin{aligned} \frac{dx}{dt} &= \sigma(y - x) \\ \frac{dy}{dt} &= rx - y - xz \\ \frac{dz}{dt} &= xy - bz. \end{aligned} \tag{4.8}$$

Lorenz discovered that this system could have extremely erratic dynamics: over a wide range of parameters, the solutions oscillate irregularly, never exactly repeating but always remaining in a bounded region of phase. Unlike stable fixed points and limit cycles, the strange attractor is not a point or a curve or even a surface–it's a fractal, with a fractional dimension between 2 and 3.[3]

In the system (4.8), σ, b and r, are positive real numbers. In the context of bifurcation theory

it is usual to fix σ and b and let r vary.

In the original context, r acted as a measure of the imposed temperature difference between the bottom of the fluid layer and the top. In the same vein, x measured the flow speed, while y and z denoted certain broad features of the temperature distribution.[4]

Some well-known properties of the Lorenz equation include:

- **Symmetry**

  If we replace (x, y) by $(-x, -y)$ in (4.8), the equations stays the same. Therefore if $\mathbf{u}(t) = (x(t), y(t), z(t))$ is a solution of (4.8), then $\mathbf{v} = (-x(t), -y(t), z(t))$ is also a solution. Thus all solutions are symmetric.

- **Fixed points**

  Like the Duffing equation, we would look at the fixed points of system (4.8). Clearly that the origin $(0, 0, 0)$ is a fixed point for all values of the parameters σ, b, and r. For $r > 1$, there is also a pair of fixed points $C^{\pm} = (\pm\sqrt{b(r-1)}, \pm\sqrt{b(r-1)}, r-1)$.

- **Stability of (0,0,0)**

  If we omit the nonlinear part of (4.8), we have the linearized system

$$\begin{cases} \dot{x} = \sigma(y - x) \\ \dot{y} = rx - y \\ \dot{z} = -bz. \end{cases} \tag{4.9}$$

  It is easy to see that on the z-direction, we have $z(t) = \exp -bt$, thus $z(t) \to 0$ exponentially as $t \to \infty$. Now looking at x and y only we have

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} -\sigma & \sigma \\ r & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}. \tag{4.10}$$

  The eigenvalues of matrix (4.10) are given by

$$\lambda_{1,2} = \frac{-(\sigma + 1) \pm \sqrt{(\sigma^2 + 1) - 4\sigma(1 - r)}}{2}. \tag{4.11}$$

If $r > 1$, then both eigenvalues are real and one is positive, thus the origin is a saddle point. If $r < 1$, then the eigenvalues are complex, and since $\sqrt{(\sigma^2+1) - 4\sigma(1-r)} < |\sigma + 1|$, thus both eigenvalues have negative real parts and thus the origin is linearly stable.

For the global stability, in [3] the author showed that for $r < 1$, the origin is globally stable. I refer the reader to [3] for the proof.

- **Stability of $C^{\pm}$**

  Now suppose that $r > 1$, then both $C^{\pm}$ are linearly stable (assuming $\sigma - b - 1 > 0$) if we have [3]

$$1 < r < \frac{\sigma(\sigma + b + 3)}{\sigma - b - 1}. \tag{4.12}$$

In Lorenz's paper, he studied the particular case

$$\sigma = 10,\ b = \frac{8}{3},\ r = 28, \tag{4.13}$$

so I will study this case as well with some added random noise. Note that with these parameters, the two fixed points $C^{\pm}$ are stable if $1 < r < 24.7368$. Thus with $r = 28$ we have some strange plots as show in below.

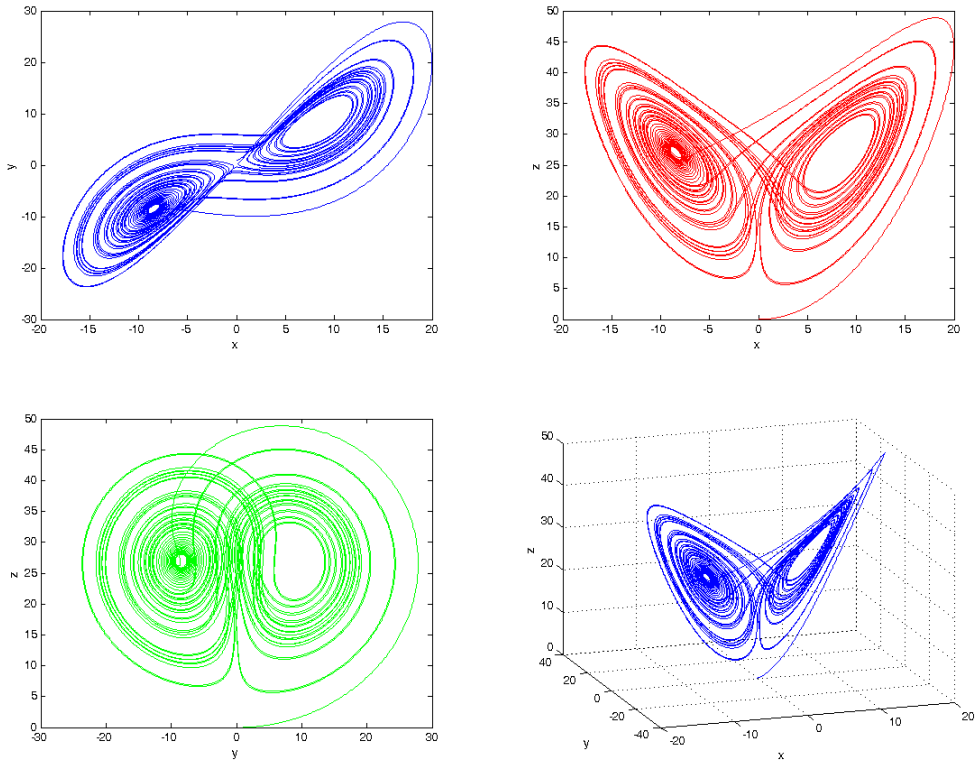

Figure 4.27: Lorenz attractor with $\sigma = 10, b = \frac{8}{3},\ r = 28$ and $x(0) = 0,\ y(0) = 1,\ z(0) = 0$.

The trajectory appears to cross itself repeatedly. We can also look at the plots of the solution on each direction:

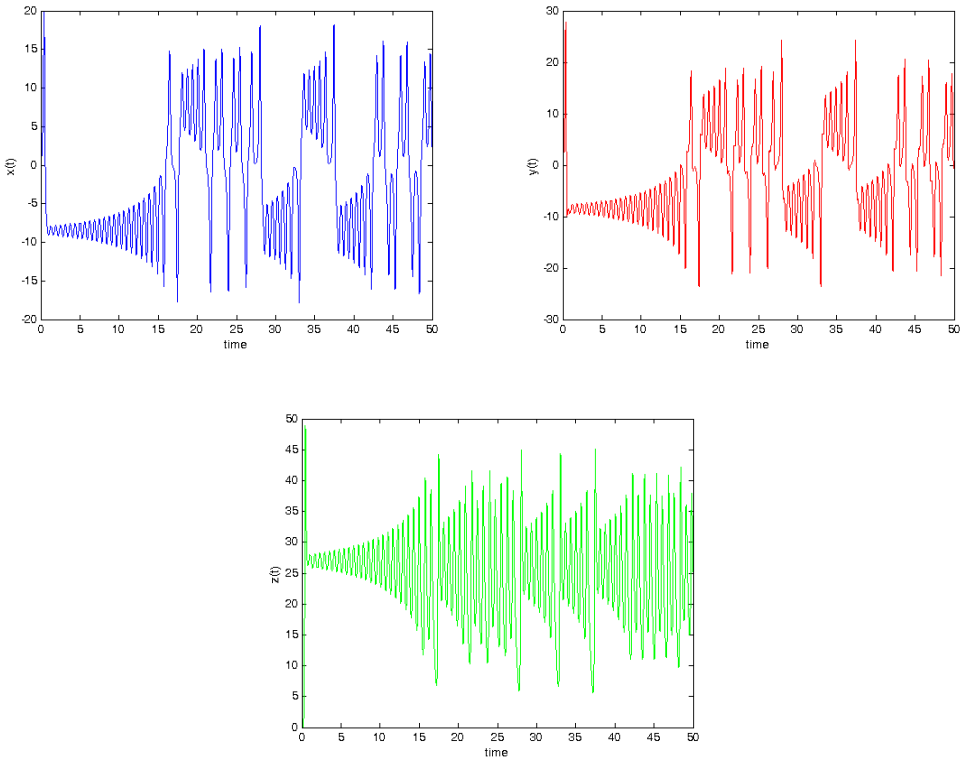

Figure 4.28: Solution for Lorenz equation with $\sigma = 10, b = \frac{8}{3}$, $r = 28$ and $x(0) = 0$, $y(0) = 1$, $z(0) = 0$.

For $x(t)$ and $y(t)$, after an initial transient the solutions gets into an irregular oscillation as $t \to \infty$, and never repeats exactly. Now if we add some random noise, will this irregular oscillation remain the same or become more chaotic?

4.2.2 Stochastic Lorenz equation. By adding a multiplicative noise on $y(t)$ and $z(t)$, we obtain a system of SDEs

$$\begin{aligned} \frac{dX_t}{dt} &= \sigma(Y_t - X_t) \\ \frac{dY_t}{dt} &= rX_t - Y_t - X_t Z_t + \delta_1 Y_t \frac{dW_1}{dt} \\ \frac{dZ_t}{dt} &= X_t Y_t - bZ_t + \delta_2 Z_t \frac{dW_2}{dt} \end{aligned} \tag{4.14}$$

where ϵ_1 and ϵ_2 are positive real numbers, and W_1, W_2 are both one-dimensional Wiener processes.

With same value of parameters and initial values and by adding different amount of random noise, I got some interesting plots. Let us look at the case when $\delta_1 = \delta_2 = 0.2$.

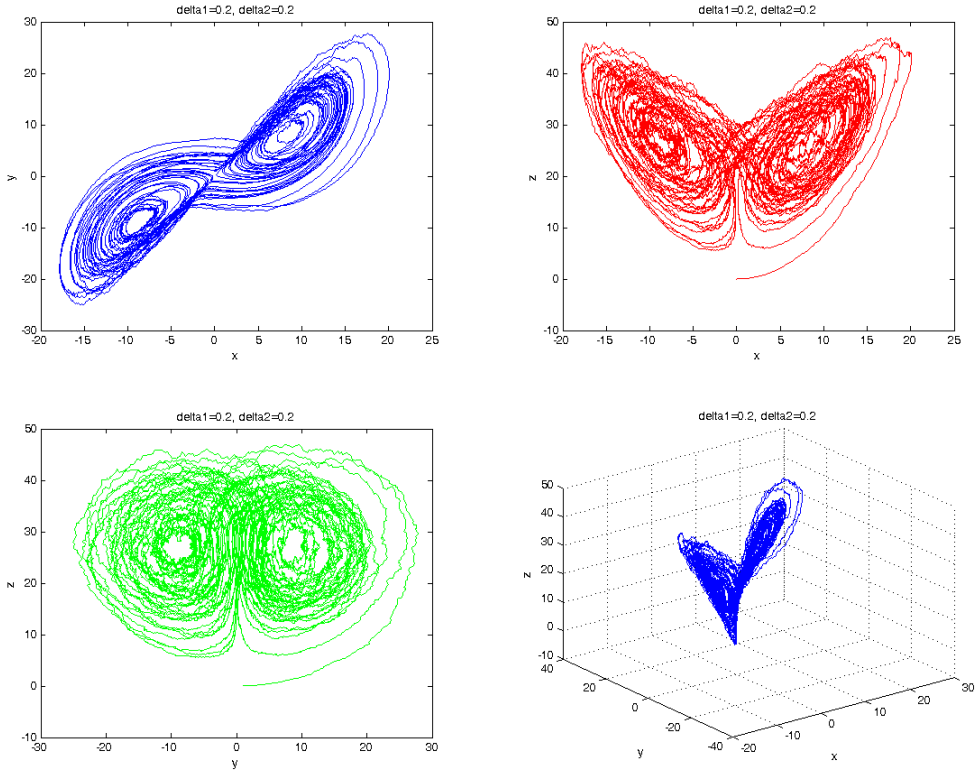

Figure 4.29: Lorenz attractor with $\delta_1 = \delta_2 = 0.2$.

The shapes of the attractor are really close to the deterministic case, trajectories are jagged due to the random noise. Looking at the solution on each direction we obtain figure 4.30:

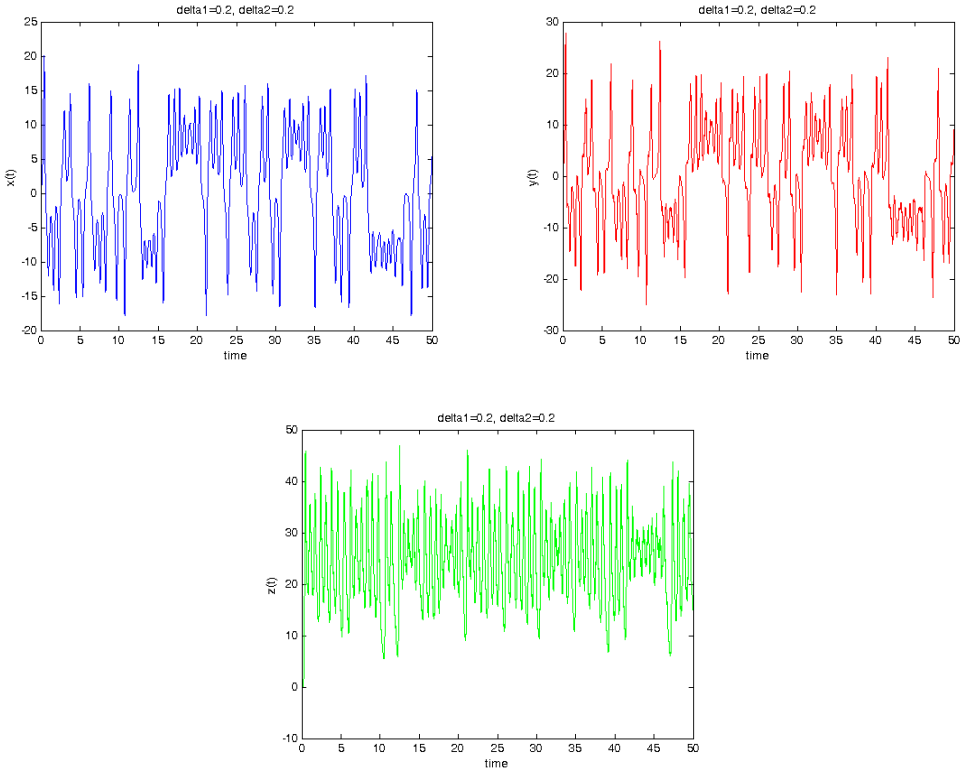

Figure 4.30: Solution for Lorenz equation with $\delta_1 = \delta_2 = 0.2$.

