8-2010

Conservation based uncertainty propagation in dynamic systems

Lillian J. Ratliff
University of Nevada, Las Vegas

CONSERVATION BASED UNCERTAINTY PROPAGATION IN DYNAMIC SYSTEMS

by

Lillian J Ratliff

Bachelor of Science in Electrical Engineering
University of Nevada, Las Vegas
2008

Bachelor of Science in Mathematics
University of Nevada, Las Vegas
2008

A thesis submitted in partial fulfillment
of the requirements for the

Masters of Science in Engineering
Department of Electrical and Computer Engineering
Howard R. Hughes College of Engineering

Graduate College
University of Nevada, Las Vegas
August 2010

UNLV
UNIVERSITY OF NEVADA LAS VEGAS

THE GRADUATE COLLEGE

We recommend the thesis prepared under our supervision by

Lillian J. Ratliff

entitled

Conservation Based Uncertainty Propagation in Dynamic Systems

be accepted in partial fulfillment of the requirements for the degree of

Master of Science in Engineering
Electrical and Computer Engineering

Pushkin Kachroo, Committee Chair

Mehdi Ahmadian, Committee Member

Rama Venkat, Committee Member

Yahia Bagzouz, Committee Member

Mohammad Trabia, Graduate Faculty Representative

Ronald Smith, Ph. D., Vice President for Research and Graduate Studies
and Dean of the Graduate College

August 2010

ABSTRACT

Conservation Based Uncertainty Propagation in Dynamic Systems

by

Lillian J Ratliff

Dr. Pushkin Kachroo, Examination Committee Chair
Professor of Electrical and Computer Engineering
University of Nevada, Las Vegas

Uncertainty is present in our everyday decision making process as well as our understanding of the structure of the universe. As a result an intense and mathematically rigorous study of how uncertainty propagates in the dynamic systems present in our lives is warranted and arguably necessary. In this thesis we examine existing methods for uncertainty propagation in dynamic systems and present the results of a literature survey that justifies the development of a conservation based method of uncertainty propagation. Conservation methods are physics based and physics drives our understanding of the physical world. Thus, it makes perfect sense to formulate an understanding of uncertainty propagation in terms of one of the fundamental concepts in physics: conservation. We develop that theory for a small group of dynamic systems which are fundamental. They include ordinary differential equations, finite difference equations, differential inclusions and inequalities, stochastic differential equations, and Markov chains. The study presented considers uncertainty propagation from the initial condition where the initial condition is given as a prior distribution defined within a probability structure. This probability structure is preserved in the sense of measure. The results of this study are the first steps into a generalized theory for uncertainty propagation using conservation laws. In addition, it is hoped that the

results can be used in applications such as robust control design for everything from transportation systems to financial markets.

ACKNOWLEDGMENTS

I would like to thank the National Science Foundation (NSF) for providing the funding through the NSF graduate research fellowship program (grfp) for the research presented in this thesis. I would also like to thank the Center for Vehicle Systems and Safety at Virginia Tech, and NASA for providing funding prior to the NSF grfp. I would like to thank my advisor, Dr. Pushkin Kachroo for the guidance he provided. In addition, I would like to thank all my committee members for participating.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Uncertainty is present in every aspect of our lives. Its is present in our everyday decision making process as well as our understanding of the structure of the universe. According to the economist George Shackle,

> In a predestinate world, decision would be illusory; in a world of a perfect foreknowledge, empty, in a would without natural order, powerless. Our intuitive attitude to life implies non-illusory, non-empty, non-powerless decision.... Since decision in this sense excludes both perfect foresight and anarchy in nature, it must be defined as choice in face of bounded uncertainty [1].

Uncertainty is a major element in the dynamical systems we study. Often we have incomplete information of the model for the dynamic system or of the initial and boundary conditions. As a result, it is important to develop a clear picture, in terms of scientific structure, of uncertainty. More than that, we should be able to incorporate our understanding of uncertainty into the control and design problem so that, as mathematicians and scientists, we can develop systems to handle uncertain events. In particular, propagation of uncertainty from the initial condition is essential to the design problem because the design of a system to handle variations in the initial conditions clearly improves the overall operation of the machine by making it more robust. The goal of studying how uncertainty propagates through a system is ultimately to understand how random variation or lack of knowledge affects the robustness of the system that is being modeled.

1.1 Definition of Uncertainty

Uncertainty can be defined in many ways. Thus, it is important to formulate a specific definition of uncertainty as it is studied in the research contained in this thesis. Uncertainty, itself, is the formalization of incomplete information and can be represented using a variety of methods. A representation of uncertainty can be obtained by assuming that uncertain quantities or statements are known only to the extent that their values belong to a set such as classes of models, ranges of parametric values, or knowledge within some probabilistic structure [2]. In the work presented in this thesis, we are concerned with stochastic uncertainty. By this we mean that the prior knowledge is given within some probabilistic structure. We denote this as a priori information. We aim to determine a posteriori information of future states given the a priori information about the initial state. This structure allows us to consider systems in general with given uncertainties of this type.

1.2 What is Uncertainty Propagation?

Uncertainty propagation is the study of propagation of elements of uncertainty through a system's dynamics. The result of which is a understanding of the state space of a dynamic system where uncertainty plays a role in determining the state at future times. Thus, the state is a function of the uncertainty. In a system there can be uncertainty in the parameters, the boundary conditions, the initial conditions, or the system dynamics. All of these types of uncertainty can manifest physically from noise in measurement devices, sensors, and other observation mechanisms or they may just be a result of our lack of a deterministic model for the parameters,

initial conditions, and boundary conditions. The study of propagation of uncertainty in the dynamic systems that our technologically encoded world is heavily dependent upon, allows control and system design to incorporate the inherent uncertainty into the final product whether it be as complicated as a traffic signal for an intersection or as simple as a toaster oven. The incorporation of uncertainty modeling into the standard design process will ultimately improve the systems on which we depend. Also, and more importantly, it will increase our fundamental understanding of the inherently uncertain physical world in which we live.

1.2.1 Conservation Based Uncertainty Propagation

There are many existing methods of propagating uncertainty. The method presented in this thesis is a physics based approach in which conservation laws are used as the method of propagating uncertainty. By this we mean that it preserves measure.

The dynamic system must be defined such that it is consistent with a probability structure. In order to ensure this, we consider the states to be represented as random variables. We are also given initial conditions defined on some probability space, $(\Omega, \mathcal{F}, P)$, where Ω is the sample space, $\mathcal{F}$ is the $\sigma-$algebra of subsets of Ω called events, and P is the probability measure. P is a function that maps $\mathcal{F}$ to the interval $[0, 1] \subset \mathbb{R}$ such that

1. $P(A) \geq 0$ for all $A \in \mathcal{F}$.

2. P is $\sigma-$additive. For disjoint events $\{A_n, n \geq 1\} \in \mathcal{F}$ we have

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n). \tag{1.1}$$

3. $P(\Omega) = 1$.

The initial condition must also be given as a random variable. Given the initial conditions defined on $(\Omega, \mathcal{F}, P)$, we construct a method for propagating uncertainty according to the system dynamics. This method preserves measure in the sense that it is consistent with the probabilistic structure on which the probability measure is defined. In this way we develop a conservation based method for uncertainty propagation where uncertainty is given as a prior and propagation is determined from the dynamics of the system under consideration.

1.3 Problem Statement

In this section a formal statement of the problem at hand is made. The purpose of this section is to explicitly define the problem of uncertainty propagation from the initial condition in dynamic systems using conservation. Conceptually, we mean that when given an evolution operator that maps an initial state space to a state space at future time t and given the probability structure on the initial state, for time $t \geq 0$, we define a probability structure on the state at time t such that the evolution operator is rendered measure preserving.

For the problem of uncertainty propagation, as studied in this thesis, we are given some dynamics and an initial condition which is unknown but we know some information defined within a probability structure, $(\Omega, \mathcal{F}, P)$. In particular, the initial condition is given as a random variable. Recall that a random variable is a measurable function from one measure space to another. We have a priori information about the initial condition, given as a probability density function, $u(X, 0)$. Using the dynamics,

we propagate uncertainty through the system by defining a probability structure on the state space in the future using the evolution operator which is measure preserving.

1.4 Outline of the Thesis

This thesis is divided into the following chapters. Each of the chapters on the different dynamic systems studied includes the development of the theory for uncertainty propagation in that system.

1. Chapter 1 is an introduction chapter in which we introduce the concept of uncertainty, uncertainty propagation and the problem statement as addressed in this thesis. Also, an outline of the thesis is presented.

2. Chapter 2 presents the background information including existing methods for uncertainty propagation and the results of a literature survey.

3. Chapter 3 presents examples of the types of dynamic systems studied in this thesis.

4. Chapter 4 presents ordinary differential equations and the theory for uncertainty propagation in ordinary differential equations. It also includes a derivation of the Liouville equation and its solution.

5. Chapter 5 presents finite difference equations and the theory for uncertainty propagation in finite difference equations.

6. Chapter 6 presents differential inclusions and inequalities and the theory for uncertainty propagation in differential inclusions and inequalities.

7. Chapter 7 presents stochastic differential equations and the theory for uncertainty propagation in stochastic differential equations. It also includes an introduction to and a discussion of Itô calculus as well as a derivation of the Fokker-Planck equation.

8. Chapter 8 presents Markov chains and the theory for uncertainty propagation in Markov chains. The theory for uncertainty propagation in Markov chains is presented through a study of finite state machines that have the Markov property.

9. Chapter 9 is an application chapter where uncertainty is propagated through Burgers' equation.

10. Chapter 10 is an application chapter in which the Liouville equation is solved numerically for some examples of ODEs.

11. Chapter 11 is an application chapter in which functions of random variables are applied for propagation of uncertainty in FDEs.

12. Chapter 12 summarizes the results of the research presented in this thesis.

1.5 Contributions

The work presented in this thesis makes several contributions. First, it presents a methods of uncertainty propagation in dynamic systems which take into account the fundamental physics behind the dynamic system resulting in understanding the geometry of the uncertainty. Having an understanding of the geometry of uncer-

tainty allows for qualitative analysis of dynamic systems under uncertainty. More specifically, the contributions of this thesis according to chapter are as follows:

1. In Chapter 4, the method for propagating uncertainty in ODEs is stated precisely. The material in this chapter was compiled from different sources and presented in a concise way.

2. The contribution of the work in Chapter 5 is the application of functions of random variables to the problem of uncertainty propagation in finite difference equations. The theory of functions of random variables was adapted from the work by Papoulis presented in [3]. The application of the theory of functions of random variables to the problem of uncertainty propagation in finite difference equations is new and constitutes a contribution. A second contribution of the work in this chapter is the development of the theory for the conservation form.

3. All of the work in Chapter 6 is a contribution in the sense that a conservation based theory for uncertainty propagation in differential inclusions and inequalities is first presented here. The work in this chapter is original work and the potential for application is immense.

4. In Chapter 7 on uncertainty propagation in stochastic differential equations, the contribution is similar to Chapter 4 in the sense that the work is not original. However, Chapter 7 is a concise statement on how to use the Fokker-Planck equation for uncertainty propagation in stochastic differential equations. The work was compiled from two main resources by Evans and Tanaka [4] and [5].

5. In Chapter 8, the contribution is the formulation of the problem of uncertainty propagation in Markov chains using existing knowledge about the transition properties of Markov chains. In addition, it makes the contribution of constructing the conservation form of the problem of uncertainty propagation in Markov chains.

6. Lastly, the contribution of Chapter 9 on Burgers' equation is the development of a method for uncertainty propagation from the initial condition. The work in this chapter is original.

Overall, the work in this thesis attempts to address the problem of uncertainty propagation in dynamic systems from a fundamental point of view providing a better understanding of the evolution of uncertainty and how it effects a systems behavior.

CHAPTER 2

BACKGROUND

We include a brief description of the existing methods for uncertainty propagation in an effort to provide a background and justification for this physics based approach. A literature survey has been conducted and is presented here in order to determine the relative merits of the existing methods of uncertainty propagation and to determine where and how the study presented in this thesis fits within the exiting literature and state of uncertainty modeling research. We conclude that physics based approach is needed.

2.1 Existing Methods

There are several existing methods currently used to model uncertainty as well as to propagate uncertainty in dynamic systems. The ones which are seen most frequently in the literature are the Monte Carlo method, polynomial chaos, and Bayesian inference.