The solution still has irregular oscillations like in the deterministic case, but they oscillate more often. Note also that in the deterministic case, during the oscillations the solution increases from a small value to a large value, but with added random noise, this does not happen.

Now if we increase the strength of the random noise on the y-direction and keep the z-direction same, we get figure 4.31:

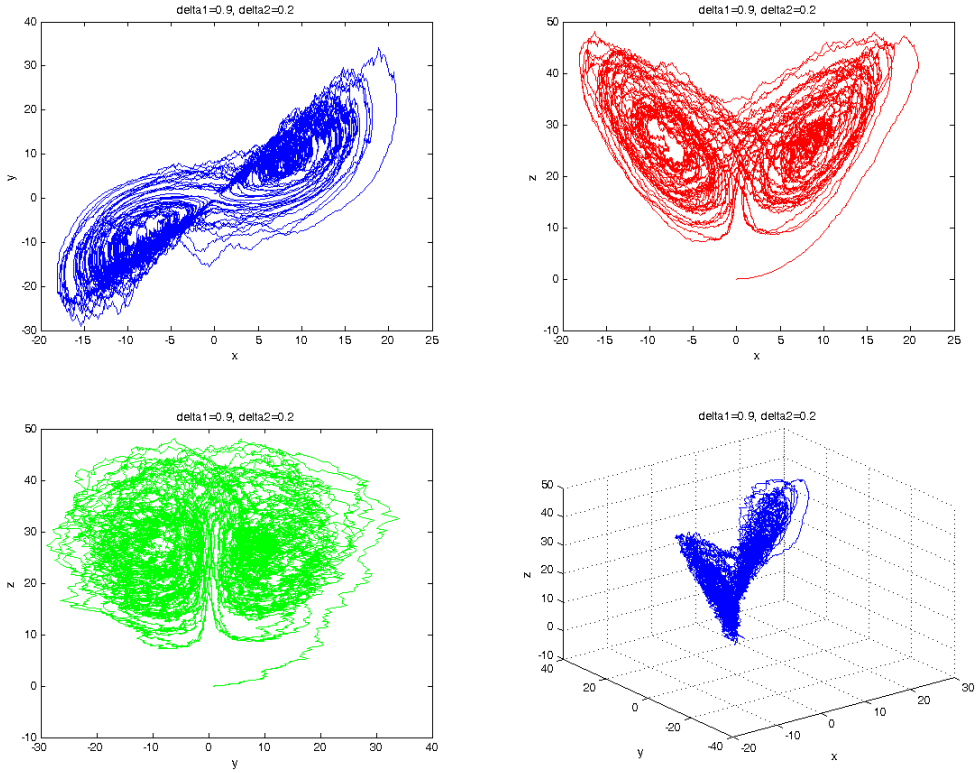

Figure 4.31: Lorenz attractor with $\delta_1 = 0.9,\ \delta_2 = 0.2$.

The main shape of the attractor is still clear, but in the center part of the attractor the trajectories are getting chaotic. The solution on each direction is shown in figure 4.32

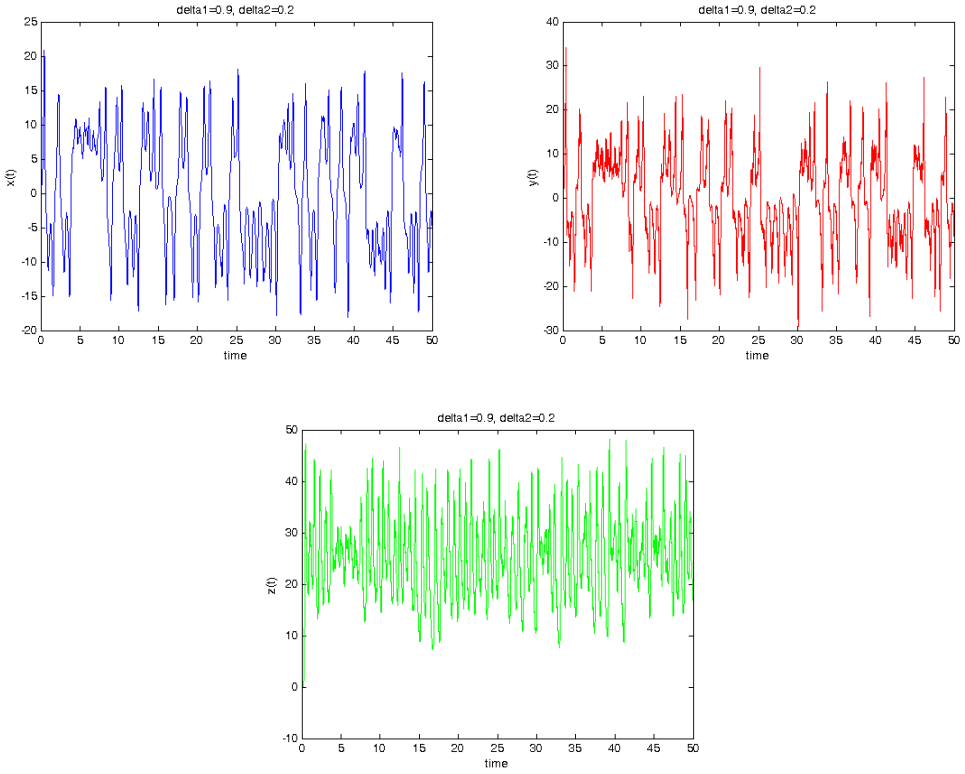

Figure 4.32: Solution for Lorenz equation with$\delta_1 = 0.9,\ \delta_2 = 0.2$.

The solutions look like the previous case but with more frequent oscillation. This is what causes the chaotic behavior at the center of the attractor.

By increasing the strength of the random noise on the z-direction also, I got following:

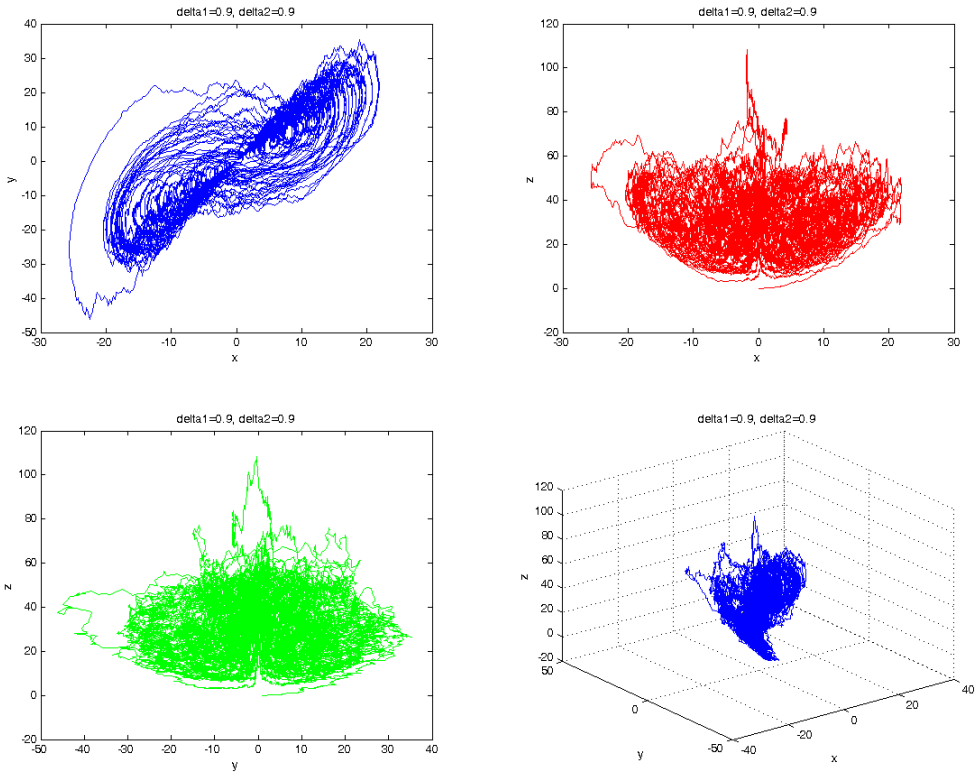

Figure 4.33: Lorenz attractor with $\delta_1 = \delta_2 = 0.9$.

Note that projections of the three-dimensional trajectory on the x-z and y-z plane have a completely different shape from the deterministic case. This is due to the chaotic behavior on each direction as shown in figure 4.34:

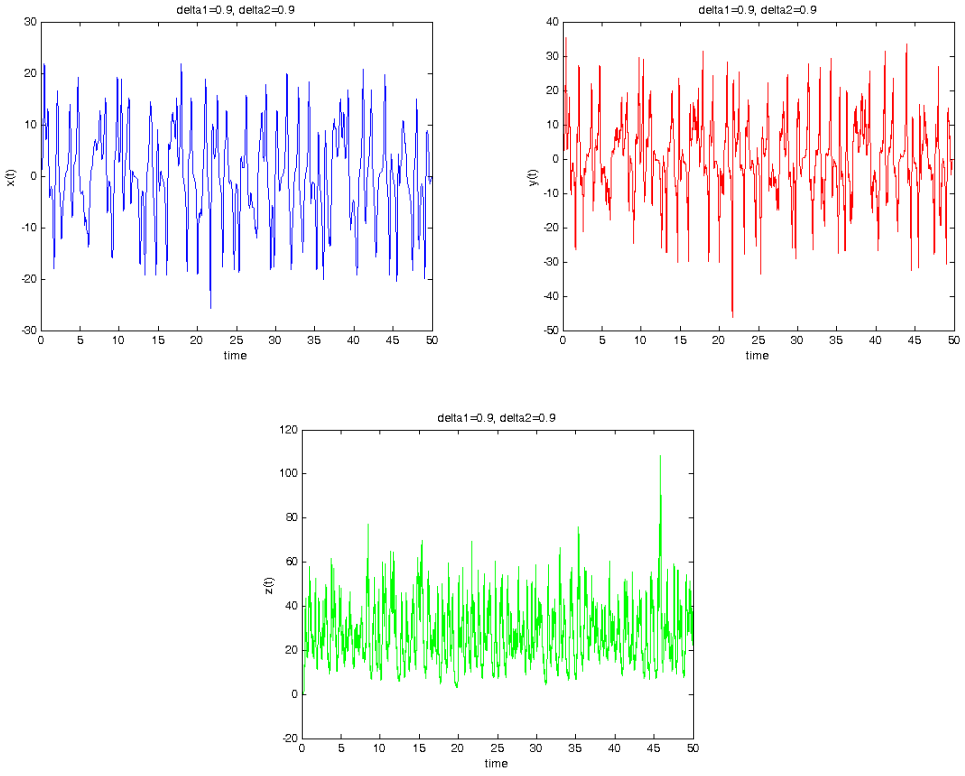

Figure 4.34: Solution for Lorenz equation with $\delta_1 = \delta_2 = 0.9$.

If we compare these plots with figure (4.28), the solution is more jagged and oscillating more quickly.

What if the random noise is really big in one direction and small in another. Taking $\delta_1 = 5$ and $\delta_2 = 0.2$, I predicted that the shape of the attractor is going to be completely different from the deterministic case and that the solution oscillates more drastically on each direction. My result showed that I was only half right. The shape of the attractor is different and the trajectory goes to the point $(0, 0, 0)$ as shown in figure 4.35.

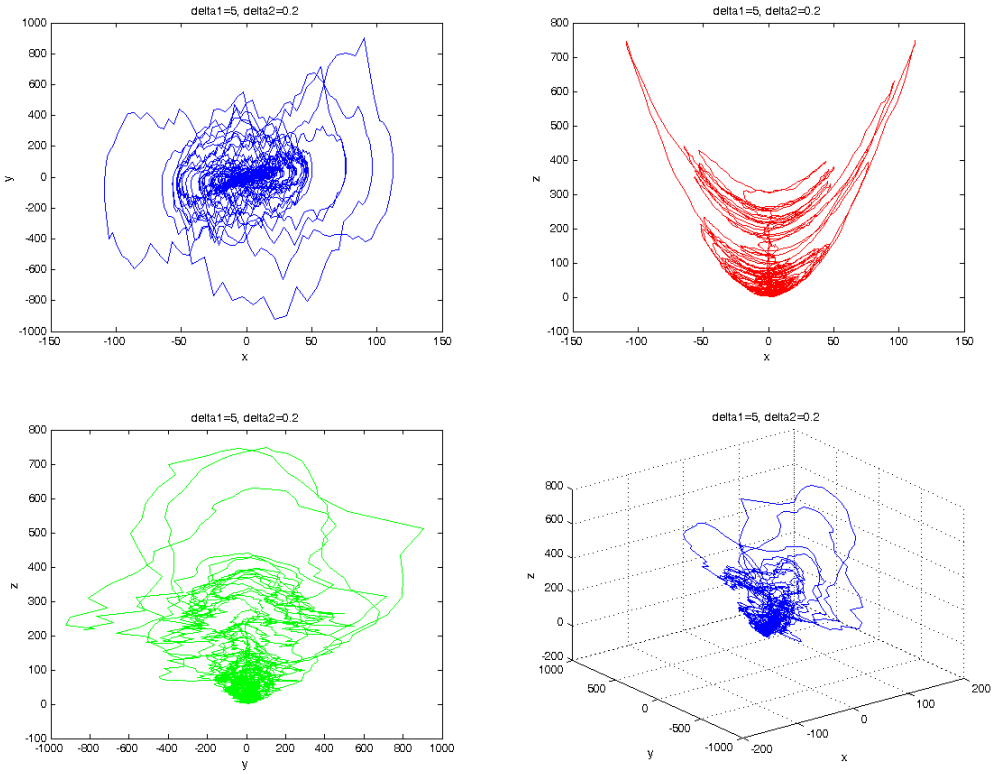

Figure 4.35: Lorenz attractor with $\delta_1 = 5,\ \delta_2 = 0.2$.