2.1.1 Monte Carlo Method

The Monte Carlo method is a useful stochastic technique for modeling uncertainty by using random inputs. Its is often used in financial modeling where uncertainty is present. Monte Carlo methods are tedious to use with manual calculations. A large portion of the computational complexity of the Monte Carlo method comes from the need to make estimates of probability distributions. Computation through parallel processing and modern computational practices improves the speed of the large number of calculations necessary, but Monte Carlo is essentially a brute-force

method in which a large number of samples are taken from the distributions of the uncertain elements and are run through the system dynamics. Using the Monte Carlo method it is possible to estimate parameters and calculate statistical moments as well as estimate the distribution. While it is possible to estimate parameters and calculate moments, it is difficult to conclude from the results of Monte Carlo simulation the underlying geometry of the uncertainty as it evolves through the system dynamics.

2.1.1.1 Mathematical Description of the Monte Carlo Method

First, we present the Monte Carlo method through a simple example. We will solve the integral

$$I = \int_a^b f(x)dx, \tag{2.1}$$

where f is assumed integrable in the classic sense. Take N random samples from the interval $[a, b]$. For each sample x_i, we find the value $f(x_i)$. All of these values are summed. The sum is multiplied by $(b - a)$ and divided by N.

$$I = \frac{(b-a)}{N} \sum_{i=1}^{N} f(x_i) \tag{2.2}$$

The accuracy of the Monte Carlo method can be described, or quantified, through examination of statistical moments. In this example, we will look at variance. The sample variance is

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{N} (x_i - x)^2. \tag{2.3}$$

Now, let us examine the Monte Carlo method in a more general setting. Consider the estimator $\theta = E[h(X)]$, where $X = \{X_1, \ldots, X_n\}$ is a random vector in $\mathbb{R}^n$, $h(\cdot)$:

$\mathbb{R}^n \to \mathbb{R}$, and $E[|h(X)|] < \infty$. To estimate θ we first generate X_i for $i = 1, 2, \ldots, n$ and define $h_i = h(X_i)$ for each X_i. Then, we calculate $\hat{\theta_n}$.

$$\hat{\theta_n} = \frac{h_1 + \cdots + h_n}{n}, \tag{2.4}$$

where the hat denotes that $\hat{\theta}$ is an estimator for θ. We know that $\hat{\theta_n}$ is a good estimator since it is unbiased,

$$E[\hat{\theta_n}] = \frac{E[\sum_i^n h_i]}{n} = \frac{E[\sum_i^n h(X_i)]}{n} = \frac{n\theta}{n} = \theta, \tag{2.5}$$

and consistent,

$$\hat{\theta_n} \to \theta \text{ almost surely as } n \to \infty. \tag{2.6}$$

Consistency follows from the strong law of large numbers.

In this way, the Monte Carlo method can be used to propagate uncertainty from the initial condition or, more generally, propagate any uncertainties in a dynamic system. There are some advantages and disadvantages to using the Monte Carlo method. The algorithms are simple so that coding and debugging efforts are minimized [6]. As previously mentioned, a Monte Carlo method can be computationally taxing, using up a lot of resources and time. Thus, it may not be appropriate when efficiency is needed. In addition, online Monte Carlo estimation is not a feasible solution when real-time estimation is needed, as in control systems.

2.1.2 Polynomial Chaos

Polynomial Chaos was first presented by Weiner in the form of homogeneous chaos expansion in 1938 [7]. The fundamental idea is that random processes can be approximated with arbitrary accuracy by partial sums of orthogonal polynomial chaoses of random variables which are independent [8]. Polynomial Chaos is a method that uses polynomial-based stochastic space to represent and propagate uncertainty in the form of probability density functions [9]. All the uncertain parameters, variables, or components of the dynamic system are represented as random variables which are measurable functions. Each random variable, ξ, is associated with a random event, θ. The total number of random variables in the system is denoted as η_s. Each ξ is represented as a polynomial of finite dimension in terms of ξ_i. Normally a decomposition is composed of infinite terms. For practical purposes we use a finite number of terms and denote the number of terms as η_p. The single variable contributions from each uncertain variable in the system are combined into a multivariable polynomial [10]. The resulting polynomial is a representation of all the uncertainty in the system and the order of the multivariable polynomial is given by

$$P = \left(\frac{(\eta_s + \eta_p)!}{\eta_s! \eta_p!} \right) - 1. \tag{2.7}$$

In the polynomial chaos approach to simulation of propagation of uncertainty the solution is expressed as a truncated series and only one simulation is performed which is unlike the Monte Carlo method where there are a vast number of simulations [11]. As the number of terms retained in the series and the dimension of the stochastic

input increases the order of the polynomial chaos expansion increases. This is to say that the number of equations in the system that results from the polynomial chaos method increases.

2.1.2.1 Mathematical Description of Polynomial Chaos

The homogenous chaos expansion presented by Wiener uses a rescaled version of the Hermite polynomials which are given as

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} \left(e^{-x^2} \right). \tag{2.8}$$

The rescaling factor of $\sqrt{2}$ is used to achieve the probabilistic version of the Hermite polynomials which are given as

$$He_n(x) = H_n \left(\frac{x}{\sqrt{2}} \right) = (-1)^n e^{\frac{x^2}{2}} \frac{d^n}{dx^n} \left(e^{\frac{-x^2}{2}} \right). \tag{2.9}$$

The Cameron-Martin theorem states that Fourier-Hermite series converges in the L_2 sense to any L_2 functional [12]. This implies that the homogeneous chaos expansion converges to any stochastic processes with a second-order moment. Thus, the Hermite polynomial expansion provides a method of representing stochastic processes with Hermite polynomials [13].

Using the Hermite polynomials as the basis, every variable in the dynamic system is expanded along the multivariable polynomial basis. A second order random process, $X(\theta)$, with finite variance, can be used to describe polynomial chaos using the Hermite

polynomial basis as follows

$$X(\theta) = \alpha_0 He_0 + \sum_{i_1}^{\infty} \alpha_{i_1} He_1(\xi_{i_1}(\theta)) \tag{2.10}$$

$$+ \sum_{i_1}^{\infty} \sum_{i_2}^{i_1} \alpha_{i_1,i_2} He_2(\xi_{i_1}(\theta), \xi_{i_2}(\theta)) \tag{2.11}$$

$$+ \sum_{i_1}^{\infty} \sum_{i_2}^{i_1} \sum_{i_3=1}^{i_2} \alpha_{i_1,i_2,i_3} He_3(\xi_{i_1}(\theta), \xi_{i_2}(\theta), \xi_{i_3}(\theta)) \tag{2.12}$$

$$+ \cdots \tag{2.13}$$

where ξ is a random variable that is normally distributed with a zero mean and unit variance, $X \sim \mathcal{N}(0, 1)$. As we can see, there are an infinite number of terms in the expansion (2.13). For polynomial chaos, we take a finite number, P, of these terms and this results in the partial sum given in shorthand notation by

$$X(\theta) = \sum_{i=0}^{P} \beta_i He_i(\xi(\theta)). \tag{2.14}$$

Hermite polynomials are not the only basis that can be used in polynomial chaos, but they are commonly used. Hermite polynomials are in terms of Gaussian variables and are orthogonal to the weighting function. Some the other bases that are used are included in table (2.1).

A particular basis is chosen based on the dynamic system and uncertainty that is involved. After expansion of the variables using the chosen basis, the Galerkin projection is applied through integration of every component of the system in the polynomial form. The integration is performed in the appropriate space for the chosen

	Random Variables	Polynomial Basis Type	Support
Discrete	Poisson	Charlier-chaos	$\{0, 1, 2, \ldots\}$
	binomial	Krawtchouk-chaos	$\{0, 1, \ldots, N\}$
	negative binomial	Meixner-chaos	$\{0, 1, 2, \ldots\}$
	hypergeometric	Hahn-chaos	$\{0, 1, \ldots, N\}$
Continuous	Gaussian	Hermite-chaos	$(-\infty, \infty)$
	gamma	Laguerre-chaos	$[0, \infty)$
	beta	Jacobi-chaos	$[a, b]$
	uniform	Legendre-chaos	$[a, b]$

Table 2.1: Polynomial Chaos Basis Function Types [13]

polynomials. Since the Hermite polynomials form a complete orthonormal system in $L^2(\mathbb{R})$, the integration is performed in this space. Thus, the Galerkin projection for the Hermite polynomials results in the following inner product

$$\langle He_i He_j He_k \rangle = \int_{L^2} He_i He_j He_k w(\xi) d\xi \tag{2.15}$$

where $w(\cdot)$ is a weighting function of the number of uncertain variables in the system and is given by

$$w(\xi) = \left(\frac{1}{\sqrt{(2\pi)^{\eta_s}}} \right) e^{-\frac{1}{2}\xi^T \xi}. \tag{2.16}$$

The Galerkin projection is used to determine the equations for the time evolution of the spectral polynomial chaos coefficients. As the number of uncertainties grows, the polynomial chaos problem becomes more computationally intensive because the Galerkin method becomes inefficient due to the computation of the inner products used for projection. when the number of uncertainties grow, an alternative method to the Galerkin method is collocation.

Collocation is motivated by psuedo-spectral methods and is an alternative ap-

proach to solve stochastic random processes with polynomial chaos [14]. The collocation method evaluates the polynomial function at the roots of the basis polynomials which are either Legendre or Jacobi. Thus, when the dynamics are more complex, the collocation approach is more appropriate since each iteration is of a deterministic solver. In general, polynomial chaos is more advantageous than Monte Carlo in that the degree of computational resource consumption is much less. Polynomial chaos, fundamentally, still uses approximations to the original dynamics by projecting them onto lower dimensional manifolds, and thus, may be less effective in capturing the fundamental characteristics of the uncertainty.

2.1.3 Bayesian Inference

Bayesian inference is a type of statistical inference in which observations are used to infer what is known about underlying parameters. Bayesian philosophies differs from frequentist philosophies, such as Monte Carlo method, in the use of the term probability. In the frequency approach, probabilities are only used to summarize hypothetical replicate data sets, whereas in the Bayesian approach probability is used to describe all unknown quantities [15].

Using the Bayesian approach, we start with the formulation of a model for our dynamic system. It is desired that this formulation is 'adequate' to describe the system. An initial distribution, which we refer to as a prior, is formulated over the unknown parameters, initial conditions, or boundary conditions. The prior describes the incomplete information in the system, i.e. the uncertainties. Bayes' rule is applied in order to obtain a posterior distribution over the unknowns. The posterior accounts

for the initial data and the observed data. Using the posterior distribution, we can compute predictive distributions for future observations.

Often it is the case that we cannot translate subjective prior beliefs into a mathematically formulated prior. This can make the Bayesian method difficult to use. In addition, there can be computational difficulties with the Bayesian approach.

2.1.3.1 Mathematical Description of Bayesian Inference

The foundation of Bayesian inference is Bayes' theorem. First, let us recall the total probability formula. For an arbitrary event B, the total probability formula is given as

$$P(B) = \sum_{i=1}^{n} P(B|A_i)P(A_i), \tag{2.17}$$

where $A_1, A_2, \ldots, A_n$ are n mutually exclusive events with

$$P(\Omega) = \sum_{i} P(A_i) = 1 \ \ a.s. \tag{2.18}$$

Theorem 2.1.1. *(Bayes' Theorem) Given the total probability formula, we have*

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{P(B|A_1)P(A_1) + \cdots + P(B|A_n)P(A_n)}. \tag{2.19}$$

where $P(A_i)$ is the prior probability of the event A_i, $P(A_i|B)$ is the conditional probability of A_i given B, and $P(B|A_i)$ is the conditional probability of B given A_i.

Bayes' theorem allows us to evaluate the a posteriori probabilities $P(A_i|B)$ of the events A_i in terms of the a priori probabilities $P(A_i)$ and the conditional probabilities $P(B|A_i)$ [3].