However if we look at the solution on each direction, we see that instead of more radical oscillations, it oscillates around the value 0, as shown in figure 4.36:

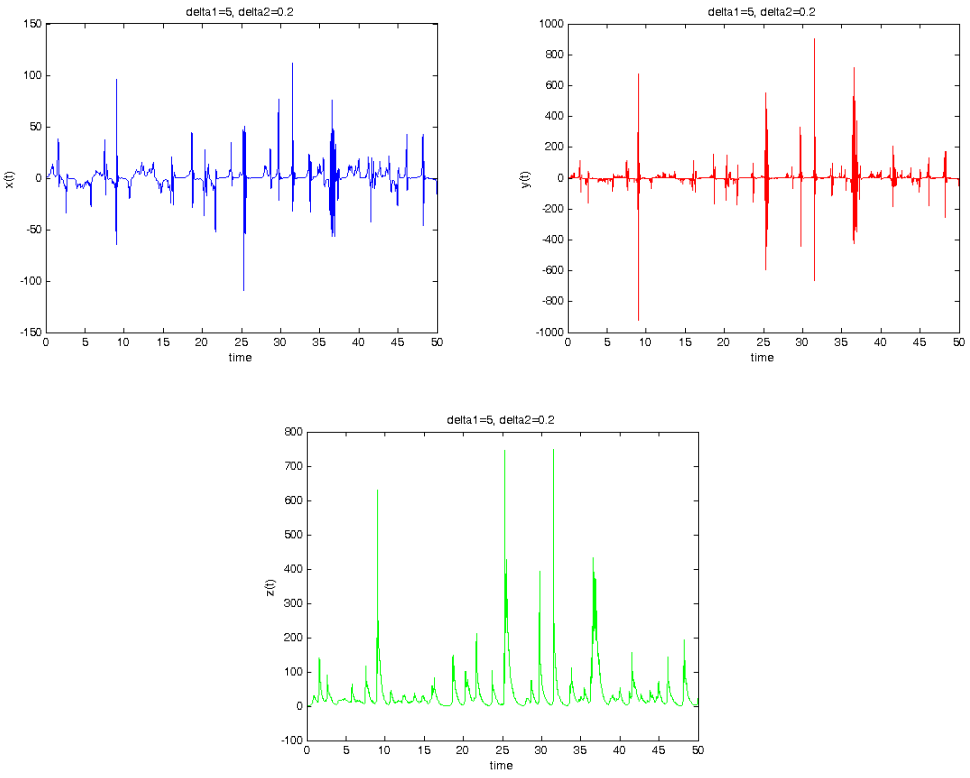

Figure 4.36: Solution for Lorenz equation with and $\delta_1 = 5,\ \delta_2 = 0.2$.

Since the random noise is too strong, it forces the solution to go to zero. I also did other simulations with differential values of δ_1 and δ_2, please see **Appendix B** for the plots.

4.3 Stochastic Pendulum Equation

4.3.1 Pendulum equation. In the absence of damping and external driving, the motion of a pendulum is modeled by

$$\ddot{\theta} + \frac{g}{L}\sin\theta = 0 \tag{4.15}$$

where θ is the angle from the downward vertical, g is the gravity constant, and L is the length of the pendulum. If we let

$$\dot{\theta} = v \tag{4.16}$$

then we have following system of ODEs

$$\begin{cases} \dot{\theta} = v \\ \dot{v} = -\frac{g}{L}\sin\theta \end{cases} \tag{4.17}$$

The fixed points are $(k\pi, 0)$, where $k \in \mathbb{Z}$. Here I only consider the case for $k = 0$ and $k = 1$. Without loss of generality, assume that $L = g$. At $(0, 0)$, the Jacobian matrix is

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \tag{4.18}$$

so the origin is a linear center [3]. At the point $(\pi, 0)$, the Jacobian matrix is

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{4.19}$$

and the eigenvalues are $\lambda = \pm 1$, thus the fixed point $(\pi, 0)$ is a saddle point. Figure 4.37 is the plot of trajectories

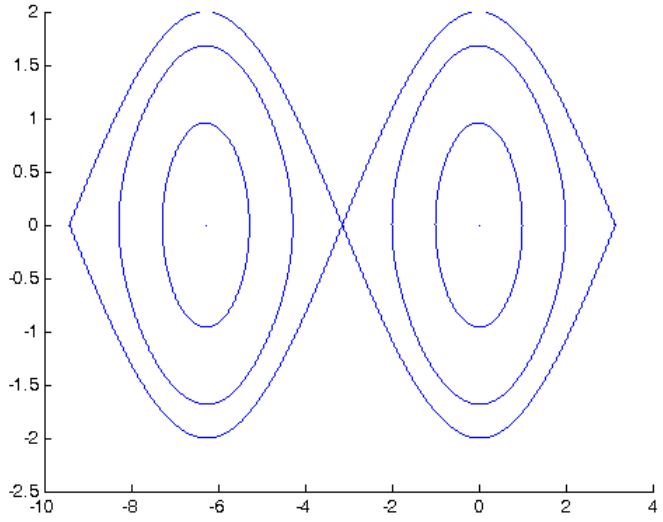

Figure 4.37: Trajectories of deterministic pendulum equation.

4.3.2 Stochastic Pendulum Equation. Now adding a random noise to system (4.17), we get a system of SDEs

$$\begin{cases} d\Theta_t = V_t dt \\ dV_t = -\frac{g}{L}\sin\Theta_t dt + \sigma V_t dW_t., \end{cases} \tag{4.20}$$

where σ is a positive real number, and W_t is a standard Wiener process.

Let us look at the trajectories with small random noise.

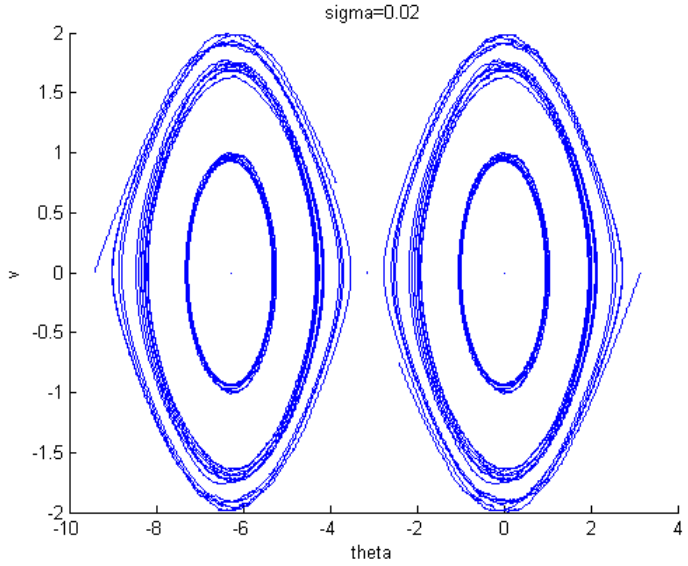

Figure 4.38: Trajectories of stochastic pendulum equation $\sigma = 0.02$.

The first thing I noticed is that the trajectories are not periodic, and they are slightly jagged due to the random noise. Also trajectory touches the point $(\pi, 0)$ in the deterministic case, but does not go close to the same point at all.

Look us just look at one trajectory with initial value $(2, 0)$. Figure 4.39 shows the deterministic case, and figure 4.40 shows the stochastic case with $\sigma = 0.02$.

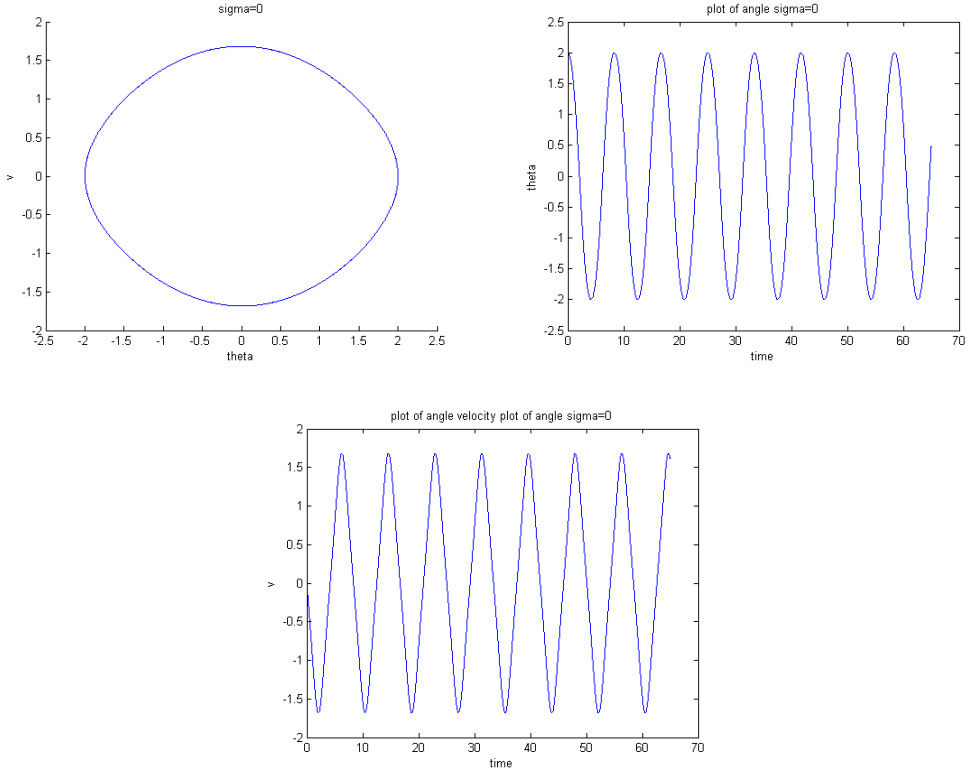

Figure 4.39: Deterministic pendulum equation one trajectory

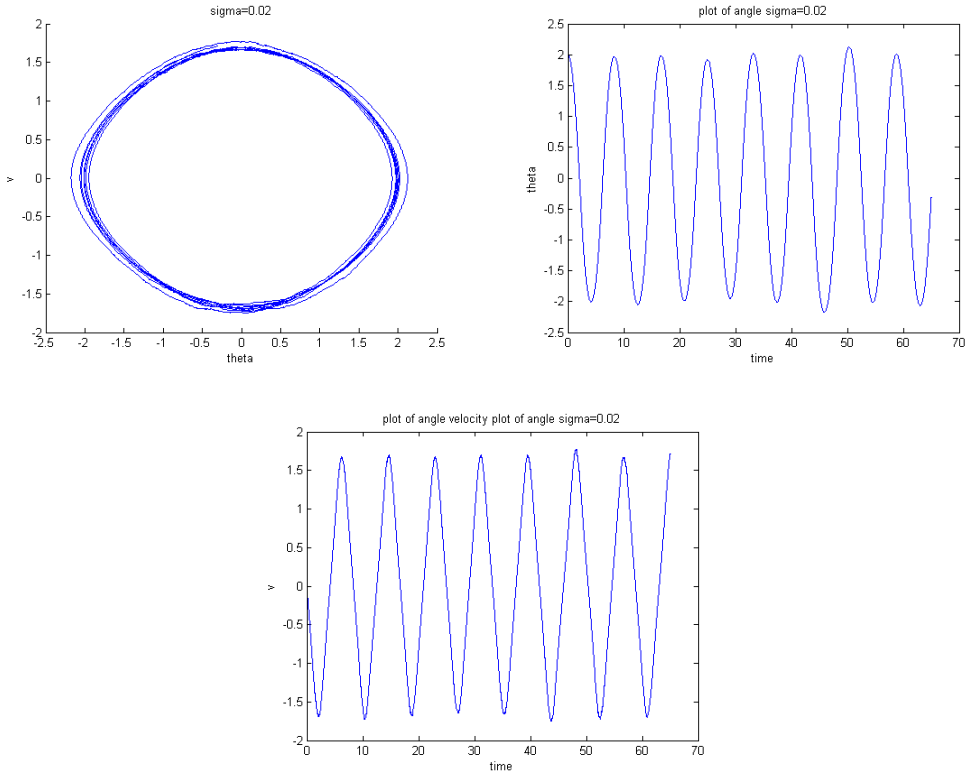

Figure 4.40: Stochastic pendulum equation one trajectory with $\sigma = 0.02$.

Due to the small random noise, the trajectory is not periodic but still has the similar shape and oscillations as in the deterministic case.

By increasing the strength of the random noise, I got the following plots in figure 4.41

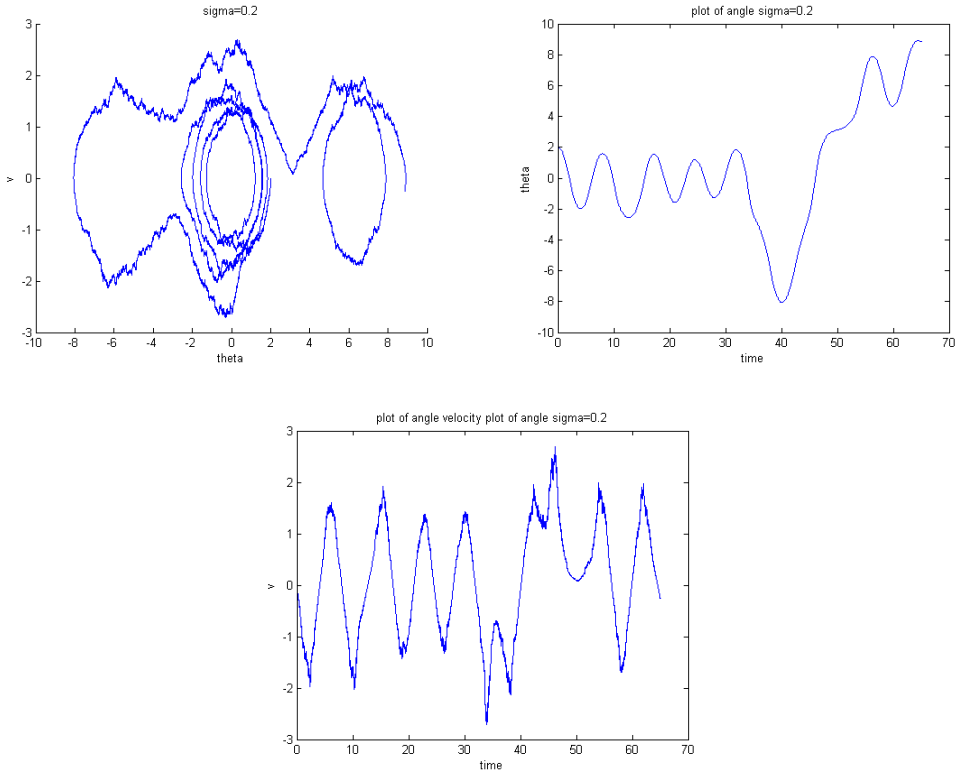

Figure 4.41: Stochastic pendulum equation one trajectory with $\sigma = 0.2$ II.

This trajectory has a totally different shape as in the deterministic case. Note also that the trajectory has two somewhat periodic states, as apposed to the deterministic case where there is only one periodic state. Now I make the random noise even stronger, the result is in figure 4.42:

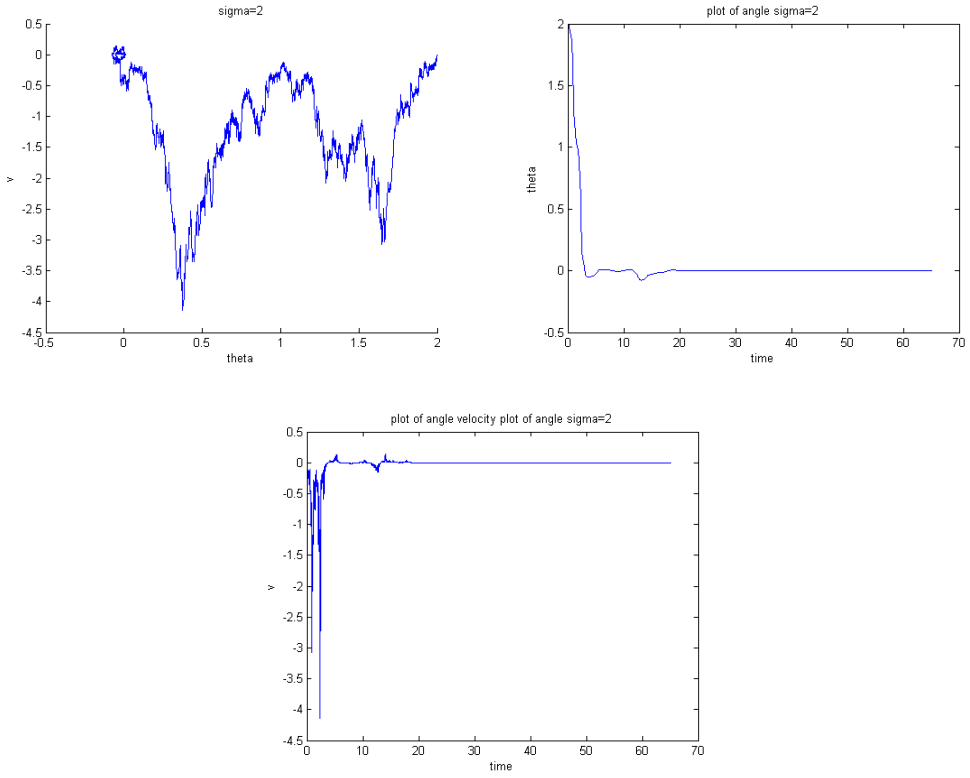

Figure 4.42: Stochastic pendulum equation one trajectory with $\sigma = 2$.

Due to the strength of the noise, the trajectory no longer oscillates. As time increases, it converges to a point near the fixed point $(0, 0)$. this can also be seen from the plot for the angle θ and the angularly velocity v. Both values go to 0 and stay there.

More plots with different strength of random noise are located in **Appendix C**.

Now let us look at the damped pendulum equation

$$\ddot{\theta} + \delta\dot{\theta} + \sin\theta = 0 \tag{4.21}$$

which yields the system

$$\begin{cases} \dot{\theta} = v \\ \dot{v} = -\sin\theta - \delta v \end{cases} \tag{4.22}$$

with the damping strength $\delta > 0$. Supposing $\delta = 0.1$, the plot shows that "the centers become spirals and saddles remain saddle".[3]

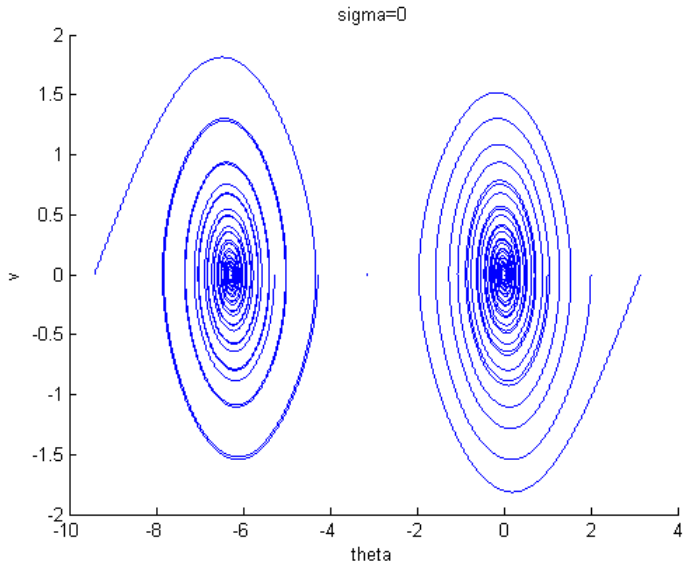

Figure 4.43: Trajectories of damped deterministic pendulum equation with $\delta = 0.1$.

Now adding a small random noise $\sigma = 0.02$, we get the plot in figure 4.44

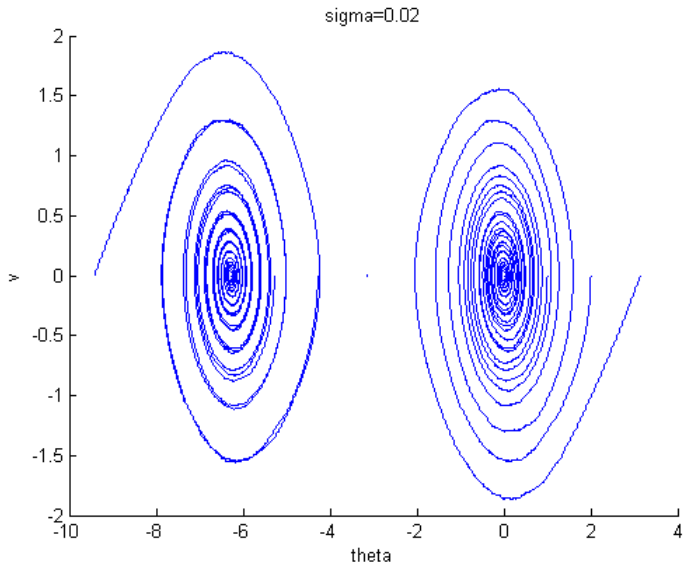

Figure 4.44: Trajectories of damped stochastic pendulum equation with $\delta = 0.1$ and $\sigma = 0.02$.

The only apparent difference is the trajectories are a little jagged due to the random noise. When the strength of the random noise is increased, we the get result in figure 4.45:

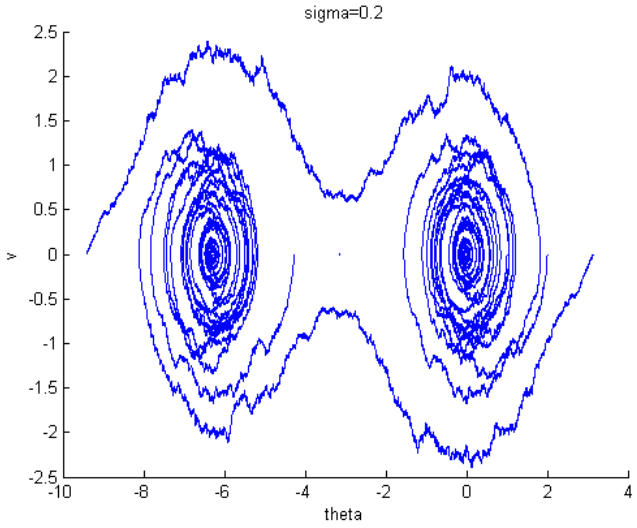

Figure 4.45: Trajectories of damped stochastic pendulum equation with $\delta = 0.1$ and $\sigma = 0.2$.

This case is very different from the deterministic case. First, around the spiral points the trajectories become chaotic. Also note that the trajectory which starts from the point $(-9.4248, 0)$ spirals to the point (0,0) instead of $(-2\pi, 0)$. Similarly, the trajectory starting from the point $(3.1215926, 0)$ spirals to the point $(-2\pi, 0)$ instead of $(0, 0)$.

The strong random noise case is shown in figure 4.46:

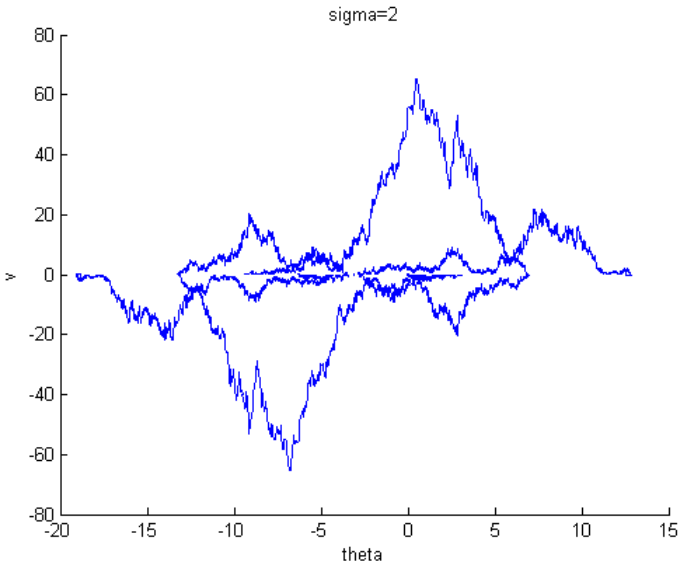

Figure 4.46: Trajectories of damped stochastic pendulum equation with $\delta = 0.1$ and $\sigma = 2$.

The structure here is different from the deterministic case. After some oscillations, all trajectories converge to the line $v = 0$. From the plot of the angle and the angular velocity, it is clear to see that both of them converge to 0 as time goes to infinity. The plots are listed in figure 4.47.

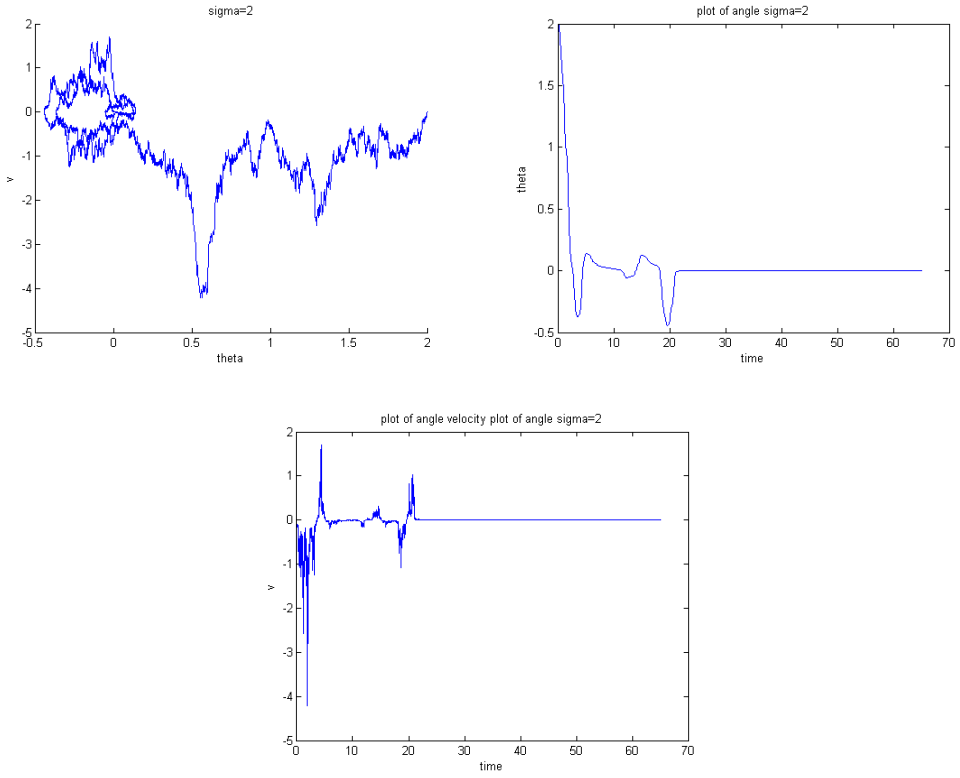

Figure 4.47: Damped stochastic pendulum equation one trajectory with $\sigma = 2$.

More plots with varying strengths of random noise for the damped stochastic pendulum equation can be found in **Appendix C**.

4.4 Pricing Spread Option

Options are traded both on exchanges and in the over-the-counter market. There are two basic types of options. A *call option* gives the holder the right to buy the underlying asset by a specified date for a specified price. A *put option* gives the holder the right to sell the underlying asset by a specified date for a specified price. A spread option is a type of option where the payoff is dependent on the difference between two underlying assets. For more details on options and spread options, please look at [5].

A spread option is entered by buying and selling equal number of options of the same class on the same underlying security but with different strike prices or expiration dates. Let $S_1(t)$ and $S_2(t)$ be the price of the two underlying assets of a spread option, then the price of a spread call option with expiration date T and strike price K at time t is

$$P_{call} = e^{-r(T-t)} \times \max(S_2(T) - S_1(T) - K, 0) \tag{4.23}$$

and the price of a spread put option with expiration date T and strike price K at time t is

$$P_{put} = e^{-r(T-t)} \times \max(K - (S_2(T) - S_1(T)), 0) \tag{4.24}$$

where r is the interest rate.

Spread option are the basic building blocks of many options trading strategies. How to price a spread option has been an popular questions for decades. If we know the price of two underlying assets at time t, then we can use formula (4.23) or formula (4.24) to price a spread option. Thus there are many SDE models to price spread options by modeling the price of the underlying assets.

The model I looked at is following [6]:

$$\begin{cases} dS_{1t} = (\mu_1 - \delta_1)S_{1t}dt + \sigma_1 S_{1t}dW_{1t} \\ dS_{2t} = (\mu_2 - \delta_2)S_{2t}dt + \rho\sigma_2 S_{2t}dW_{1t} + \sqrt{1-\rho^2}\sigma_2 S_{2t}dW_{2t} \end{cases} \tag{4.25}$$

where W_{1t} and W_{2t} are standard Wiener processes. The parameters μ_1 and μ_2 are the instantaneous expected rates of return on the two assets, δ_1 and δ_2 are the instantaneous dividend yields,

σ_1 and σ_2 are the instantaneous standard deviations of the rates of return, and ρ is the correlation coefficient [6].

In order to use this model to price the spread option, I took a certain amount of paths of the two Wiener processes and used the numerical schemes mentioned above to simulate the price of the two underlying assets for each random path. Then I took the average of these prices. In the end I used formula (4.23) or formula (4.24) (depending on the type spread option). The plots for 5000 paths and 10000 paths are shown in figures 4.48, and 4.49 respectively.