Bayes' theorem can be extended to probability densities. Recall from probability theory that the probability distribution for any x_1 is defined as $U(x_1) = P\{X \leq x_1\}$ and $U(x_2) - U(x_1) = P\{x_1 < X \leq x_2\}$ where X is a random variable. Recall that a random variable is a measurable function from a probability space, $(\Omega, \mathcal{F}, P)$, to another measurable space, and real-valued random variables are such that $\Omega \mapsto \mathbb{R}$. Now,

$$U(x) = U(x|A_1)P(A_1) + \cdots + U(x|A_n)P(A_n) \tag{2.20}$$

follows directly from the total probability formula. Replacing $U(x)$, we have

$$P\{X \leq x\} = P\{X \leq x|A_1\}P(A_1) + \cdots + P\{X \leq x|A_n\}P(A_n). \tag{2.21}$$

In addition, recall the definition of the conditional distribution of the random variable X, where $\mathcal{M}$ is defined as the conditional probability of the event $\{X \leq x\}$.

$$U_X(x|\mathcal{M}) = P\{X \leq x|\mathcal{M}\} = \frac{P\{X \leq x, \mathcal{M}\}}{P(\mathcal{M})}, \tag{2.22}$$

where $\{X \leq x, \mathcal{M}\}$ is the event consisting of all outcomes ξ such that

$$(X(\xi) \leq x) \cap (\xi \in \mathcal{M}).$$

Now, consider the the arbitrary event A where $P(A) \neq 0$ and let $\mathcal{I} = \{x_1 < X \leq x_2\}$ be an interval where $x_1 < x_2$. Since $P(\mathcal{I}) = U(x_2) - U(x_1)$, it follows from (2.22)

that

$$P(A|\mathcal{I}) = \frac{P\{A, x_1 < X \leq x_2\}}{P(\mathcal{I})} = \frac{P\{A, x_1 < X \leq x_2\}}{U(x_2) - U(x_1)}. \tag{2.23}$$

Now, since $P(A)(U(x_2|A) - U(x_1|A)) = P\{x_1 < X \leq x_2, A\}$, we have

$$P(A|\mathcal{I}) = \frac{P(A)(U(x_2|A) - U(x_1|A))}{U(x_2) - U(x_1)}. \tag{2.24}$$

Let us now consider the case where $\mathcal{I} = \{X = x\}$. Suppose that given x, $u(x) \neq 0$ where u the probability density function corresponding to the random variable X. Then, we define $P(A|X = x)$ as follows

$$P(A|X = x) = \lim_{\Delta x \to 0} P(A|x < X \leq x + \Delta x). \tag{2.25}$$

The conditional density of X, $u(x|\mathcal{I})$, is defined as the derivative of the conditional distribution $U(x|\mathcal{I})$ and is given as

$$u(x|\mathcal{I}) = \lim_{\Delta x \to 0} \frac{P\{x \leq X \leq x + \Delta x|\mathcal{I}\}}{\Delta x}. \tag{2.26}$$

Letting $x_1 = x$ and $x_2 = x + \Delta x$, from (2.24) and (2.25) we have

$$P(A|X = x) = \frac{u(x|A)P(A)}{u(x)}. \tag{2.27}$$

Rearranging and integrating (2.27), we have

$$\int_{-\infty}^{\infty} P(A|X=x)u(x)dx = \int_{-\infty}^{\infty} u(x|A)P(A)dx. \tag{2.28}$$

Given that u is a density function, it has the property that

$$\int_{-\infty}^{\infty} u(x|\mathcal{I})dx = 1. \tag{2.29}$$

From (2.29), (2.28) becomes

$$\int_{-\infty}^{\infty} P(A|X=x)u(x)dx = P(A), \tag{2.30}$$

which is the continuous version of the total probability formula. From equations (2.27) and (2.30), we have

$$u(x|A) = \frac{P(A|X=x)u(x)}{\int_{-\infty}^{\infty} P(A|X=x)u(x)dx}. \tag{2.31}$$

Equation (2.31) is Bayes' theorem for probability density functions. This means we can find not only posterior probabilities, but posterior densities when the probability density function is appropriately defined. The two forms of Bayes' theorem are the foundation of Bayesian inference and can be used to propagate uncertainties. Initial uncertainties are updated using Bayes' theorem, and thus, we are able to determine the evolution of the initial uncertainties.

2.2 Relevant Literature

As we have seen there are several existing methods for propagation of uncertainty in dynamic systems. We now present the results of a literature survey conducted as part of the research in order to deduce how the approach presented in this thesis fits in with the current state of uncertainty propagation research. Several resources were found, and each of them verifies that our general study of dynamic systems fills a void in the literature.

Uncertainty has been studied in many different contexts in the literature. In general, within the literature, uncertainty is regarded as stochastic and it has been the case that the relevant problem determines how the study of uncertainty is conducted. More specifically, in the case of uncertainty propagation, the problem almost always determines the method of for studying the evolution for uncertainty. The methods most commonly present in the literature are the Monte Carlo method, polynomial chaos, and Bayesian inference.

In [16], the uncertainty propagation is studied in the context assembly tasks. The uncertainties are represented in the form of homogeneous transforms. More generally, the real location of an object is considered to be a nominal location with a small perturbation. The probabilistic information about an objects location is incorporated into this transform representation by an error vector. Also, in the problem of assembly tasks, it is often the case that the information about the probability distribution of the error vector is incomplete. Thus, the authors use moments, mean vector and covariance matrix, to characterize the uncertainty. Using this representation, the probabilistic model of uncertainty in location of an object is propagated spatially.

This is an example of how the application defines the method for studying uncertainty.

In [17], uncertainty propagation is studied in the context of sampling of measurement devices and then applied to metrology applications. The result is an improved sampling method, by which uncertainty of the measurement device is propagated statistically throughout the computation chain [17]. Again, here the authors use various moments to characterize the uncertainty.

In [18], uncertainty propagation is studied in the methodology for scoring dangerous chemical pollutants. Uncertainty in the scoring procedure is evaluated using the law of uncertainty propagation. The authors evaluate uncertainty on the basis of a scoring procedure which utilizes the moments of the uncertain parameters. These are only a few of the examples of the current state of uncertainty propagation research. Further examples of context specific studies can be found in [19], [20], [6], [21], [2], [12], [22], [23], [17], [24], [25], [15], [9], [10], [26], [27], [11], [14], [18], [28], [8], [16], and [29].

The majority of these studies of uncertainty propagation has been done within the context of the relevant problem. Thus, it is important to develop an understanding of uncertainty within the general context of dynamic systems as well as develop a theory which allows for analysis of the geometry of the uncertainty at future times. Analysis of this type can provide useful information about the qualitative properties of a system under uncertainty. One important result of our study is that it provides the background information that is necessary for further research into the control and design problem under uncertainty.

2.3 Motivation and Research Goal

The results of the literature survey clearly indicate that there is a need for a generalized theory for uncertainty propagation in dynamic systems. A conservation based method is ideal because it is physics based, and thus, it incorporates the fundamentals of the system dynamics into the study of uncertainty. This is essential. As a result of incorporating the fundamentals of the system dynamics, we are able to determine the geometry of the uncertainty as it is evolving. As we have seen in the brief discussion of current methods used for uncertainty propagation, this is greatly lacking. Having an understanding of the geometry allows for more concrete interpretations of qualitative properties of systems under uncertainty. Through the research described in this thesis, we begin to address this concern and void in the current state of studies of uncertainty propagation.

CHAPTER 3

DYNAMIC SYSTEMS

Dynamic systems can be constructed for everything from prediction and retrodiction to planning and control. Each system is a relation among the states of the systems variables with respect to their temporal evolution, and as a result, the relation allows for determination of unknown variables from known states of other system variables. Since we cannot observe exactly the states of the dynamic system at each step along the way, there is always some element of uncertainty. We must incorporate this into our mathematical model.

3.1 Examples of Dynamic Systems

The dynamic systems studied in this thesis are introduced with simple examples. The idea is to provide a basic understanding of each system through example. In addition, the concept of stability is introduced. Understanding the conditions for stability in dynamic systems is important for the development of controllers and the design of other components based on our models of these systems. In general, the theory for the stability of dynamic systems is well known. However, the theory for uncertain systems is not well known. When uncertainty is introduced into the system the conditions for stability change. In order to understand how these conditions change we must first develop a in depth understanding of the behavior of uncertain systems. This simple fact is one reason why the work presented in this thesis is important. By developing the theory for how uncertainty propagates through dynamic systems, we are making a step forward. In this thesis, we study the following dynamic

systems:

1. Ordinary Differential Equations (ODE),
2. Finite Difference Equations (FDE),
3. Differential Inclusions and Inequalities,
4. Stochastic Differential Equations (SDE), and
5. Markov Chains

In addition, we study uncertainty propagation through Burgers' equation through evaluation of expectation of velocity term at future times.

3.2 Ordinary Differential Equations

ODEs are used to model deterministic systems where the state evolves continuously in time. We consider systems which are modeled by a finite number of ODEs given, generally, by

$$\dot{x} = \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \vdots \\ \dot{x}_n \end{bmatrix} = \begin{bmatrix} f_1(x,t) \\ f_2(x,t) \\ \vdots \\ f_n(x,t) \end{bmatrix} = f(x,t), \tag{3.1}$$

where $x_1, x_2, \ldots, x_n$ are state variables and $\dot{x}$ represents the derivative of x with respect to time t. If they are autonomous, then the value of the future state depends only on the present state, and we write the state equation as $f(x)$. This implies that

a change in in the time variable from t to $\tau - t_0$ does not change the state equation. Conversely, if the are non-autonomous, then the system is dependent on time.

3.2.1 Pendulum

A pendulum is a simple example of a dynamic system that can be modeled as an ODE. Consider the pendulum pictured in figure (3.1). From Newton's second law of

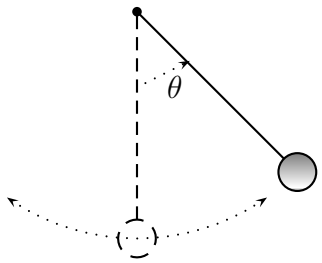


Figure 3.1: Pendulum Dynamics

motion, we have

$$m\ell\ddot{\theta} = -mg\sin\theta - k\ell\dot{\theta} \tag{3.2}$$

where m is mass, g is the usual gravitational constant, k is the coefficient of friction, θ is the angle between the vertical axis and the rod as the bob rotates about the pivot point, and ℓ is the radius of the circle along which the bob travels.

We can rewrite the model as a state space model. Let $x_1 = \theta$ and $x_2 = \dot{\theta}$. Then, we have

$$\dot{x_1} = x_2 \tag{3.3}$$

$$\dot{x_2} = -\frac{g}{\ell}\sin x_1 - \frac{k}{m}x_2 \tag{3.4}$$

The equilibrium points are the points where the state trajectory is stationary for all time. For the pendulum the equilibrium points are at $\dot{\theta} = 0$ and $\theta = \pm n\pi$.

3.2.2 Van der Pol Oscillator

The Van der Pol oscillator is an example of a dynamic system modeled by a second order ODE. It is a stable system, in the sense of Lyapnov, with a limit cycle.

Definition 3.2.1. *An equilibrium state, x_0, of a system is called stable in the sense of Lyapnov if given $\epsilon > 0$, for any t_0, there exists $\delta = \delta(t_0, \epsilon) > 0$ such that $\|x_0\| < \delta$ implies $\|x(t)\| < \epsilon$ for all $t > t_0$.*

Any perturbation results in the system returning to its limit cycle. The dynamics of a Van der Pol oscillator are given as

$$\frac{d^2x}{dt^2} - \mu(1 - x^2)\frac{dx}{dt} + x = 0, \tag{3.5}$$

and can be reformed into a system of first order ODEs.

$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ (1 - x_1)^2 x_2 - x_1 \end{bmatrix} \tag{3.6}$$

3.2.3 ODE model of Traffic Flow

Traffic can be modeled using an ODE approximation to the Lighthill-Whitham-Richards (LWR) partial differential equation model. The LWR model for traffic is the conservation equation for traffic density and Greenshield's model for velocity. It

is given as

$$\frac{\partial \rho(x,t)}{\partial t} + \frac{\partial \rho(x,t)v(x,t)}{\partial x} = 0 \tag{3.7}$$

with

$$v(x,t) = v_f \left(1 - \frac{\rho(x,t)}{\rho_m}\right) \tag{3.8}$$

where $\rho(x,t)$ is the traffic density, $v(x,t)$ is the velocity, v_f is the free-flow velocity, and ρ_m is the jam density.