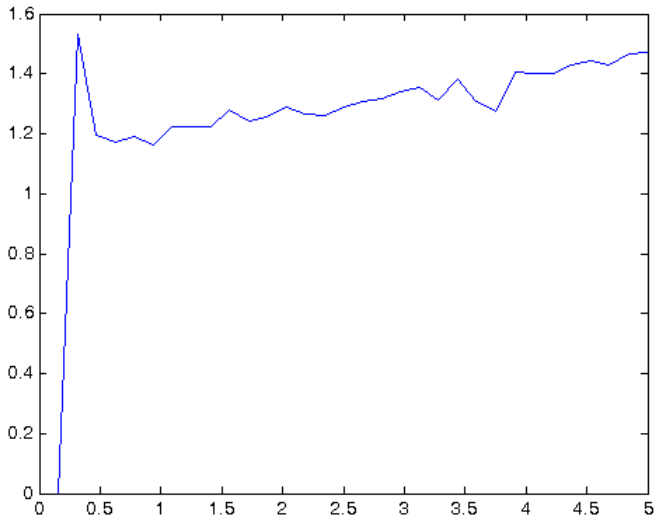

Figure 4.48: Price of a spread option, average over 5000 paths.

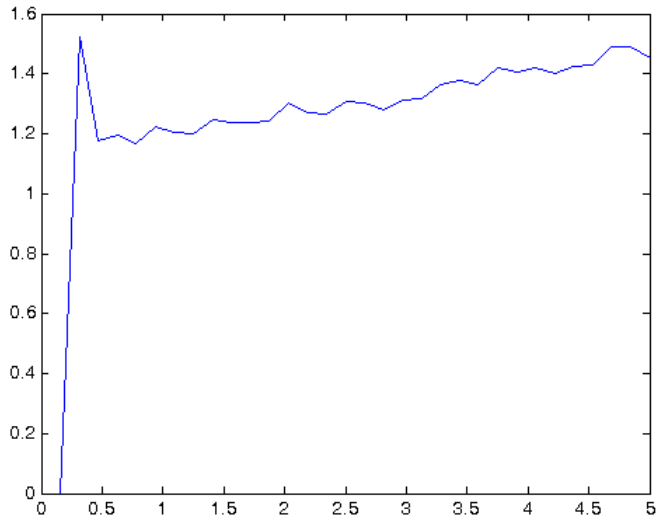

Figure 4.49: Price of a spread option, average over 10000 paths.

The question now is how close are these plots to real life. Unfortunately, this question cannot be answered here. The numerical schemes mentioned before guarantee the simulated solution is close to the real solution of SDE(s). So if the model for the price of the spread option is good, then we can say that the result could be close to real life. On the other hand, if the model is bad, then the result is most likely inaccurate.

Appendix A. Plots of Stochastic Duffing Equation

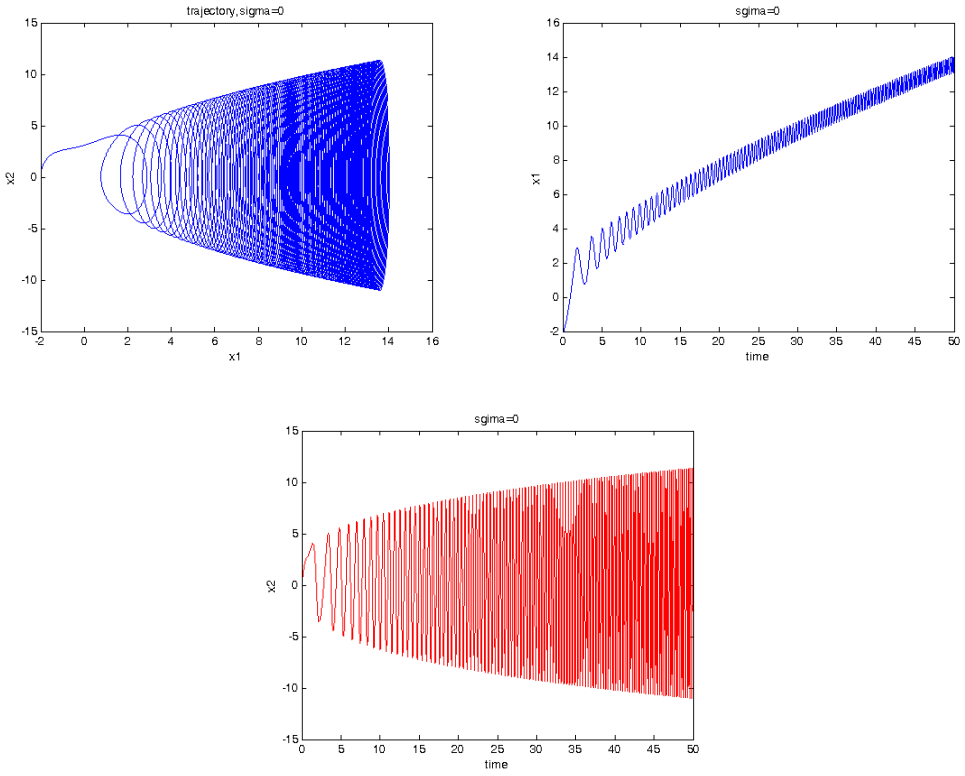

Figure A.1: Duffing no damping forced with $F(t) = t^2 + 1$.

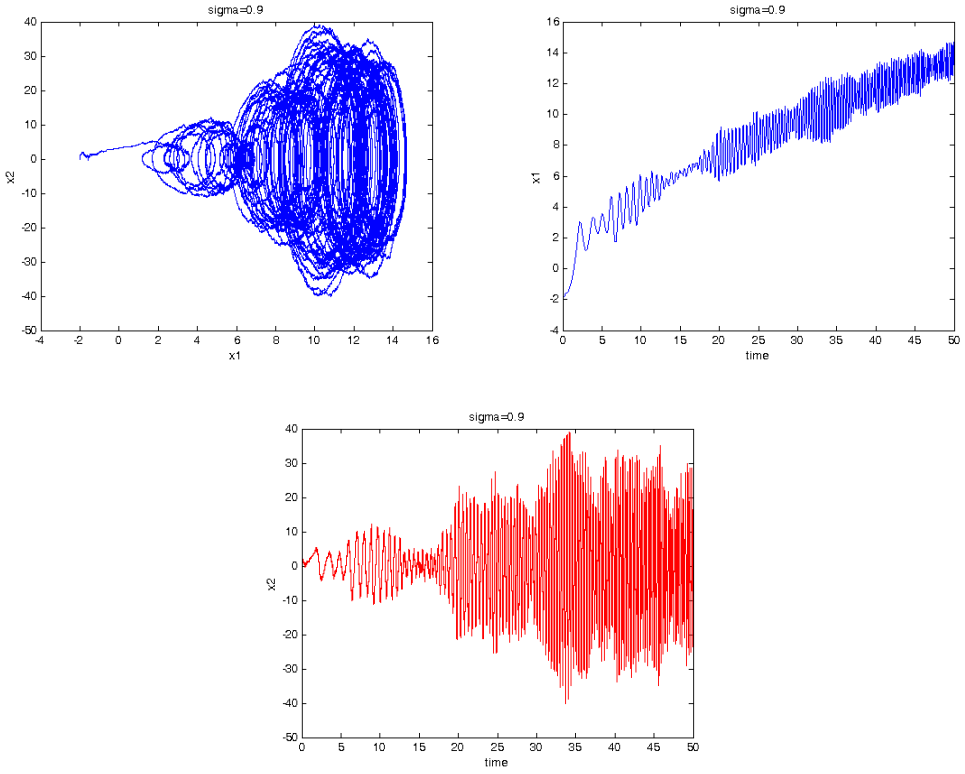

Figure A.2: Duffing no damping forced with $F(t) = t^2 + 1$ and $\sigma = 0.9$.

Appendix B. Plots of Stochastic Lorenz Equation

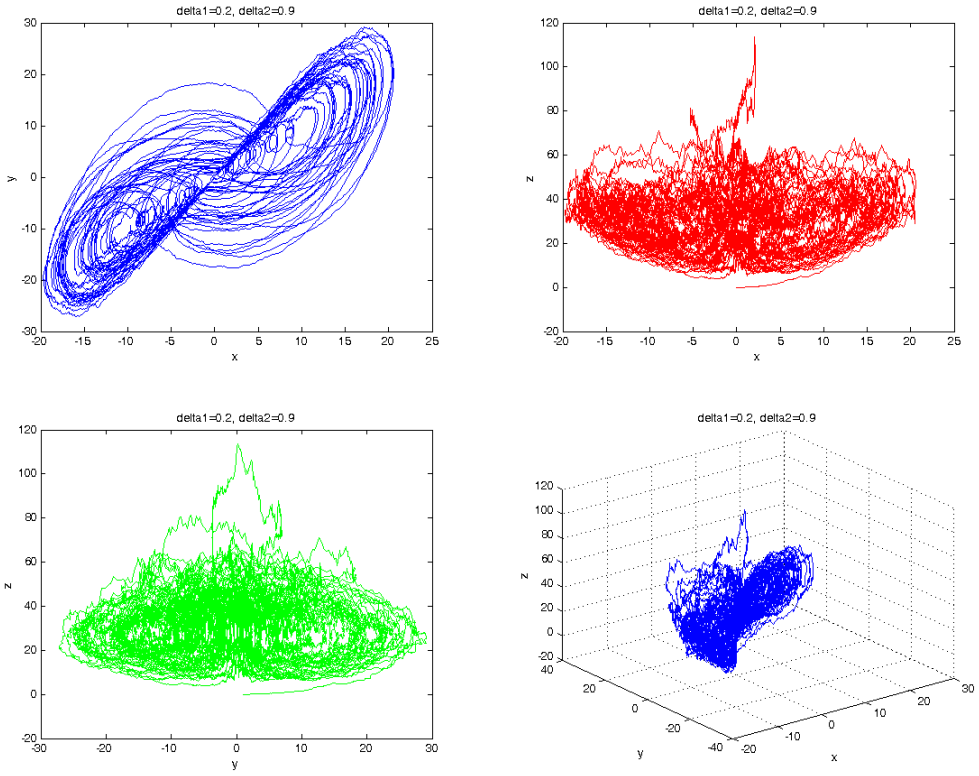

Figure B.1: Lorenz attractor with $\delta_1 = 0.2,\ \delta_2 = 0.9$.

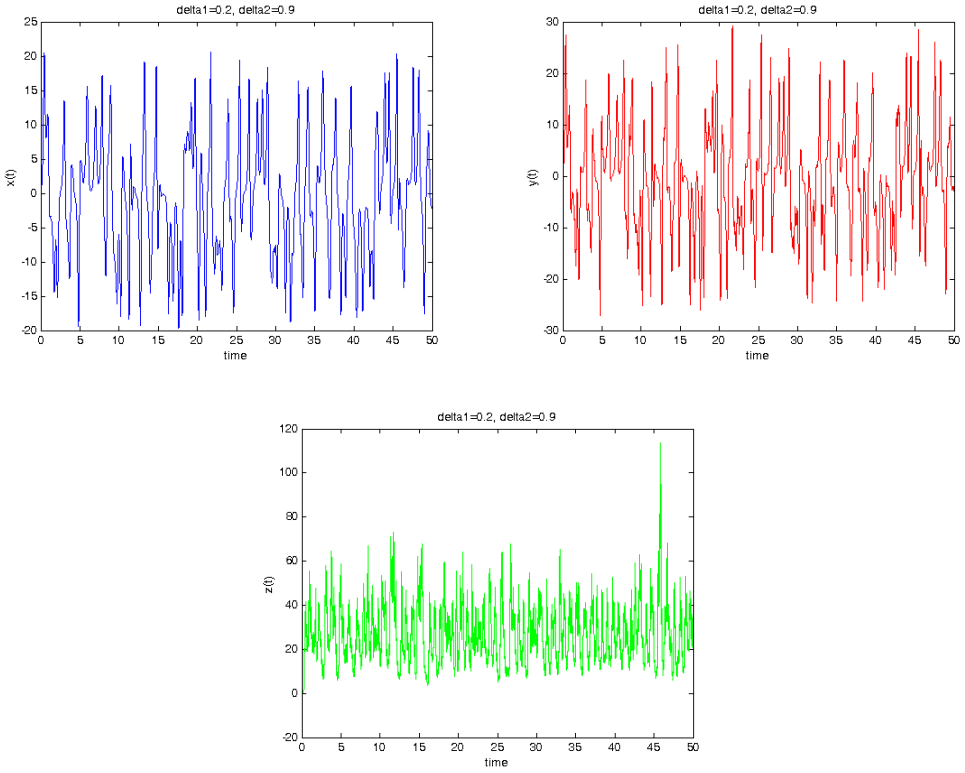

Figure B.2: Solution for Lorenz equation with $\delta_1 = 0.2,\ \delta_2 = 0.9$.

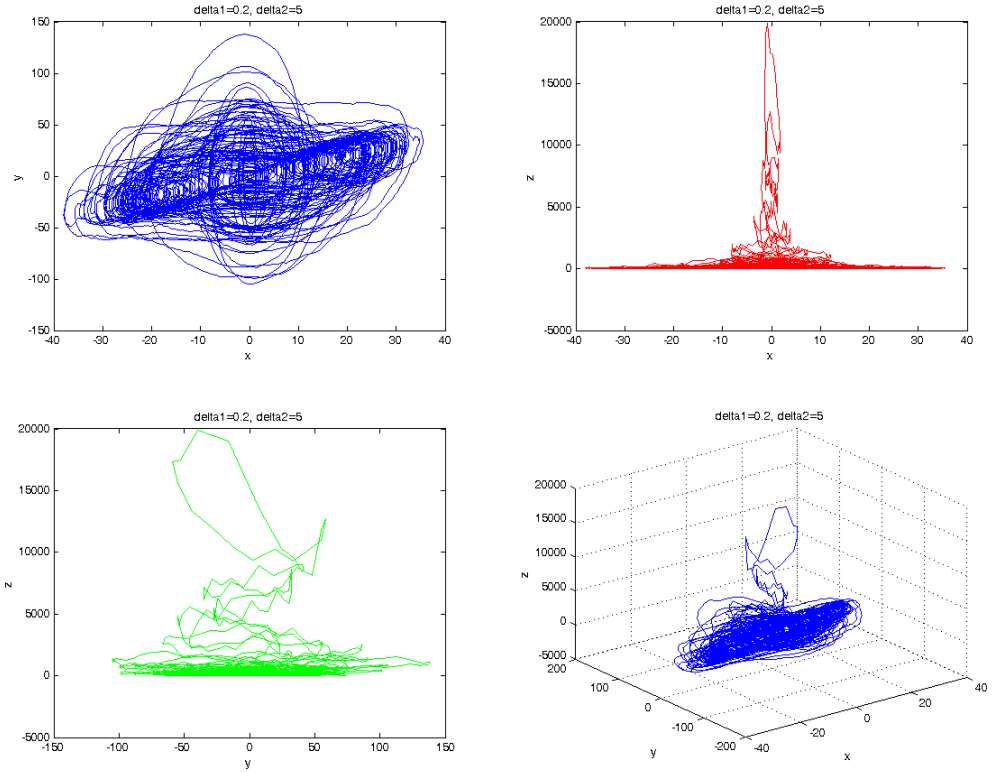

Figure B.3: Lorenz attractor with$\delta_1 = 0.2,\ \delta_2 = 5$.

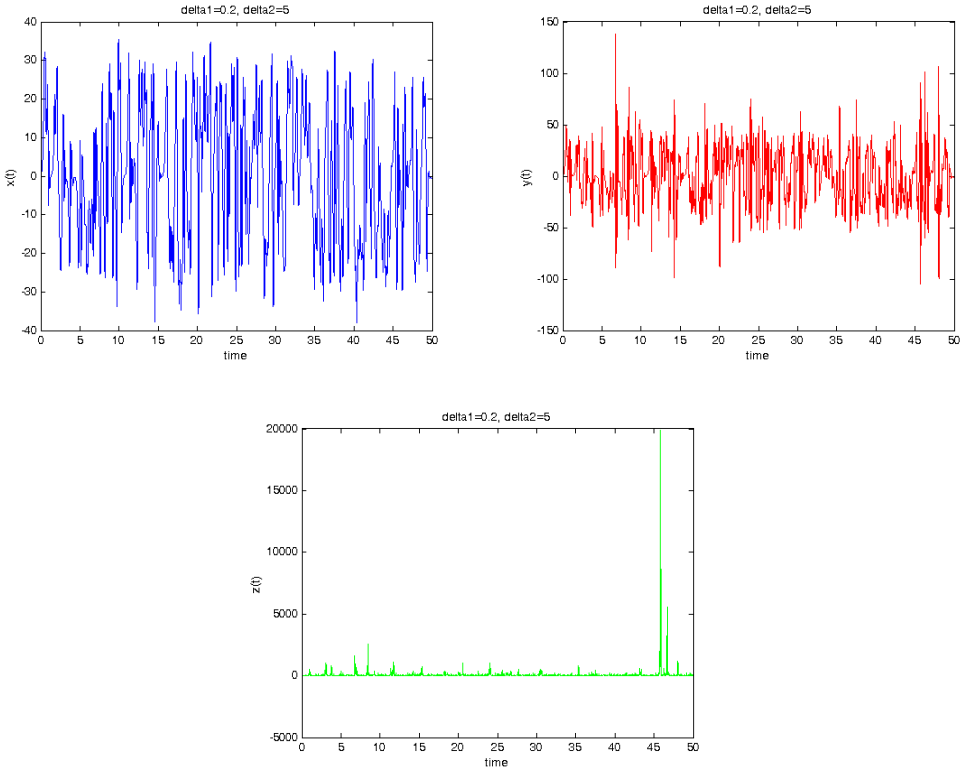

Figure B.4: Solution for Lorenz equation with $\delta_1 = 0.2,\ \delta_2 = 5$.

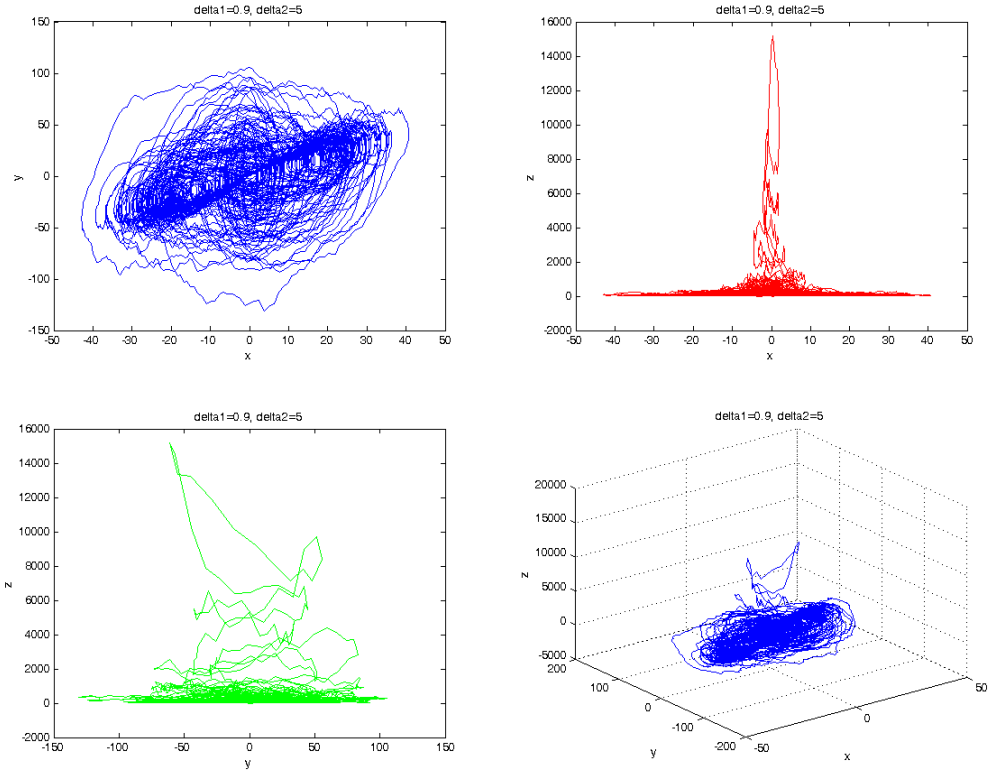

Figure B.5: Lorenz attractor with $\delta_1 = 0.9,\ \delta_2 = 5$.

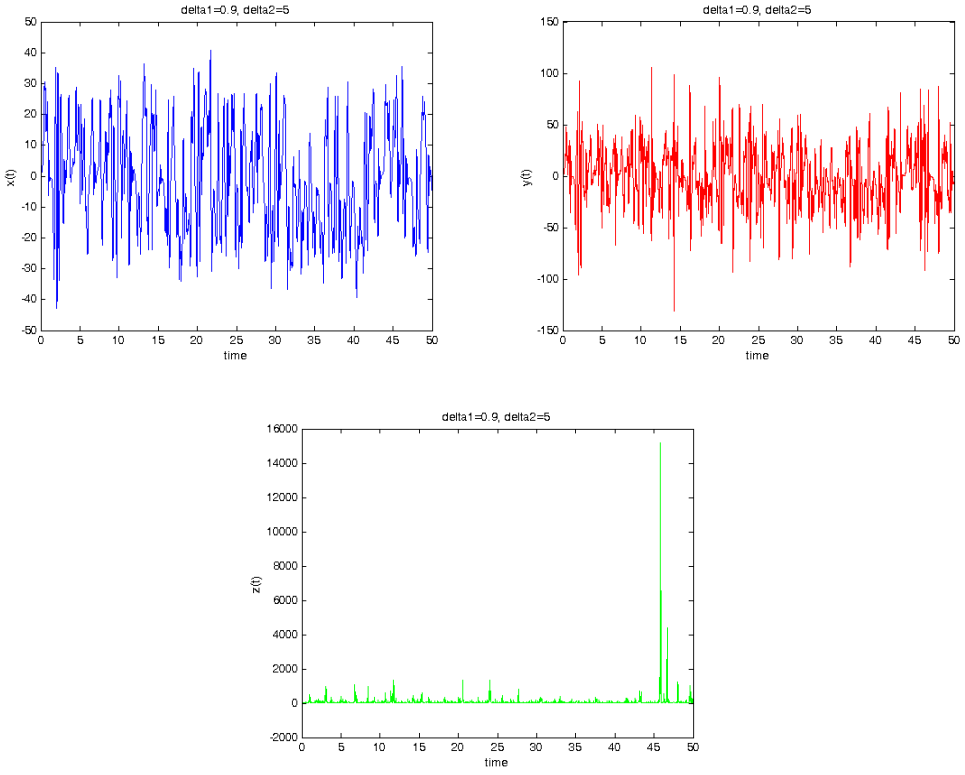

Figure B.6: Solution for Lorenz equation with $\delta_1 = 0.9,\ \delta_2 = 5$.

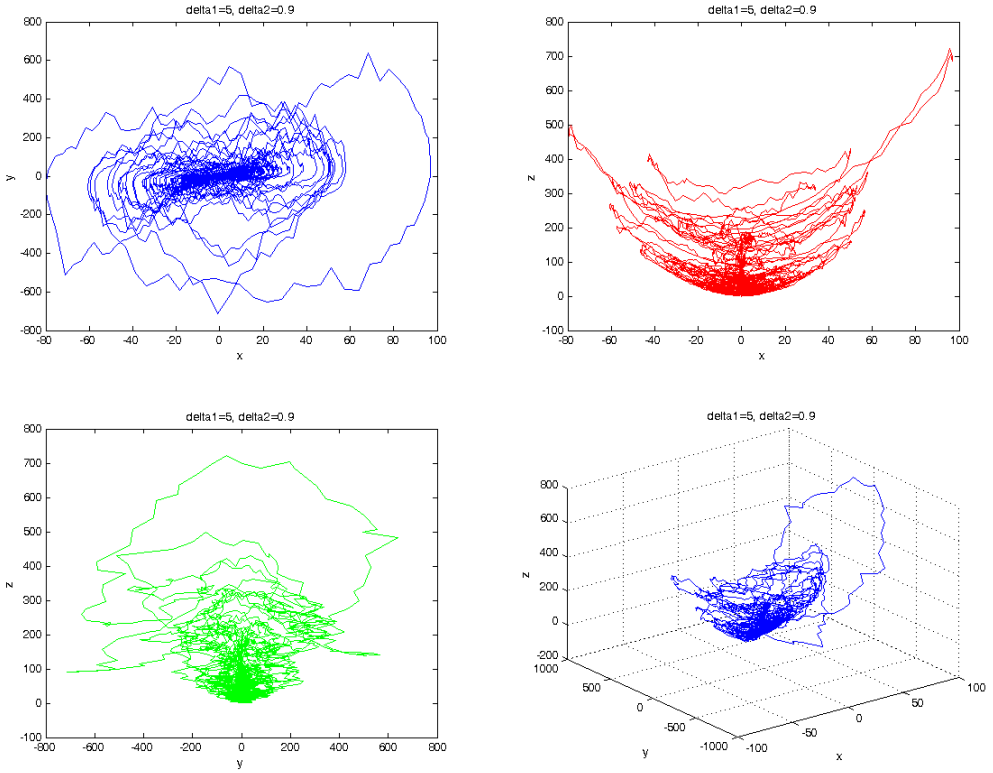

Figure B.7: Lorenz attractor with $\delta_1 = 5,\ \delta_2 = 0.9$.

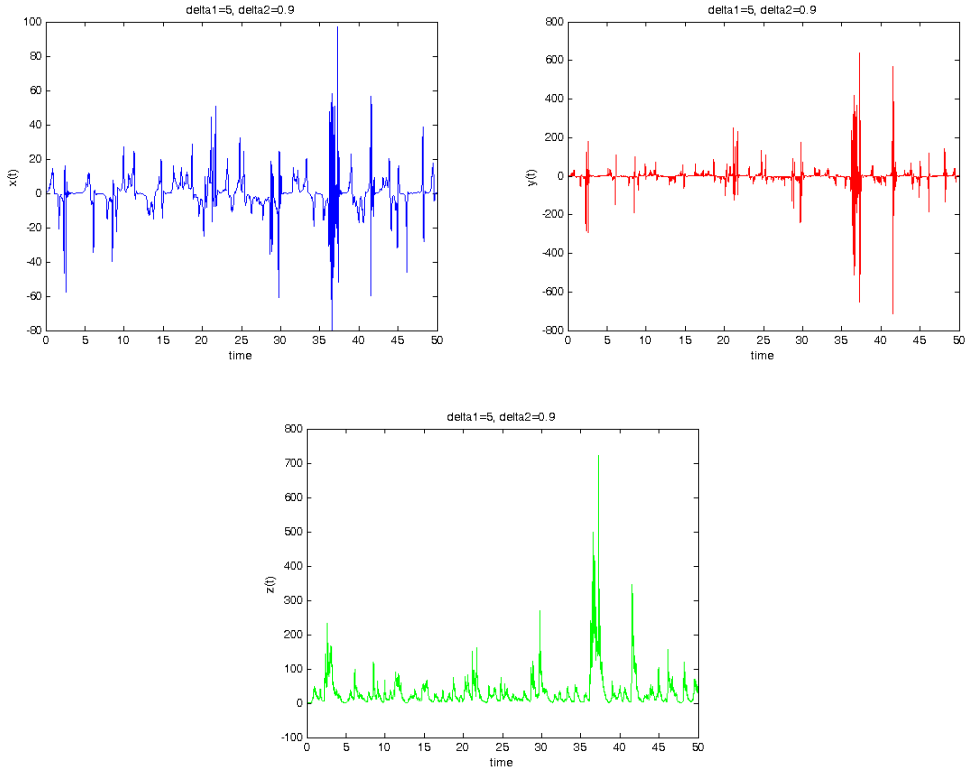

Figure B.8: Solution for Lorenz equation with $\delta_1 = 5,\ \delta_2 = 0.9$.