Now, let us construct the system of ODEs which approximate the LWR model for traffic flow. For the sake of simplicity, we will discritize the section of road under

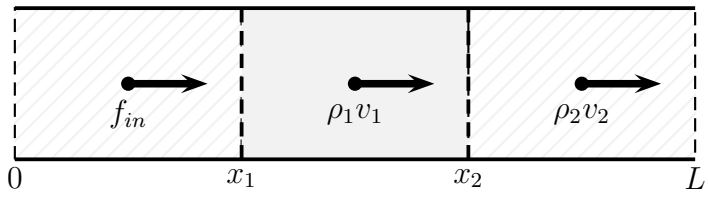


Figure 3.2: Discritization of Single-Lane Road

consideration into one section and two boundary sections. Figure (3.2) shows how the section of road is discritized. The system of ODEs which approximates the LWR model is given in equation (3.9).

$$\dot{\rho} = \begin{bmatrix} \dot{\rho_1} \\ \\ \dot{\rho_2} \end{bmatrix} = \begin{bmatrix} f_{in} - \rho_1 v_1 \\ \\ \rho_1 v_1 - \rho_2 v_2 \end{bmatrix} = F(\rho, t) \tag{3.9}$$

The ODE approximation is determined by taking the difference of the flux into a cell and the flux out of the cell and equating that with the change in the density with

respect to time.

3.3 Finite Difference Equations

Finite difference equations (FDE) are the discrete time analog of differential equations. FDEs are based on fundamental difference operations. We define a difference equation with the relation

$$y_{k+r} = F(k, y_k, y_{k+1}, \ldots, y_{k+r+1}) \tag{3.10}$$

where $k, r \in \mathbb{N}$ and r is the order of the difference equation. In order for the problem to be solvable, we must have sufficient initial information which in this case means we must have $y_1, y_2, \ldots, y_r$. Each successive value of y is determined based on the previous values. Following this method, it is easy to see how to evolve the system.

3.3.1 FDEs Given by Affine Transformations

The first example of an FDE that we present is an affine transformation. An affine is a transformation of the form

$$x \mapsto ax + b. \tag{3.11}$$

Thus, the FDE example that we consider is given by

$$x[k+1] = ax[k] + b, \tag{3.12}$$

where k is the discrete time index and $a, b \in \mathbb{R}$. In general, we require that the initial condition, $x[0]$, be given. Clearly, we can see how the system evolves iteratively.

Consider

$$x[1] = ax[0] + b$$

$$x[2] = ax[1] + b = a^2x[0] + ab + b$$

$$\vdots$$

$$x[k] = ax[k-1] + b = a^kx[0] + a^{k-1}b + a^{n-2}b + \cdots + a^2b + ab + b.$$

In general, the solution is

$$x[k] = a^kx[0] + \sum_{j=0}^{k-1} a^{k-1-j}b. \tag{3.13}$$

3.3.2 FDE Model of Traffic Flow

Recall the ODE model for traffic flow.

$$\begin{bmatrix} \dot{\rho_1} \\ \\ \dot{\rho_2} \end{bmatrix} = \begin{bmatrix} f_{in} - \rho_1 v_1 \\ \\ \rho_1 v_1 - \rho_2 v_2 \end{bmatrix} \tag{3.14}$$

Using Euler's method, we can construct a system of FDEs to model traffic flow. The first equation in the system becomes

$$\frac{\rho_1[t+1] - \rho_1[t]}{\Delta t} = f_{in} - \rho_1[t]v_1[t]. \tag{3.15}$$

Rearranging, we get

$$\rho_1[t+1] = \rho_1[t] + \Delta t(f_{in} - \rho_1[t]v_1[t]). \tag{3.16}$$

Performing similar operations on the second equation in the ODE system, we can now write the FDE system as

$$\begin{bmatrix} \rho_1[t+1] \\ \rho_2[t+1] \end{bmatrix} = \begin{bmatrix} \rho_1[t] + \Delta t(\rho_1[t]v_1[t] - f_{in}) \\ \rho_1[t] + \Delta t(\rho_1[t]v_1[t] - \rho_2[t]v_2[t]) \end{bmatrix} \tag{3.17}$$

3.4 Differential Inequalities and Inclusions

Differential inequalities and inclusions are set-valued maps. Set-valued maps are total relations in which every input is associated with multiple outputs. This means that set-valued maps are not injective, but they can be represented as functions if we consider point-sets. Differential inequalities are generalizations of standard inequalities and inclusions are further generalizations.

3.4.1 Examples of Differential Inclusions

In control theory, set-valued maps provide a nice framework for modeling systems. The map f describes the dynamics of the system. $f(x, \phi)$ is the velocity of the system where x is the state of the system and ϕ is the control. The set-valued map Φ describes a feedback map assigning to the state x the subset $\Phi(x)$ of possible controls. Thus, the map $F(x)$, a subset of feasible velocities, is defined as

$$F(x) := f(x, \Phi(x)) = \{f(x, \phi)\}_{\phi \in \Phi(x)}. \tag{3.18}$$

The control system that is governed by the family of differential equations given by

$$\dot{x}(t) = f(x(t), \phi(t)) \quad \text{where } \phi(t) \in \Phi(x(t)) \tag{3.19}$$

is equivalent to the differential inclusion given by

$$\dot{x}(t) \in F(x(t)). \tag{3.20}$$

Optimization studies are another area where uniqueness of the solution is lacking [30]. For example, let W be a function such that $X \times Y \mapsto \mathbb{R}$. Now, consider the family of minimization problems

$$\forall y \in Y, \;\; V(y) := \inf_{x \in X} W(x, y). \tag{3.21}$$

The function V is called the value function. For every $y \in Y$, we define

$$G(y) := \{x \in X | W(x, y) = V(y)\} \tag{3.22}$$

to be a subset of solutions to the minimization problems in (3.21).

3.4.2 Examples of Differential Inequalities

Differential inequalities are a special case of differential inclusions; they are less abstract. The differential inequalities studied in this thesis are finite dimensional dynamical systems.

Let $\mathcal{I} \subset \mathbb{R}$ be an interval and $\mathcal{O}$ be an open set in $\mathbb{R}^n$. Differential inequalities

are given by

$$f^m(x,t) \leq \dot{x} \leq f^M(x,t) \tag{3.23}$$

where $f^m \in C[\mathbb{R}^n, \mathcal{I} \times \mathcal{O}]$, $f^M \in C[\mathbb{R}^n, \mathcal{I} \times \mathcal{O}]$, and $f^m(x,t) \leq f^M(x,t) \quad \forall (x,t) \in \mathcal{I} \times \mathcal{O}$. If the differential inequality is a system then the inequalities are interpreted componentwise.

Let us consider a simple example of a differential inequality. Consider the example of the system of ODEs describing traffic flow. Now, if we modify the boundary conditions so that the input flux of traffic is given by the inequality $f_m \leq f_{in} \leq f^M$, then we get the following differential inequality

$$\begin{bmatrix} f_m - \rho_1 v_1 \\ \\ \rho_1 v_1 - \rho_2 v_2 \end{bmatrix} \leq \begin{bmatrix} \dot{\rho_1} \\ \\ \dot{\rho_2} \end{bmatrix} \leq \begin{bmatrix} f^M - \rho_1 v_1 \\ \\ \rho_1 v_1 - \rho_2 v_2 \end{bmatrix} \tag{3.24}$$

3.5 Stochastic Differential Equations

SDEs are mathematical models in which randomness is present in the system dynamics. Allowing for randomness in some of the coefficients of a differential equation or in the system itself, we obtain a more realistic mathematical model [31].

Before considering examples of SDEs, we introduce the method for mathematically representing the noise terms. Generally, the mathematical model for a SDE is

$$\begin{cases} dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dW_t \\ \\ X(0) = x_0 \end{cases} \tag{3.25}$$

where the subscript indicates time dependence, σ and μ are given functions, and dW_t

is the noise term. We say that $X(\cdot)$ solves (3.25) if

$$X(t) = x_0 + \int_0^t \mu(X(s))ds + \int_0^t \sigma(X(s))dW \quad \forall t > 0. \tag{3.26}$$

Here we consider the case where the noise is one-dimensional. W_t is a stochastic process that we use to describe the noise. We assume that it has the following properties [4]:

1. $W_0 = 0$ a.s.,
2. $W_t - W_s \sim \mathcal{N}(0, t-s) \quad \forall t \geq s \geq 0$, and
3. $\forall \ 0 < t_1 < t_2 < \cdots < t_n$, the random variables $W_{t_1}, W_{t_2} - W_{t_1}, \ldots, W_{t_n} - W_{t_{n-1}}$ are independent,

From the above properties, we can derive that $E[W_t] = 0$ for all t, and $E[W_t^2] = t$ for all t where $E(\cdot)$ is the expectation operator. The process W is commonly referred to as a Wiener process.

Now, equation (3.26) also requires that we define the stochastic integral. In this thesis, we present the Itô stochastic integral. The theory for SDEs can also be developed around the Stratonovich stochastic integral. At this point in the thesis we take for granted that we can perform this type of integration. A more detailed discussion on the definition of such an integral is included in the chapter on SDEs.

3.5.1 Population Growth

We will study a simple example of population growth where the coefficient for growth has some randomness. In general this is a more realistic model for population

growth. Consider a simple population growth model

$$\frac{dP}{dt} = a(t)P(t), \quad P(0) = A, \tag{3.27}$$

where $P(t)$ is the size of the population at time t, and $a(t)$ is the relative rate of growth at time t. The relative growth rate $a(t)$ may not be completely known. It may be subject to some random environmental effects. In this case we can express $a(t)$ as

$$a(t) = r(t) + N \tag{3.28}$$

where $N = \gamma W(t)$ is a noise term. γ is a constant and $W(t)$ is 'white noise'. We do not know the specific behavior of the noise term. Instead, we know its probability distribution. Let us assume $r(t) = r$ is a constant. Following the Itô interpretation, we have

$$dP(t) = rP(t)dt + \gamma P(t)dB(t). \tag{3.29}$$

3.6 Markov Chains

A Markov chain is a class of sequences of random variables taking values in a finite or countable set, called the state space, and satisfying the Markov property [32]. The Markov property refers to the characteristic of a sequence of random variables, or chain, where the conditional probability distribution of future states of the system are conditionally independent of the past states, i.e.

$$P(X_{n+1} = x_{n+1}|X_0 = x_0, \ldots, X_n = x_n) = P(X_{n+1} = x_{n+1}|X_n = x_n) \tag{3.30}$$

where $P(\cdot)$ is the probability function and X_k is the state of the system at time k. Time is considered a discrete quantity in the theory of Markov chains. The following is an important example of a Markov Chain.

3.6.1 Two-State Markov Chain

The following example was adapted from [33]. Let us consider a machine having two potential states: broken or operational. We can model the state of machine as a Markov chain. We denote the state space as $\mathcal{S} = \{1, 0\}$ where state '1' symbolically corresponds to the machine being operational and state '0' symbolically corresponds to the machine being broken.

Now, assume that if the machine is broken at time n, then with probability p the machine will be operational at time $n + 1$. Conversely, assume that if the machine is operational at time n, then with probability q the machine will be broken at time $n + 1$. Let $\rho_0(0)$, where the subscript indicates $X = 0$, denote the probability that the machine is broken initially. Thus, we have

$$P(X_{n+1} = 1 | X_n = 0) = p, \tag{3.31}$$

$$P(X_{n+1} = 0 | X_n = 1) = q, \tag{3.32}$$

and

$$P(X_0 = 0) = \rho_0(0). \tag{3.33}$$

As a result of there only being two states, 0 and 1, we have the following

$$P(X_{n+1} = 0|X_n = 0) = 1 - p, \tag{3.34}$$

$$P(X_{n+1} = 1|X_n = 1) = 1 - q, \tag{3.35}$$

and

$$P(X_0 = 1) = p_0(1) = 1 - \rho_0(0). \tag{3.36}$$

Now, we can compute the probability that the machine will be broken, $P(X_n = 0)$, or will operational, $P(X_n = 1)$, at time n. Clearly,

$$P(X_{n+1} = 0) = P(X_n = 0 \cap X_{n+1} = 0) + P(X_n = 1 \cap X_{n+1} = 0) \tag{3.37}$$

because if $X_{n+1} = 0$, then either $X_n = 0$ or $X_n = 1$ since there are only two states. We also know from the definition of conditional probability,

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \tag{3.38}$$

that

$$P(X_n = 0 \cap X_{n+1} = 0) = P(X_n = 0)P(X_{n+1} = 0|X_n = 0) \tag{3.39}$$

and

$$P(X_n = 1 \cap X_{n+1} = 0) = P(X_n = 1)P(X_{n+1} = 0|X_n = 1). \tag{3.40}$$

By substituting equations (3.39) and (3.40) into equation (3.37) and applying equa-

tions (3.34) and (3.32), we have

$$P(X_{n+1} = 0) = (1-p)P(X_n = 0) + qP(X_n = 1) \tag{3.41}$$

$$= (1-p)P(X_n = 0) + q(1 - P(X_n = 0)) \tag{3.42}$$

$$= (1-p-q)P(X_n = 0) + q \tag{3.43}$$

$$= (1-p-q)\rho_0(0) + q \tag{3.44}$$

We now can calculate the probability that the machine is broken at time $n = 2$.

$$P(X_2 = 0) = (1-p-q)P(X_1 = 0) + q \tag{3.45}$$

$$= (1-p-q)^2\rho_0(0) + q(1 + (1-p-q)) \tag{3.46}$$

By induction, we can calculate the probability that the machine is broken at time n.

$$P(X_n = 0) = (1-p-q)^n\rho_0(0) + q\sum_{k=0}^{n-1}(1-p-q)^k \tag{3.47}$$

Since the summation term in (3.47) is a geometric series such that

$$\sum_{k=0}^{n-1}(1-p-q)^k = \frac{1-(1-p-q)^n}{p+q}, \tag{3.48}$$

we have

$$P(X_n = 0) = \frac{q}{p+q} + (1-p-q)^n\left(\rho_0(0) - \frac{q}{p+q}\right). \tag{3.49}$$

In a similar fashion, we can derive the equation for the probability that the machine

is operational at time n.

$$P(X_n = 1) = \frac{p}{p+q} + (1-p-q)^n \left(\rho_0(1) - \frac{p}{p+q} \right). \tag{3.50}$$

3.7 Summary

Dynamic systems are important for modeling the physical world. This chapter presented a small group of dynamic systems for which we have developed the theory for uncertainty propagation. Each of the dynamic systems presented in this chapter are discussed in detail in the following chapters.

CHAPTER 4

ORDINARY DIFFERENTIAL EQUATIONS

In this chapter we review the concept of ordinary differential equations and how they are used in mathematical modeling. We also present the theory for how to propagate uncertainty in the initial condition through ODEs where this initial uncertainty is given within a probability structure. The method for propagating uncertainty is based on the conservation principle. The theory is compiled from several resources, including fundamental papers by Martin Ehrendorfer, and presented here in a concise form. In addition, the problem, as presented here, is consistent with the overall problem statement for uncertainty propagation in dynamic systems given in chapter 1. As a result, we have confirmation of the importance and significance of developing a theory of uncertainty propagation that is based on the conservation principle.

4.1 What are Ordinary Differential Equations?

ODEs are arguably the most important mathematical model for dynamic systems due to their simplicity and accuracy in depicting dynamics. A differential equation is an equation involving an unknown function and its derivatives [34]. An ODE is a differential equation if the unknown function depends only on one independent variable.

In general, we denote an ODE model for a dynamic system using the following equation:

$$\dot{x} = F(x, t), \tag{4.1}$$

where x is the state vector, the equations in F are the state equations, and $\dot{x}$ denotes

the derivative of the state with respect to time. ODEs can be autonomous or non-autonomous. The ODE represented by equation (4.1) is non-autonomous because the state equations depend on time t. Autonomous ODEs have state equations of the form $F(x)$.

4.2 Propagation of Uncertainty in ODEs

Uncertainty propagation in ODEs is performed using the Liouville equation (LE) [23], [35], [36], [37]. The LE is the mathematical formulation of the concept of conservation of density in the state manifold. The state manifold is the space in which all possible states of the system are represented. Each unique point in the state manifold corresponds to one possible state. From a statistical point of view, state manifold is estimated by an ensemble. An ensemble consists of a large number of realizations of a system considered all at once. Each realization represents a possible state of the real system. The local density points in this type of system lie on exactly one trajectory and obey Liouville's theorem which states that the state manifold distribution function is constant with respect to time along the trajectories of the system. Thus, we can take the local density points as constant. In the following section we will discuss the development of the LE for the density function.

4.3 Derivation of Liouville Equation

The derivation of the Liouville equation as presented here was adapted from the following references: [23], [35], [36], and [37]. Consider the N-dimensional dynamical

system with the state vector $X(t)$ given by

$$\dot{X} = F(X) \tag{4.2}$$

where $X(t) = (X_1(t), X_2(t), \ldots, X_N(t))$, the dot denotes the total derivative with respect to time t, and $F(X)$ describes how the state vector evolves in the state manifold.

Consider the probability space $(\Omega, \mathcal{F}, P)$. In general, let $u(X,t)$ be the density function describing the probability of state X occurring at time t. X is a random variable, mapping $\mathcal{F}$ to $\mathbb{R}$, and the density function $u(\cdot)$ is defined appropriately within the probability structure. Here, we take the probability structure to be $(\mathbb{R}^N, \mathcal{L}, P)$ where $\mathcal{L}$ is the σ-algebra on $\mathbb{R}^N$ composed of Lebesgue measurable subsets. The reason we take the sample space, Ω, to be $\mathbb{R}^N$ is that the state, X, belongs to $\mathbb{R}^N$ and the events on which we can define probability are those belonging to $\mathcal{L}$. We may note here that the Borel σ-algebra, $\mathcal{B}$, is a subset of $\mathcal{L}$.

Let the initial condition, given by $X(0)$, be unknown, but its statistical properties be known through its probability density function (pdf), $u_X(0)$. Thus, $u(X,0) = u_X(0)$.

The most important characteristic of the density function is that its integral over the whole state space is unity.

$$\int u(X,t)dX = 1 \tag{4.3}$$

Equation (4.3) implies the probability conservation equation in general since at any time t this property holds. Thus, as time evolves, the mass of the original density function remains the same. Here, we are concerned with understanding the geometry of the density function, thereby, allowing for a fundamental understanding of qualitative properties, such as stability, of the system under uncertainty. Briefly, let us consider this concept in a heuristic manner. Consider the case where the initial uncertainty diffuses. Figure 4.1 presents this example. We know that this system

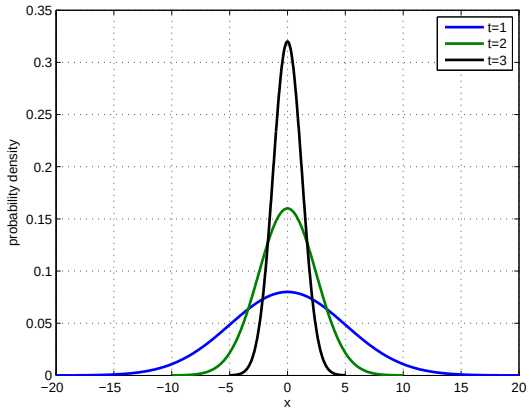


Figure 4.1: Example of diffusion of the pdf.

is unstable since the pdf diffuses across the real line. Conversely, consider the case where the prior evolves toward the Dirac delta function, $\delta(x - x_0)$. This indicates the system is stable since the probability that the state occupies $x = x_0$, in this case, converges to unity, i.e. $p(x_0) \to 1$. This is similar to the system converging to some equilibrium point when the initial state is within the support of the prior function. This example shows us that the geometry of the pdf is important and the fundamental principle presented in equation (4.3) informs us about the geometry.

When the dynamics of a system are Hamiltonian in nature, (4.3) implies the Li-

ouville equation[38]. In addition, $u(X,t) \geq 0$ for all X and t. The LE applied to the problem of uncertainty propagation allows for us to examine how the uncertainty in the initial condition, given as a pdf, evolves while preserving the probability measure. In this way, we are applying the conservation methodology to the problem of uncertainty propagation in ODEs.

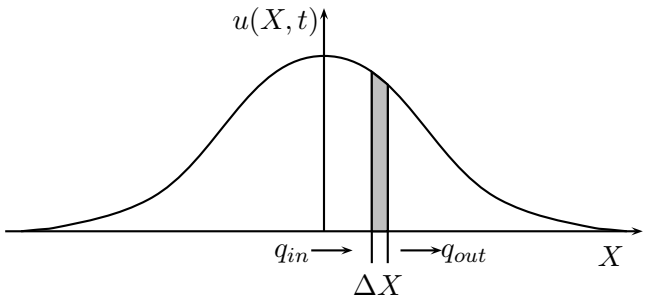


Figure 4.2: Conservation of Flow through ΔX.

In fluid dynamics, the Euler equations correspond to the Navier-Stokes equations with zero viscosity and heat conduction terms. When written in conservation form they represent conservation of mass, energy and momentum. Euler's equations can be applied to the system given in (4.2) in order to derive the LE.

Euler's equations are formulated by taking an infinitesimal volume of the state space and examining the flux through its surfaces. Taking the density function constructed above as a mass to be conserved, the Euler equations can be used to describe the continuity equation for $u(X,t)$ given the dynamics $F(X)$.

Using the Eulerian method just described, we will develop the LE in one dimen-

sion. Consider the one dimensional system

$$\dot{X} = F(X). \tag{4.4}$$

The net flow through the volume must be equal to the change in volume under $u(X,t)$,

$$\frac{\partial u(X,t)}{\partial t}\Delta X = q_{in} - q_{out}, \tag{4.5}$$

where q_{in} and q_{out} denote the flow in and flow out of the volume respectively. We can

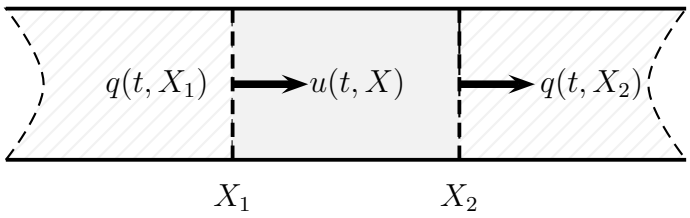


Figure 4.3: Conservation of Flow through ΔX: A Differential View.

replace q_{in} and q_{out} by the equations for the flow entering and leaving the volume, ΔX,

$$q_{in} = u(X,t)F(X), \tag{4.6}$$

$$q_{out} = u(X + \Delta X, t)F(X + \Delta X). \tag{4.7}$$

Substituting equations (2.4) and (2.5) into (2.3) gives,

$$\frac{\partial u(X,t)}{\partial t}\Delta X = u(X,t)F(X) - u(X + \Delta X, t)F(X + \Delta X). \tag{4.8}$$

Rearranging and taking the limit results in the LE in one dimension

$$\frac{\partial u(X,t)}{\partial t} = -\frac{\partial u(X,t)F(X)}{\partial X}. \tag{4.9}$$

Extending the LE to higher dimensions, the net flow is now given by the divergence of the flux through the volume. The generalized LE is

$$\frac{\partial u(X,t)}{\partial t} + \sum_{i=1}^{n} \frac{\partial}{\partial X_i}[u(X,t)F_i(X)] = 0, \tag{4.10}$$

where $X = \{X_1, X_2, \ldots, X_n\}$ and $F_i(X)$ is the i-th component of $F(X)$. The LE is linear in its first derivatives of the variable u; thus, it is an inhomogeneous, semi-linear, first-order partial differential equation. The independent variables are the components of the state vector, $X_1, X_2, \ldots, X_n$ and time t. The dependent variable is the density u.

4.4 Solution to the Liouville Equation

In this section, the LE will be solved giving the general form of the analytical solution to (4.10) explicitly. The LE is a semi-linear partial differential equation (PDE). Thus, we can use the method of characteristics to find the solution to the PDE.

4.4.1 Solution By Method of Characteristics

In the previous section, we defined the system given in (4.2). The solution, $X(t)$, to (4.2) is a function of the initial condition Ξ and time t:

$$X = X(\Xi, t). \tag{4.11}$$

Here we will denote the initial condition, $X(0)$, as Ξ. Assuming that the solution to (4.2) exists and is unique, there is a unique Ξ corresponding to every $X(t)$. Specifically,

$$\Xi = \Xi(X, t) \tag{4.12}$$

and

$$\Xi = X(0). \tag{4.13}$$

Given equations (4.11) and (4.12), the relationship between the initial state and the state of the system is injective.