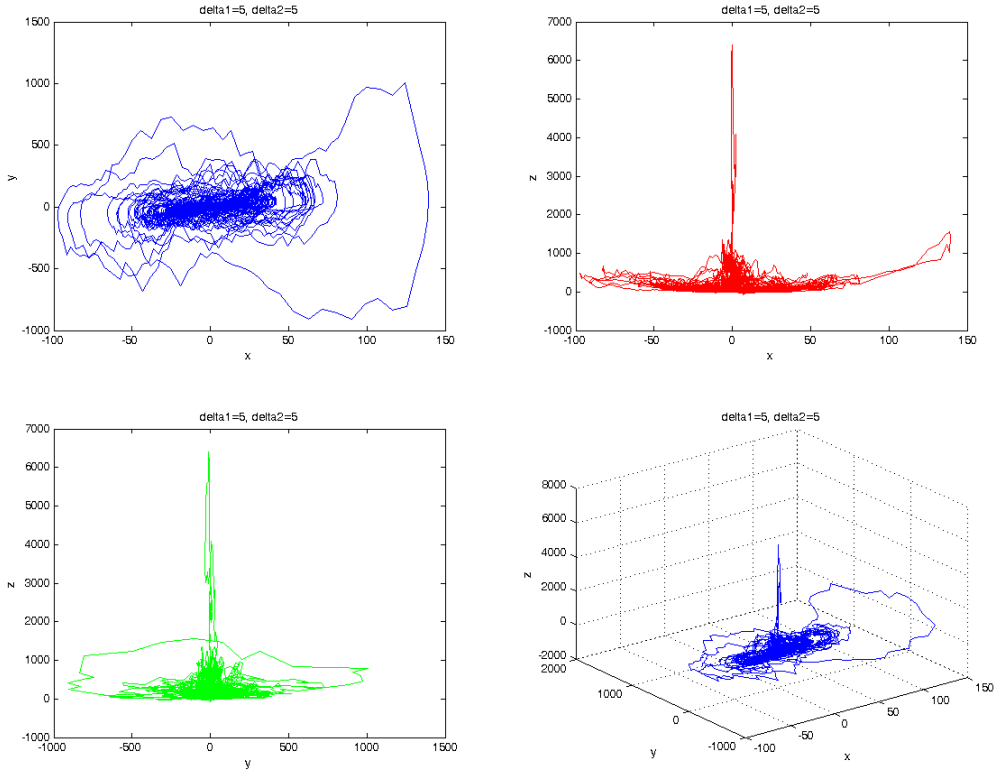

Figure B.9: Lorenz attractor with $\delta_1 = \delta_2 = 5$.

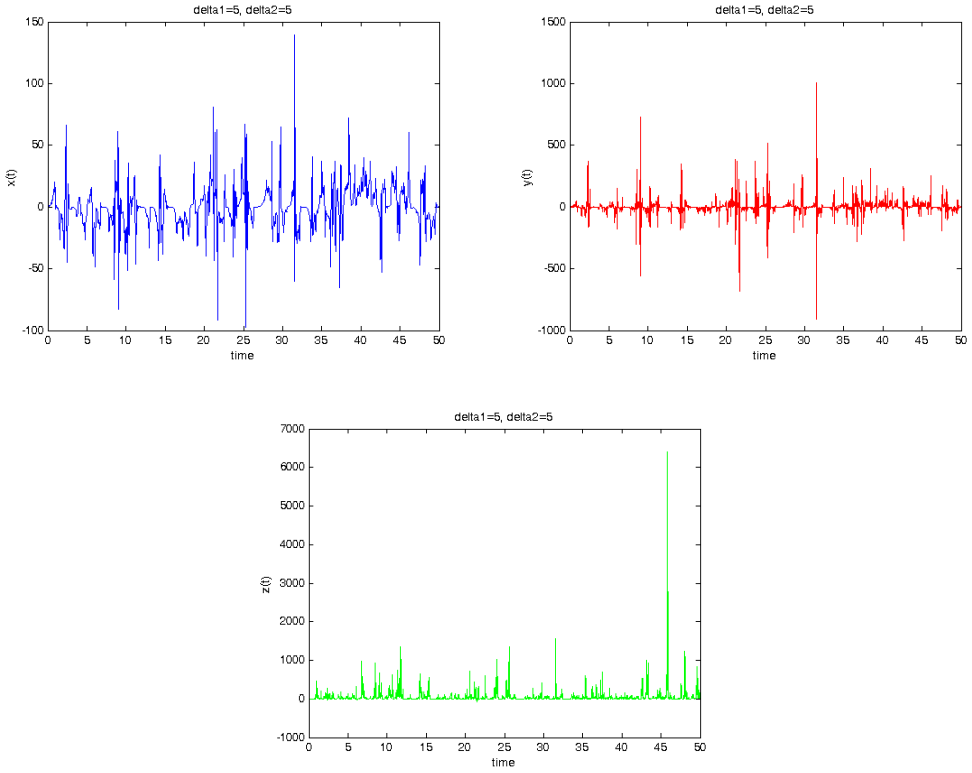

Figure B.10: Solution for Lorenz equation with $\delta_1 = \delta_2 = 5$.

Appendix C. Plots of Stochastic Pendulum Equation

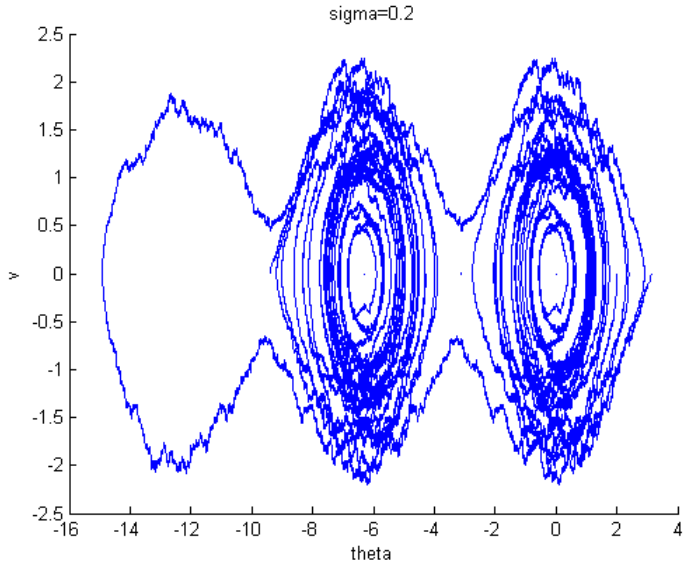

Figure C.1: Trajectories of stochastic pendulum equation $\sigma = 0.2$.

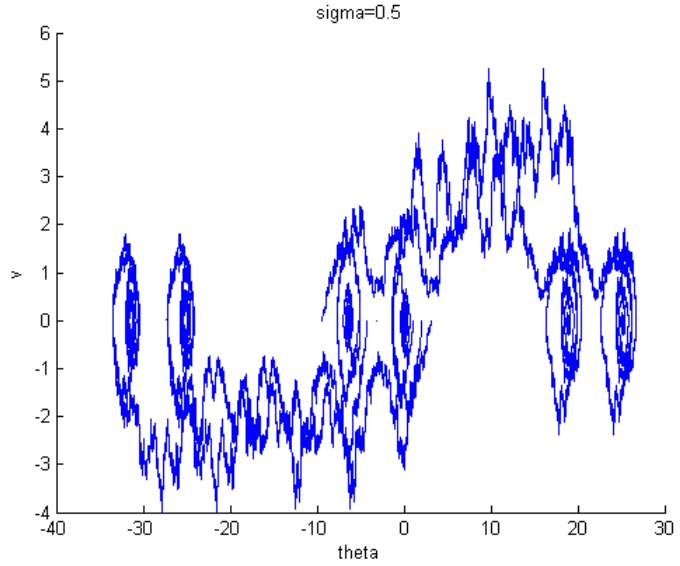

Figure C.2: Trajectories of stochastic pendulum equation $\sigma = 0.5$.

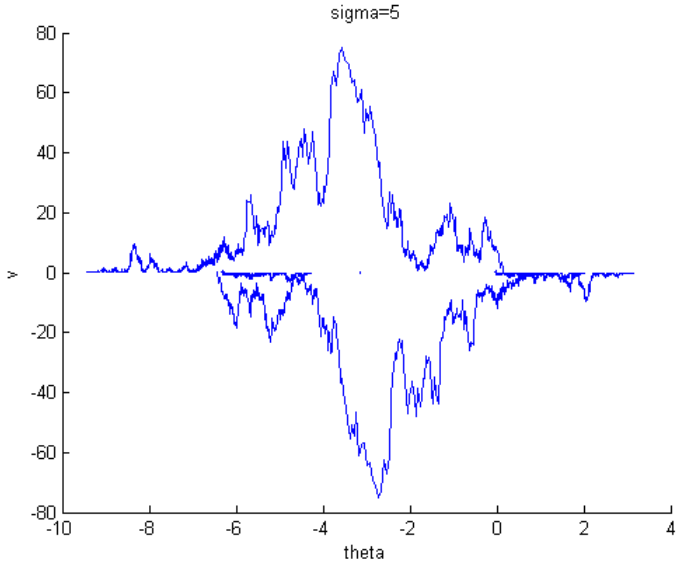

Figure C.3: Trajectories of stochastic pendulum equation $\sigma = 5$.

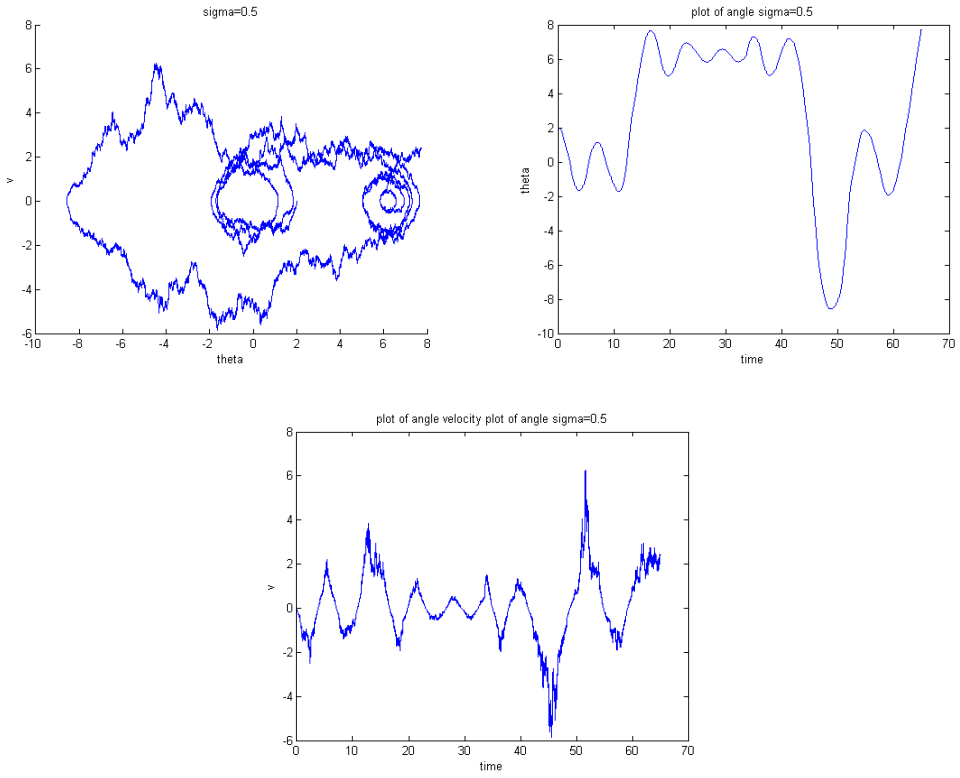

Figure C.4: Stochastic pendulum equation one trajectory with $\sigma = 0.5$.

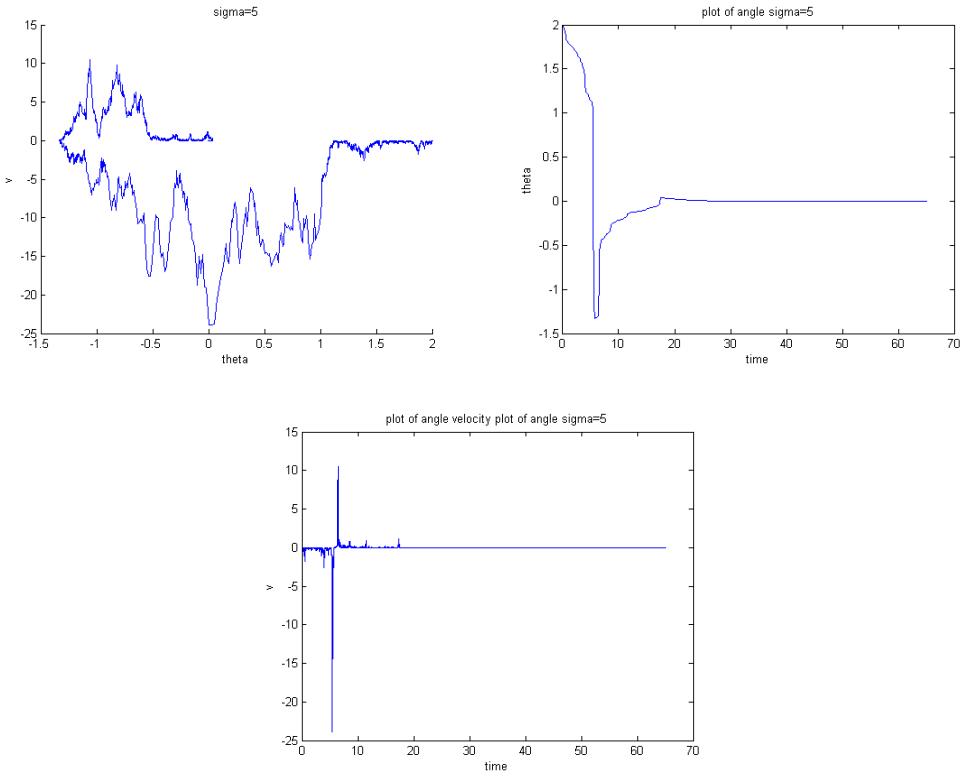

Figure C.5: Stochastic pendulum equation one trajectory with $\sigma = 5$.

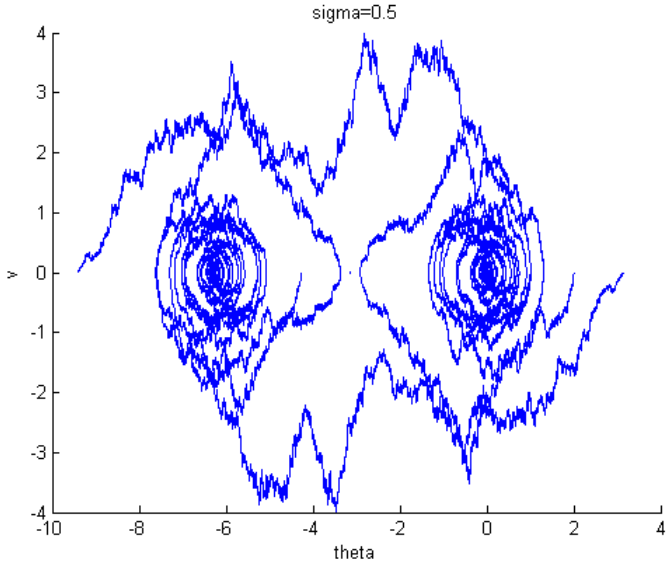

Figure C.6: Trajectories of damped stochastic pendulum equation with $\delta = 0.1$ and $\sigma = 0.5$.