By expanding (4.10) we get the following equation:

$$\frac{\partial u(X,t)}{\partial t} + \sum_{i=1}^{n} F_i(X) \frac{\partial u(X,t)}{\partial X_i} + u(X,t) \sum_{i=1}^{n} \frac{\partial F_i(X)}{\partial X_i} = 0. \tag{4.14}$$

By examining the first two terms on the left side, we can see that they form the full derivative of the pdf, $u(X, t)$, with respect to time. Now, (4.14) can be rewritten as

$$\frac{du(X,t)}{dt} = -u(X,t)\psi(X), \tag{4.15}$$

where

$$\psi(X) = \sum_{i=1}^{n} \frac{\partial F_i(X)}{\partial X_i}. \tag{4.16}$$

Using separation of variables and integration, (4.15) becomes

$$\int_{u(\Xi,0)}^{u(X,t)} \frac{1}{u(X,\zeta)} du(X,\zeta) = -\int_0^t \psi(X(\zeta)) d\zeta. \tag{4.17}$$

When evaluating $u(X,t)$ at a given time t and state X, X and t are independent. The solution to the LE is

$$u(X,t) = u_o(\Xi(X,t)) \exp\left[-\int_0^t \psi(\hat{X}(\zeta)) d\zeta\right]. \tag{4.18}$$

where

$$u(X(0),0) = u_0(\Xi(X,t)), \tag{4.19}$$

and $\Xi(X,t)$ is the initial condition of the system for a given X and t in the associated state manifold.

4.4.2 Verification of Solution

Now that we have the solution (4.18), we will plug it back into (4.10) to verify that it satisfies the LE. First, we will define

$$g(\Xi,t) := -\int_0^t \psi[X(\Xi,\zeta)] d\zeta. \tag{4.20}$$

Now, we will start by taking partial derivatives of $u(X,t)$ with respect to X and t. Plugging the solution (4.18) into the left hand side (LHS) of (4.10), gives

$$\frac{\partial}{\partial t}\left[u_0(\Xi,t)e^{g(\Xi,t)}\right]+\sum_{k=1}^{N}F_k(X)\left[\frac{\partial}{\partial X_k}\left(u_0(\Xi,t)e^{g(\Xi,t)}\right)\right]+\left(u_0(\Xi,t)e^{g(\Xi,t)}\right)\psi(X). \quad (4.21)$$

We can eliminate $e^{g(\Xi,t)}$ from each term in the LHS so that we have

$$\frac{\partial}{\partial t}\left[u_0(\Xi,t)\right]+\sum_{k=1}^{N}F_k(X)\left[\frac{\partial}{\partial X_k}u_0(\Xi,t)\right]+u_0(\Xi,t)\psi(X). \quad (4.22)$$

Define (4.22) to be Υ. Now, we must show that Υ vanishes, verifying that $u(X,t)$ as define in (4.18) satisfies (4.10). We start by applying the chain rule to the derivatives in Υ.

$$\Upsilon=\sum_{i=1}^{N}\frac{\partial u_0(\Xi)}{\partial \Xi_i}\frac{\partial \Xi_i(X,t)}{\partial t}+\sum_{i=1}^{N}\sum_{j=1}^{N}F_j(X)\frac{\partial u_0(\Xi)}{\partial \Xi_i}\frac{\partial \Xi_i(X,t)}{\partial X_j}$$

$$+u_0(\Xi)\frac{\partial g(\Xi,t)}{\partial t}+u_0(\Xi)\sum_{j=1}^{N}F_j(X)\frac{\partial g(\Xi,t)}{\partial X_j}+u_0(\Xi)\psi(X). \quad (4.23)$$

After gathering like terms, rearranging we have

$$\Upsilon=\sum_{i=1}^{N}\frac{\partial u_0(\Xi)}{\partial \Xi_i}\Upsilon_1+u_0(\Xi)\Upsilon_2, \quad (4.24)$$

where Υ_1 and Υ_2 are defined as

$$\Upsilon_1:=\frac{\partial \Xi_i(X,t)}{\partial t}+\sum_{j=1}^{N}F_j(X)\frac{\partial \Xi_i(X,t)}{\partial X_j} \quad (4.25)$$

and

$$\Upsilon_2 := \frac{\partial g(\Xi,t)}{\partial t} + \sum_{j=1}^{N} F_j(X)\frac{\partial g(\Xi,t)}{\partial X_j} + \psi(X). \tag{4.26}$$

We desire that Υ go to zero. Thus, Υ_1 and Υ_2 must vanish since $u_0(\Xi)$ is arbitrary.

We will now consider the total derivatives of X_j and Ξ_l.

$$dX_j(\Xi,t) = \sum_{m=1}^{N} \frac{\partial X_j(\Xi,t)}{\partial \Xi_m} d\Xi_m + \frac{\partial X_j(\Xi,t)}{\partial t} dt, \tag{4.27}$$

$$d\Xi_l(X,t) = \sum_{n=1}^{N} \frac{\partial \Xi_l(X,t)}{\partial X_n} dX_n + \frac{\partial \Xi(X,t)}{\partial t} dt. \tag{4.28}$$

Now, plugging (4.27) into (4.28) gives

$$d\Xi_l(X,t) = \sum_{n=1}^{N} \left\{ \frac{\partial \Xi_l(X,t)}{\partial X_n} \left[\sum_{m=1}^{N} \frac{\partial X_n(\Xi,t)}{\partial \Xi_m} d\Xi_m + \frac{\partial X_n(\Xi,t)}{\partial t} dt \right] \right\} + \frac{\partial \Xi_l(X,t)}{\partial t} dt. \tag{4.29}$$

By rearranging (4.29) we have

$$\begin{aligned} d\Xi_l(X,t) = & \sum_{n=1}^{N}\sum_{m=1}^{N} \frac{\partial X_n(\Xi,t)}{\partial \Xi_m}\frac{\partial \Xi_l(X,t)}{\partial X_n} d\Xi_m \\ & + \left[\sum_{n=1}^{N} \frac{\partial \Xi_l(X,t)}{\partial X_n}\frac{\partial X_n(\Xi,t)}{\partial t} + \frac{\partial \Xi_l(X,t)}{\partial t} \right] dt. \end{aligned} \tag{4.30}$$

Because the product of the Jacobian of the transformation and its inverse is unity, the first term in the RHS of (4.30) is equal to the LHS of (4.30). Thus, the second

term in the RHS is equivalent to 0. Rearranging the second term gives the equality

$$\frac{\partial \Xi_l(X,t)}{\partial t} = -\sum_{n=1}^{N} \frac{\partial \Xi_l(X,t)}{\partial X_n} \frac{\partial X_n(\Xi,t)}{\partial t}. \tag{4.31}$$

We can now plug (4.31) into Υ_1.

$$\Upsilon_1 = -\sum_{n=1}^{N} \frac{\partial \Xi_l(X,t)}{\partial X_n} \frac{\partial X_n(\Xi,t)}{\partial t} + \sum_{j=1}^{N} F_j(X) \frac{\partial \Xi_i(X,t)}{\partial X_j} \tag{4.32}$$

$$= \sum_{n=1}^{N} \frac{\partial \Xi_i(X,t)}{\partial X_n} \left[F_n(X) - \frac{\partial X_n(\Xi,t)}{\partial t} \right]. \tag{4.33}$$

Notice the term in the brackets on the RHS of (4.33). By (4.11), $X(\Xi,t)$ is a solution to (4.2). Thus, the term in the brackets must go to zero. Therefore, Υ_1 vanishes.

Now, we must show that Υ_2 vanishes. Since Υ_2 depends on $g(\Xi,t)$, we must examine its derivatives. The partial derivative of $g(\Xi,t) = g[\Xi(X,t)] = g(X,t)$ with respect to t is

$$\frac{\partial g(\Xi,t)}{\partial t} = -\frac{\partial}{\partial t} \int_0^t \psi\{X[\Xi,\zeta]\} d\zeta \tag{4.34}$$

$$= -\psi(X) - \int_0^t \frac{\partial}{\partial t} \psi\{X[\Xi,\zeta]\} d\zeta \tag{4.35}$$

$$= -\psi(X) - \int_0^t \sum_{i=1}^{N} \frac{\partial \psi(X)}{\partial X_i} \frac{\partial X_i[\Xi,\zeta]}{\partial t} d\zeta \tag{4.36}$$

$$= -\psi(X) - \int_0^t \left[\sum_{i=1}^{N} \frac{\partial \psi(X)}{\partial X_i} \sum_{j=1}^{N} \frac{\partial X_i(\Xi,\zeta)}{\partial \Xi_j} \frac{\partial \Xi_j(X,t)}{\partial t} \right] d\zeta. \tag{4.37}$$

The partial derivative of $g(\Xi, t)$ with respect to X_i is as follows:

$$\frac{\partial g(\Xi,t)}{\partial X_i} = -\frac{\partial}{\partial X_i}\int_0^t \psi\{X[\Xi(X,t),\zeta]\}d\zeta \tag{4.38}$$

$$= -\int_0^t \sum_{k=1}^{N} \frac{\partial \psi(X)}{\partial X_k}\frac{\partial X_k[\Xi,\zeta]}{\partial X_i}d\zeta \tag{4.39}$$

$$= -\int_0^t \left[\sum_{k=1}^{N} \frac{\partial \psi(X)}{\partial X_k}\sum_{m=1}^{N}\frac{\partial X_k(\Xi,\zeta)}{\partial \Xi_m}\frac{\partial \Xi_m(X,t)}{\partial X_i}\right]d\zeta. \tag{4.40}$$

Now, we plug (4.37) and (4.40) into Υ_2.

$$\Upsilon_2 = -\psi(X) - \int_0^t \left[\sum_{i=1}^{N} \frac{\partial \psi(X)}{\partial X_i}\sum_{j=1}^{N}\frac{\partial X_i(\Xi,\zeta)}{\partial \Xi_j}\frac{\partial \Xi_j(X,t)}{\partial t}\right]d\zeta$$

$$-\sum_{k=1}^{N} F_i(X)\int_0^t \left[\sum_{l=1}^{N} \frac{\partial \psi(X)}{\partial X_l}\sum_{m=1}^{N}\frac{\partial x_l(\Xi,\zeta)}{\partial \Xi_m}\frac{\partial \Xi_m(X,t)}{\partial X_k}\right]d\zeta + \psi(X) \tag{4.41}$$

Rearranging and eliminating terms, we get

$$\Upsilon_2 = -\sum_{i=1}^{N}\sum_{j=1}^{N} \frac{\partial \psi(X)}{\partial X_i}\left[\frac{\partial \Xi_j(X,t)}{\partial t} + \sum_{k=1}^{N} F_k(X)\frac{\partial \Xi_j(X,t)}{\partial X_k}\right]\int_0^t \frac{\partial X_i(\Xi,\zeta)}{\partial \Xi_j}d\zeta. \tag{4.42}$$

The term in the LHS of (4.42) in the brackets is just Υ_1, which we have shown to vanish. Thus, $\Upsilon_2 \to 0$. Now that we have shown that Υ_1 and Υ_2 vanish, it is obvious that Υ vanishes. We have shown that (4.18) satisfies (4.10); thus, (4.18) is the general solution of the LE.

4.5 Summary

In summary, the LE is the appropriate conservation based method for propagating uncertainty in the initial condition through ODEs [23], [35], [36], [37]. This chapter provides the derivation and verification of the solution to the LE. Computation of the solution requires numerical techniques. The contribution of this chapter is that it provides the complete problem statement and solution of uncertainty propagation in ODEs.

CHAPTER 5

FINITE DIFFERENCE EQUATIONS

In this chapter we review the concept of difference equations and how they are used in mathematical modeling. We also propose a method for propagating uncertainty through deterministic FDEs with uncertainty in the initial condition defined within a probabilistic structure. The method for propagating uncertainty is based on the theory of functions of random variables and can be written in conservation form. The contribution of this chapter is the application of the theory of functions of random variables to the problem of uncertainty propagation in FDEs.

5.1 What are Finite Difference Equations?

FDEs are the discrete time analog of differential equations. In some cases a discrete model that leads to a FDE may be more natural than a continuous model that leads to a differential equation. An example of this phenomena is population growth. Population growth in a species whose generations do not overlap and propagate at regular intervals is more accurately modeled using a discrete time model.

We define a difference equation with the relation

$$x_{k+r} = F(k, x_k, x_{k+1}, \ldots, x_{k+r+1}) \tag{5.1}$$

where $k, r \in \mathbb{N}$ and r is the order of the difference equation. In order for the problem to be solvable, we must have sufficient initial information which in this case means we must have $x_1, x_2, \ldots, x_r$. Each successive value of x is determined based on the previous values. Following this method, it is easy to see how to evolve the system.

FDEs can be both linear and nonlinear. The FDE in (5.1) is called linear if F is a linear function of x_k where linear is defined in the standard way. Otherwise, it is called nonlinear. A solution to the equation (5.1) is a sequence of numbers $\{x_k\}$ that satisfy the FDE for each k. In this sense, $\{x_k\}$ is equivalent to the state trajectory.

5.2 Uncertainty Propagation in Finite Difference Equations

In order to formulate the problem statement, without loss of generality, we consider a system in which the dynamics are only dependent on the current state, i.e. a first order FDE, and the state vector is one dimensional. By this we mean that we are given a FDE of the form

$$X_{k+1} = F(k, X_k), \tag{5.2}$$

where X_k is shorthand notation for $X[k]$, the X_i's are random variables defined on the probability space $(\Omega, \mathcal{F}, P)$, and the initial condition is given as the random variable X_0. The probability space here is $(\mathbb{R}, \mathcal{L}, P)$, where $\mathcal{L}$ is the σ-algebra of the Lebesgue measurable sets defined on $\mathbb{R}$. Thus, X_i is the i-th random variable such that $\mathcal{L} \mapsto \mathcal{L}$. We do not have sufficient initial information to know X_0 exactly. We do, however, know its pdf which is given by $u_0 = u(X, 0)$ and defined within the appropriate probability structure.

Now, the method for propagating uncertainty through FDEs from the initial condition uses functions of random variables. Functions of random variables are the basis for the method of uncertainty propagation through FDEs. The theory of functions of random variables is developed through examining first functions of one random

variable and then functions of multiple random variables. In the following sections we consider first order FDEs and then higher order FDEs following which we consider systems of FDEs.

5.3 1-D Finite Difference Equation

Given

$$X_{k+1} = g(X_k) \tag{5.3}$$

with $X_0 = u_0$ as the initial condition where $u_0 = u(X, 0)$ is a pdf, we can apply the theory of functions of random variables to solve for the pdf of X_1 in terms of u_0 based on the dynamics of the system given in (5.3). Iterating we can solve for X_k for any time k.

First, let us introduce the theory of functions of one random variable. The theory for functions of random variables as presented here is adapted from [3]. In general, if we are given

$$Y = g(X) \tag{5.4}$$

where X is a random variable and $g(x)$ is a real valued function, then Y is also a random variable defined by (5.4). Now, let X be continuous and let $u_X(x)$ be its pdf. We call $y = g(x)$ a transformation. If $g(x)$ is a continuous, monotonic function, then the transformation is injective and has the inverse transformation

$$x = g^{-1}(y) = h(y). \tag{5.5}$$

The pdf for Y is given by

$$u_Y(y) = u_X(x) \left| \frac{dx}{dy} \right|. \tag{5.6}$$

If $g(x)$ is not injective, then there is more than one solution and we find $u_Y(y)$ by obtaining all the real roots of $y = g(x)$

$$y = g(x^{(1)}) = g(x^{(2)}) = \ldots = g(x^{(n)}) = \ldots, \tag{5.7}$$

where the $x^{(i)}$ is the i-th root, and calculating

$$u_Y(y) = \sum_i \frac{u_X(x^{(i)})}{|g'(x^{(i)})|} \tag{5.8}$$

where $' = \frac{d}{dx}$.

Now, let us go back and consider the FDE in equation (5.3). Since the initial condition is a random variable, $u_0 = u(X, 0)$, we can find x_{k+1} by transforming the random variable by the following method. If (5.3) is injective, i.e. there is one solution to (5.3), then the pdf of X_{k+1} is given by

$$u_{X_{k+1}}(x_{k+1}) = u_{X_k}(x_k) \left| \frac{dx_k}{dx_{k+1}} \right|. \tag{5.9}$$

We know for a monotonic, continuous function with the inverse transformation $x_k = g^{-1}(x_{k+1}) = h(x_{k+1})$ we have

$$u_{X_{k+1}}(x_{k+1}) = u_{X_k}(x_k) \left| \frac{dx_k}{dx_{k+1}} \right| = u_{X_k}[h(x_{k+1})] \left| \frac{dh(x_{k+1})}{dx_{k+1}} \right|. \tag{5.10}$$

If $g(x)$ is not injective, then there is more than one solution to (5.3). In this case g is a non-monotonic function. Thus, we must obtain all the roots, $x_k^{(i)}$, of $x_{k+1} = g(x_k)$ so that

$$u_{X_{k+1}}(x_{k+1}) = \sum_i \frac{u_{X_k}(x_k^{(i)})}{|g'(x_k^{(i)})|}. \tag{5.11}$$

Equation (5.10) and (5.11) are the results of applying the theory of random variables to FDEs where the dynamics are monotonic and non-monotonic respectively. It is clear that using this method we are able to obtain the pdf of the future state for any time $k+1 \in \mathbb{N}$. The pdf $u_{X_{k+1}}(x_{k+1})$ is defined within the probability structure at time $k+1$ and renders the evolution operator measure preserving.

5.4 Higher Dimensional Finite Difference Equations

First, we consider 2-D FDEs. Let is be given the following FDE:

$$X_{k+2} = g(X_{k+1}, X_k), \tag{5.12}$$

where the initial data is the pdfs u_{X_1} and u_{X_0}. In order to find a solution to the uncertainty propagation problem for 2-D FDEs we must consider the theory on functions of two random variables.

In the theory of functions of two random variables, in general, we are given

$$Z = g(X, Y), \tag{5.13}$$

where X, Y, and Z are random variables. Now, let X and Y be continuous with pdfs

$u_X(x)$ and $u_Y(y)$ respectively. We use the distribution of Z, which is given by

$$U_Z(z) = P\{Z \leq z\}, \tag{5.14}$$

to determine the pdf of Z, $u_Z(z)$. We know from probability theory that $u_Z(z) = \frac{d}{dz}U_Z(z)$. Equation (5.14) tells us that in order to find U_Z we need to find the probability of the event $\{Z \leq z\}$. Since we want to consider the region in the xy-plane such that $g(x, y) \leq z$ and this region may not be simply connected, we define D_z to be the set

$$D_z := \{(X, Y)|g(x, y) \leq z\}. \tag{5.15}$$

Now, we find the probability mass in this region and this mass is given by

$$U_Z(z) = P\{(X, Y) \in D_z\} = \iint_{D_z} u_{XY}(x, y)dxdy. \tag{5.16}$$

The pdf $u_Z(z)$ is the joint density function, $u_{XY}(x, y)$, of X and Y, and can be found by differentiating U_Z. An alternative method is to determine the region ΔD_z which is the region in the xy-plane such that

$$z < g(x, y) \leq z + dz. \tag{5.17}$$

Explicitly, we define $\Delta D_z := \{(X, Y)|z < g(x, y) \leq z + dz\}$. Thus,

$$u_Z(z)dz = P\{z < Z \leq z + dz\} = \iint_{\Delta D_z} u_{XY}(x, y)dxdy. \tag{5.18}$$

Application of Leibiniz's rule allows us to compute u_Z explicitly.

Theorem 5.4.1. *(Leibinz's Rule) Given that a function f and the partial derivative $\frac{\partial f}{\partial x}$ are both continuous over the region $[x_1, x_2] \times [y_1, y_2]$, we have*

$$\frac{d}{dx}\int_{y_1}^{y_2} f(x,y)dy = \int_{y_1}^{y_2} \frac{\partial}{\partial x} f(x,y)dy. \tag{5.19}$$

Let us take an example to see how this computation is carried out. Consider the function $Z = X + Y$. The region D_z is the set $\{(X,Y)|x + y \leq z\}$. Thus,

$$U_Z(z) = \int_{-\infty}^{\infty}\int_{-\infty}^{z-y} u_{XY}(x,y)dxdy. \tag{5.20}$$

We need to differentiate in order to compute u_Z. This means we will need to apply Leibinz's rule. The result is

$$u_Z(z) = \int_{-\infty}^{\infty} u_{XY}(z-y,y)dy. \tag{5.21}$$

The case when the random variables X and Y are independent gives us nice results.

Theorem 5.4.2. *Let X and Y be two independent random variables. Then the density of their sum, $Z = X + Y$ equals the convolution of their respective densities [3].*

$$u_Z(z) = \int_{-\infty}^{\infty} u_X(z-y)u_Y(y)dy = \int_{-\infty}^{\infty} u_X(x)u_Y(z-x)dx \tag{5.22}$$

From the theory for functions of two random variables, we can generalize to func-

tions of multiple random variables. For example, given a function of three random variables, $W = g(X, Y, Z)$, we can compute the pdf $u_W(w)$ using the following integral:

$$u_W(w)dw = P\{w < W \leq w + dw\} = \iiint_{\Delta D_w} u_{XYZ}(x, y, z)dxdydz, \tag{5.23}$$

where $D_w = \{(X, Y, Z)|g(x, y, z) \leq w\}$.

Now, let us recall the FDE in equation (5.12) with the initial data u_{X_1} and u_{X_0}. We will use the theory of functions of random variables to determine the pdf of future states. We define $\Delta D_{x_{k+2}} := \{(X_k, X_{k+1})|x_{k+2} < g(x_k, x_{k+1}) \leq x_{k+2} + dx_{k+2}\}$. Thus,

$$u_{X_{k+2}}(x_{k+2}) = \frac{d}{dx_{k+2}} \iint_{\Delta D_{x_{k+2}}} u_{X_k X_{k+1}}(x_k, x_{k+1})dx_k dx_{k+1}. \tag{5.24}$$

5.5 Systems of Finite Difference Equaitons

For systems of FDEs, we again use functions of random variables. For the sake of simplicity, we will only consider systems of two FDEs. Consider the system

$$\begin{bmatrix} X_{k+1} \\ Y_{k+1} \end{bmatrix} = \begin{bmatrix} g(X_k, Y_k) \\ h(X_k, Y_k) \end{bmatrix} \tag{5.25}$$

with initial data given as the joint density function of random variables X_0 and Y_0, $u_{X_0Y_0}$. To solve this we need to develop the theory pertaining to two functions of two random variables.

Define the system to be

$$\begin{bmatrix} Z \\ W \end{bmatrix} = \begin{bmatrix} g(X,Y) \\ h(X,Y) \end{bmatrix}. \tag{5.26}$$

The joint density and distribution associated with W and Z are $u_{WZ}(w,z)$ and $U_{WZ}(w,z)$ respectively. We know the joint density of X and Y. It is given as $u_{XY}(x,y)$. As in the previous section, we must find the region such that $g(x,y) \leq z$ and $h(x,y) \leq w$.

$$D_{wz} := \{(X,Y) | g(x,y) \leq z, h(x,y) \leq w\} \tag{5.27}$$

The probability of the event $\{(W \leq w) \cap (Z \leq z)\}$ is

$$F_{WZ}(w,z) = P\{(W \leq w) \cap (Z \leq z)\} = \iint\limits_{D_{wz}} u_{XY}(x,y)dxdy. \tag{5.28}$$

Now, in order to find the pdf u_{WZ} we must use the Jacobian.

$$J(x,y) := \begin{vmatrix} \frac{\partial g(x,y)}{\partial x} & \frac{\partial g(x,y)}{\partial y} \\ \frac{\partial h(x,y)}{\partial x} & \frac{\partial h(x,y)}{\partial y} \end{vmatrix} \tag{5.29}$$

Equation (5.29) is the Jacobian of the system (5.26). The following theorem tells us how we can express the density function in terms of the Jacobian and the distribution [3].

Theorem 5.5.1. *Given that $u_{XY}(x,y)$ is finite and the members of the sequence*

$\{(x_n, y_n)\}$ are all real solutions to the system

$$\begin{bmatrix} z \\ w \end{bmatrix} = \begin{bmatrix} g(X,Y) \\ h(X,Y) \end{bmatrix}, \tag{5.30}$$

then

$$u_{WZ}(w,z) = \frac{u_{XY}(x,y)}{|J(x_1,y_1)|} + \cdots + \frac{u_{XY}(x,y)}{|J(x_n,y_n)|} + \cdots \tag{5.31}$$

If the system (5.30) has no real solutions, then $u_{WZ}(w,z) = 0$.