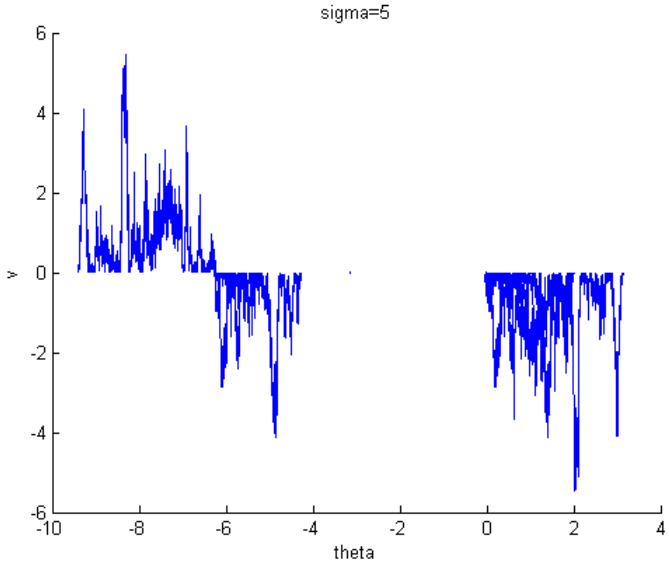

Figure C.7: Trajectories of damped stochastic pendulum equation with $\delta = 0.1$ and $\sigma = 5$.

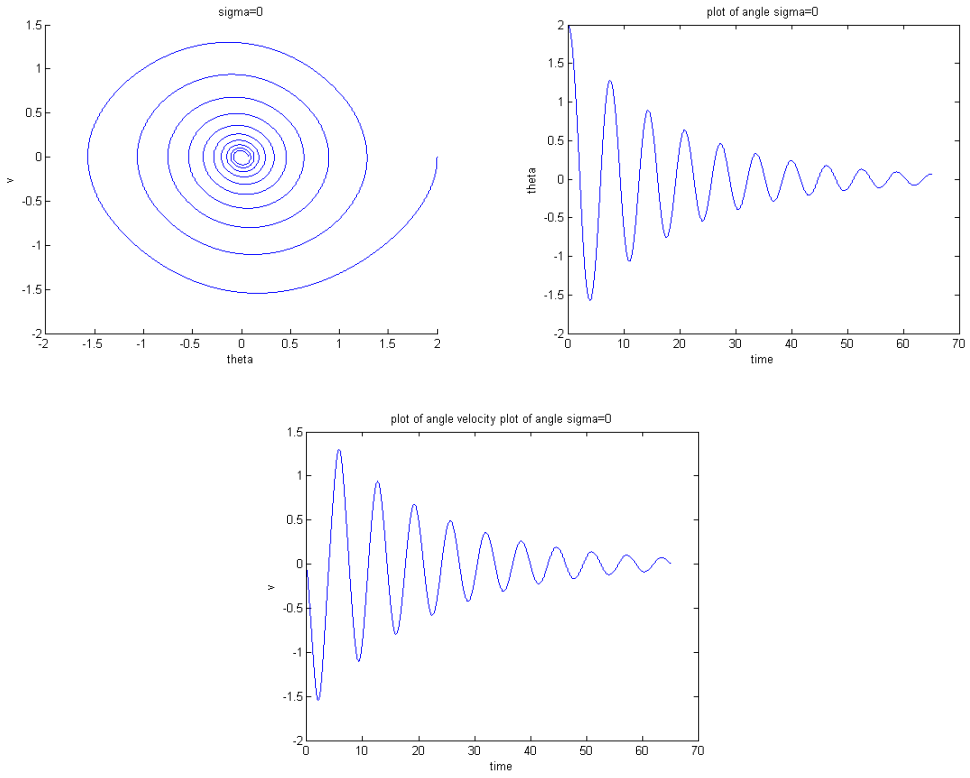

Figure C.8: Damped deterministic pendulum equation one trajectory.

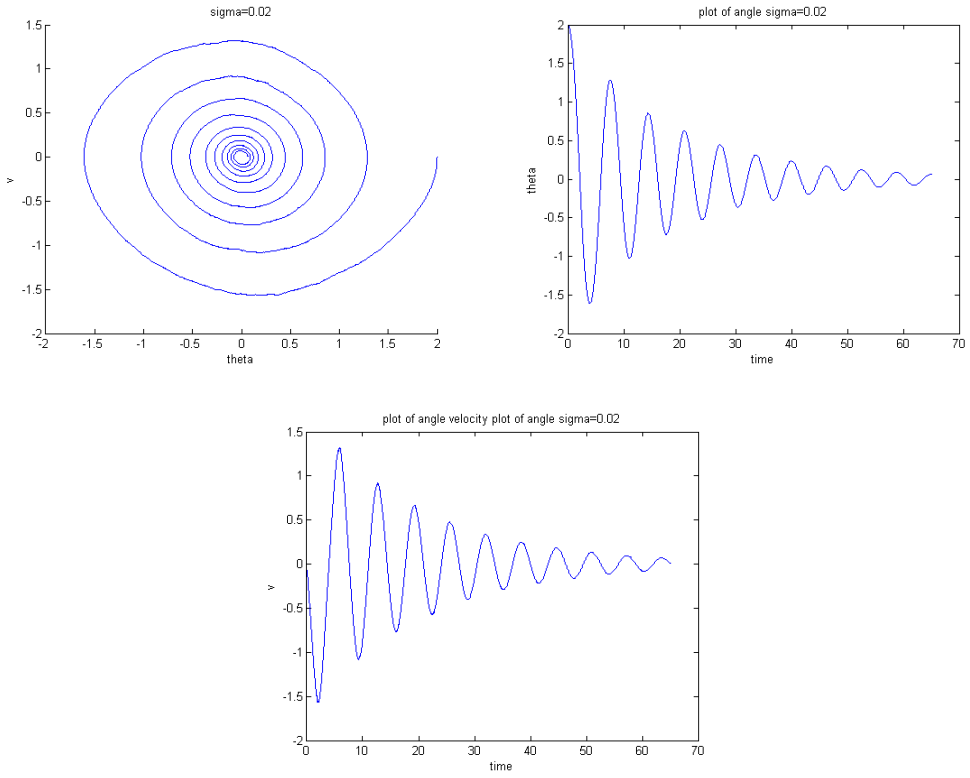

Figure C.9: Damped deterministic pendulum equation one trajectory with $\sigma = 0.02$.

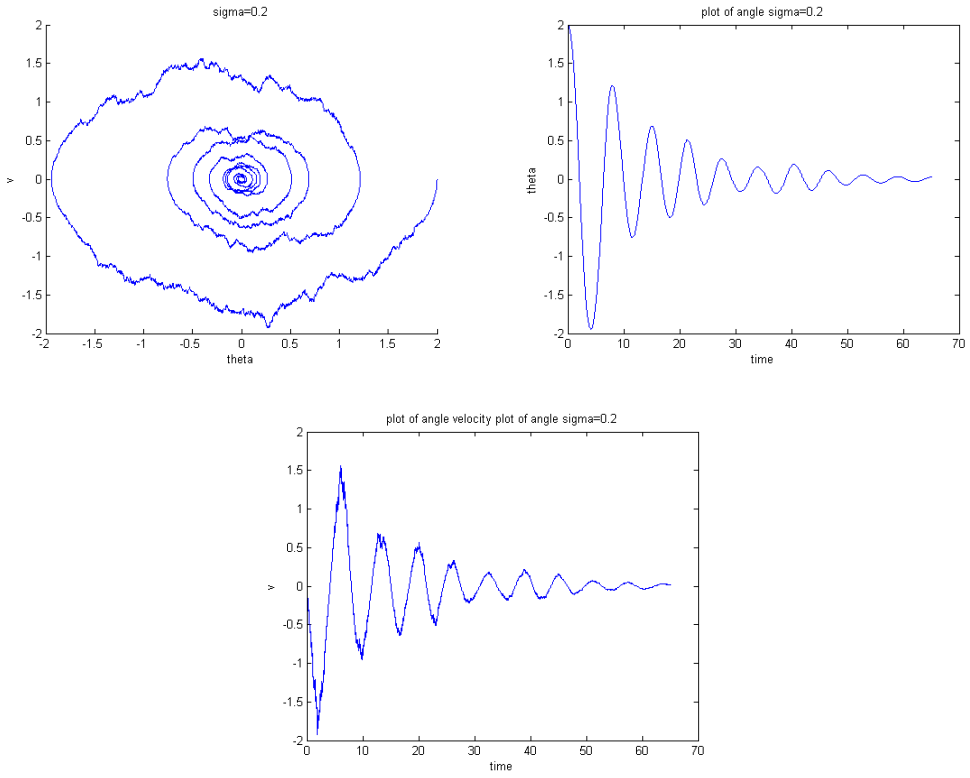

Figure C.10: Damped deterministic pendulum equation one trajectory with $\sigma = 0.2$.

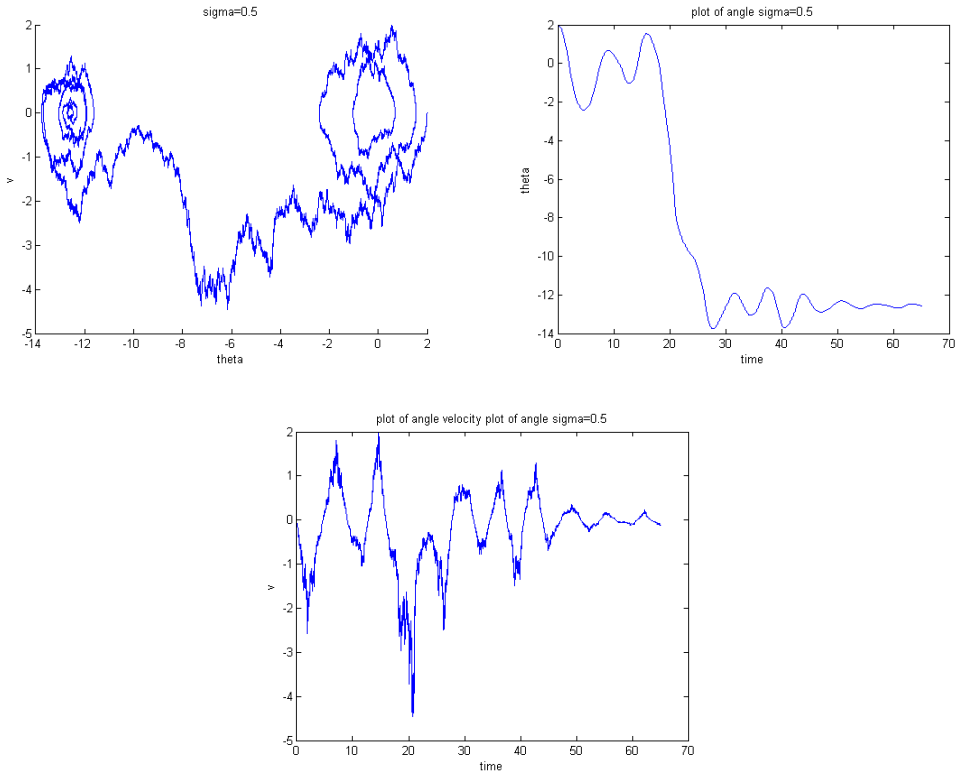

Figure C.11: Damped deterministic pendulum equation one trajectory with $\sigma = 0.05$.

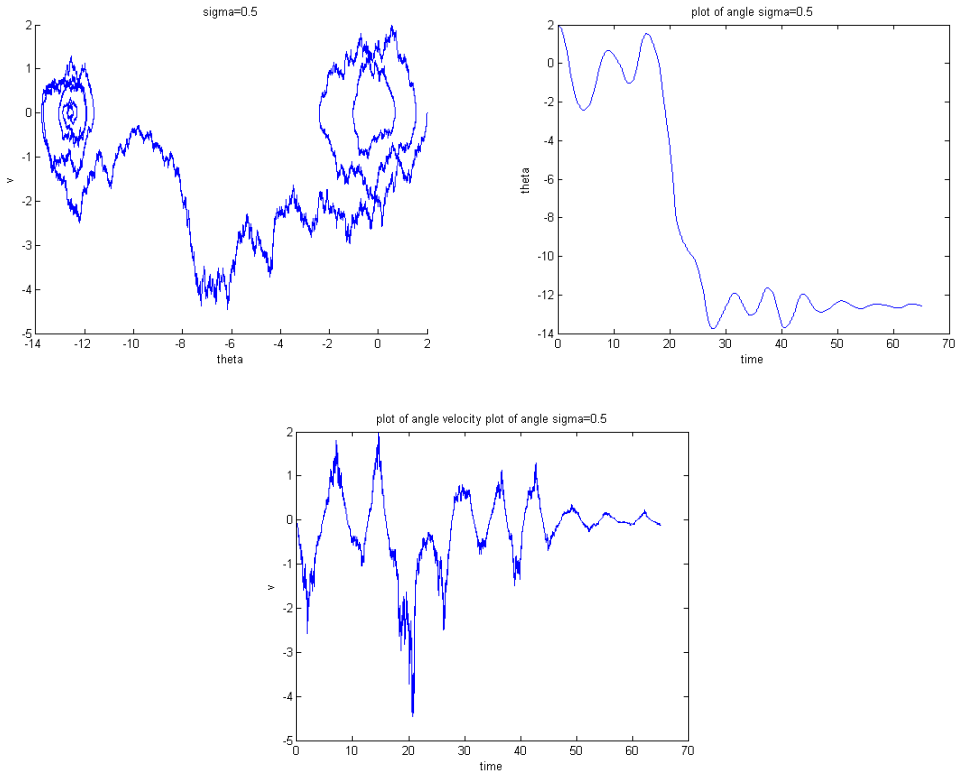

Figure C.12: Damped deterministic pendulum equation one trajectory with $\sigma = 5$.

Bibliography

[1] Eckhard Platen Peter E. Kloeden. *Numerical Solution of Stochastic Differential Equations*. Springer-Verlag Berlin Heidelberg New York, 1999.

[2] Bernt Oksendal. *Stochastic Differential Equations: An Introduction with Applications*. Springer-Verlag Berlin Heidelberg New York, 6th edition, 2005.

[3] Steven H. Strogatz. *Nonlinear Dynamics and Chaos with Applications to Physics, Biology, Chemistry, and Engineering*. Perseus Publishing, 1984.

[4] David Acheson. *From Calculus to Chaos An Introduction to Dynamics*. Oxford University Press Inc., New York, 1998.

[5] John C. Hull. *Options, Futures, and other Derivatives*. Prentice Hall, 6th edition, 2005.

[6] Neil D. Pearson. An efficient approach for pricing spread option. *The Journal of Derivatives Fall*, 1995.