A proof of this theorem can be found in [3]. The theorem has a condition which needs further explanation. It says that we require u_{XY} to be finite. By this we mean that there are no point or line masses. This is consistent with our probability structure and they way we define the probability measure. Thus, this condition is met and in our application we are free to use this theorem.

Now, we return to the system of FDEs and apply the theory for two functions of two random variables. Given the system in equation (5.25) and the initial data $u_{X_0Y_0}$ which is finite, applying theorem (5.5.1) we get

$$u_{X_{k+1}Y_{k+1}}(x_{k+1}, y_{k+1}) = \frac{u_{X_kY_k}(x_k,y_k)}{|J(x_k^{(1)}, y_k^{(1)})|} + \cdots + \frac{u_{X_kY_k}(x_k,y_k)}{|J(x_k^{(n)}, y_k^{(n)})|} + \cdots \tag{5.32}$$

for the sequence of solutions $\{(X_k^{(n)}, Y_k^{(n)})\}$ at time $k \in \mathbb{N}$. The result is that for any future time $k+1 \in \mathbb{N}$, using iteration we are able to find the pdf $u_{X_{k+1}Y_{k+1}}$. In this way we are able to evolve the initial uncertainty. The pdf given in equation (5.32) is defined within $(\Omega, \mathcal{L}, P)_k$ and is consistent with our initial probability structure. It is

constructed in such a way that the method we use to propagate uncertainty is made a measure preserving map. This leads us to a discussion on the conservation form.

5.6 Conservation Form

We now consider how this method of uncertainty propagation is consistent with the conservation principle. The conservation method can be applied to finite difference equations to propagate uncertainty. We write uncertainty propagation FDEs in conservation form as follows:

$$\int_x^{x+\Delta x} u_{X_{k+1}}(x_{k+1})dx_{k+1} - \int_x^{x+\Delta x} u_{X_k}(x_k)dx_k$$

$$= \int_{g^{-1}(x)}^{g^{-1}(x+\Delta x)} u_{X_{k-1}}(x_{k-1})dx_{k-1} - \int_x^{x+\Delta x} u_{X_{k+1}}(x_{k+1})dx_{k+1}. \tag{5.33}$$

The heuristic explanation for equation (5.33) is as follows. The left-hand side of

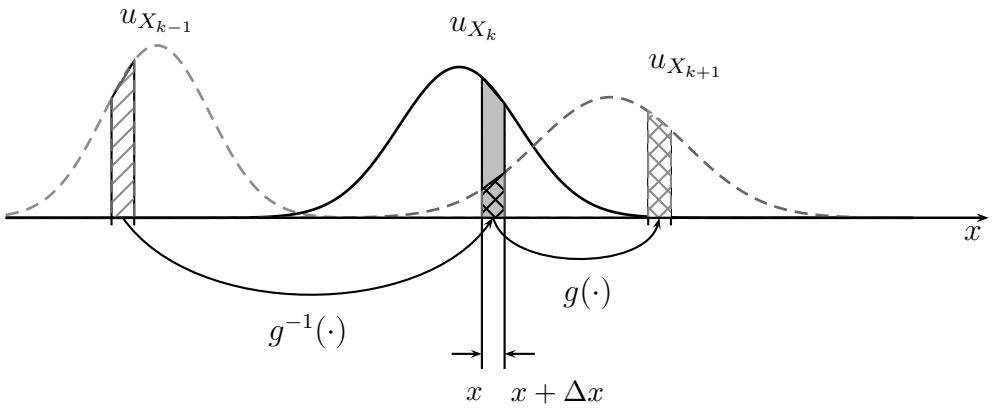


Figure 5.1: FDE Conservation

the equation is the mathematical formulation of the change in mass on the interval $[x, x + \Delta x]$ in one time step, k to $k + 1$ which we denote as Δk. This should equal

the flow of mass out of the interval $[x, x + \Delta x]$ at time k less the flow of mass into the same interval. This describes the right-hand side of equation (5.33). The flow of mass into the interval is the mass under the curve $u_{X_{k-1}}$ and we determine this mass by integrating over the interval $[g^{-1}(x), g^{-1}(x + \Delta x)]$ where g^{-1} denotes the inverse of the system dynamics. It is clear that the probability measure is preserved since the probability spaces at each time step are consistent with one another. Also, the conservation law as stated in equation (5.33) holds.

5.7 Summary

In summary, we have developed a theory for how to propagate uncertainty in the initial condition through one dimensional FDEs, multiple dimensional FDEs, and systems of FDEs . The theory of functions of random variables is the basis for the method of uncertainty propagation and is not original work. The contribution of this chapter is the application of the theory of functions of random variables to the problem of uncertainty propagation in FDEs as well as the development of the conservation form for the problem of uncertainty propagation in FDEs. The method developed is consistent with the probability structure defined in accordance with the initial data and it preserves the probability measure.

CHAPTER 6

DIFFERENTIAL INEQUALITIES AND INCLUSIONS

The dynamics of evolving processes is often subject to perturbations. Including uncertainties and disturbances into the mathematical model inevitable leads to set-valued maps and inclusions. Control theory and the evolution of macro-systems under uncertainty, constitute very strong motivations for extending differential equations to differential inclusions and inequalities [39].

Differential inequalities and inclusions are dynamic systems where the vector fields are given by set-valued maps. Set-valued maps are, by nature, ill-posed problems or inverse problems because the uniqueness requirement for well-posedness cannot, in general, be met.

In this chapter we define what differential inclusions and inequalities are in general, and then, we develop a previously nonexistent theory for propagating uncertainty in a small class differential inclusions defined by an important class of differential equations.

6.1 Set-Valued Maps

Set-valued maps have many of the same properties as functions do and thus it is important to note them. Essentially, set-valued maps violate the well-posed principle, i.e. they violate either existence of a solution or its uniqueness. For example, let us say we have a function which describes how the state of a system evolves:

$$\dot{x} = f(x). \tag{6.1}$$

Often, it may be the case that the inverse of $f : X \to 2^X$ does not exist, but we can always define a set of solutions for f^{-1}:

$$f^{-1}(\dot{x}) := \{x \in X | f(x) = \dot{x} \in 2^X\}, \tag{6.2}$$

where f^{-1} is a symbolic representation for the inverse function which may not exist. In this case, we say symbolically $f^{-1} : 2^X \to 2^X$. In the study of set-valued maps, the third condition of the well-posed problem is retained, i.e. the 'stability' condition. This is the condition that the set of solutions depend continuously on the initial data.

6.2 What are Differential Inclusions?

Differential inclusions are models of dynamic systems in which the velocities are not uniquely determined by the state of the system. Consider the differential equation in Banach space given by

$$\dot{x} = f(x, t) \tag{6.3}$$

where X is a Banach space, $A \subset X$, f maps $A \times \mathbb{R}^+$ to the tangent space of X, and $t \in \mathbb{R}^+$. It is often the case where the function f on the right-hand side of (6.3) may be multi-valued. Problems of this type are systems of differential inclusions in Banach space. Systems of this type take the form

$$\dot{x}(t) \in F(x, t), \tag{6.4}$$

where F is a set-valued function with subsets as its images in the Banach space.

6.2.1 What are Differential Inequalities?

Differential inequalities are a special class of dynamical systems determined by differential inclusions and they take the form

$$\dot{x}(t) \in F(x,t), \tag{6.5}$$

where $F(x,t) := [f^m(x,t), f^M(x,t)]$ with $f_m \leq f^M$, $x \in \mathbb{R}^n$, $f^m \in C[\mathbb{R}^n \times \mathbb{R}^+, \mathbb{R}^n]$, $f^M \in C[\mathbb{R}^n \times \mathbb{R}^+, \mathbb{R}^n]$, and $t \in \mathbb{R}^+$. The inequalities are evaluated pointwise.

We call a function $\phi \in C[[t_0, t_1), G]$, $[t_0, t_1) \subset T$, a solution of (6.5) if for all $t \in [t_0, t_1)$,

1. $\dot{\phi}(t) = \frac{\partial \phi}{\partial t}(t)$ exists and
2. $f^m(\phi(t), t) \leq \dot{\phi}(t) \leq f^M(\phi(t), t)$.

We are guaranteed solutions to the differential inequality given in (6.5) since for any $f \in C[T \times G, \mathbb{R}^n]$ satisfying

$$f^m(x,t) \leq f(x,t) \leq f^M(x,t) \;\; \forall (x,t) \in T \times G, \tag{6.6}$$

any solution of the equation $\dot{x}(t) = f(x,t)$ is clearly a solution to (6.5).

6.3 Uncertainty Propagation through differential inclusions

Propagation of uncertainty through differential inclusions, as presented here, will be for a particular class of differential inclusions where the boundary functions f^M

and f_m are of the form

$$f^i(x) = -\alpha_i x, \tag{6.7}$$

for $\alpha_i > 0$ and where $i \in \{m, M\}$. Given a differential inclusion of the form

$$\dot{x} \in F(x,t) \tag{6.8}$$

where $F(x,t) = [f_m(x,t), f^M(x,t)]$ for some $f_m(x,t) \leq f^M(x,t)$, the initial condition is given as a random variable $X(0) = X_0$ defined on the appropriate probability space $(\Omega_0, \mathcal{L}, P_0)$, and the inequalities are defined pointwise. The σ-algebra $\mathcal{L}$ is the σ-algebra of Lebesgue measurable sets defined on $\mathbb{R}$. The random variable $X_t(x)$ is the measurable function such that $\mathbb{R} \mapsto \mathbb{R}$ and it is the random variable describing the state of the system. We must define a random variable of this type so that we can define the density function for the state at future times that is consistent with our probability structure.

In this thesis, we study differential inclusions where the initial probability measure for the random variable corresponding to the initial state, X_0 is defined on a countable point-set which we denote as $\mathcal{I}_0$.

$$\mathcal{I}_0 := \{x_1, x_2, \ldots, x_n, \ldots\} \tag{6.9}$$

The initial density function is defined on the points in $\mathcal{I}_0$ so that

$$\sum_{i=1}^{\infty} P(x_i) = 1, \tag{6.10}$$

where $P(\cdot)$ is the probability measure.

In this thesis we study differential inclusions where the dynamics are of the form given in equation (6.7) and we know that the solutions to ODEs of that form are stable. Also, we know that if we are given two ODE problems,

$$\dot{x_1} = -\alpha x_1 \quad \text{with} \quad x_1(0) = a \tag{6.11}$$

and

$$\dot{x_2} = -\alpha x_2 \quad \text{with} \quad x_2(0) = b, \tag{6.12}$$

where $x_1(0) < x_2(0)$ and $\alpha > 0$, then the solutions have the property that $x_1(t) <$

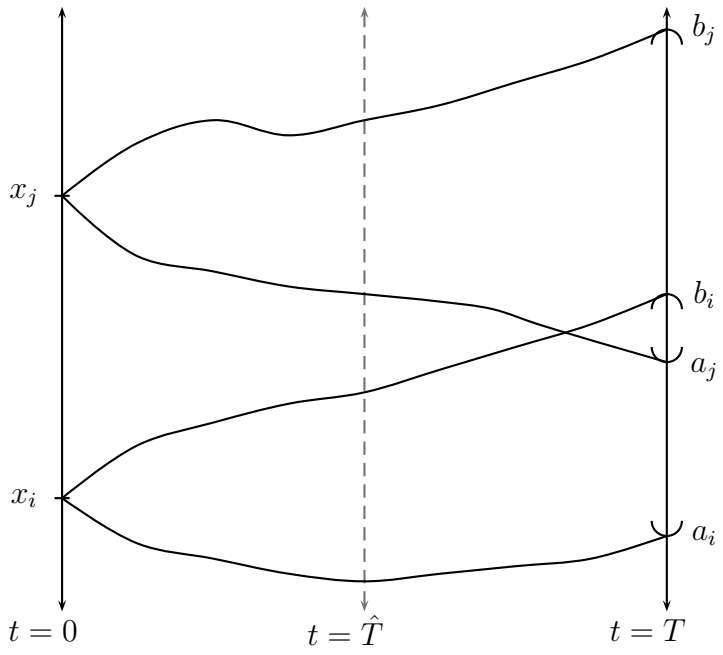


Figure 6.1: Trajectories generated by f^M and f_m for an initial point x.

$x_2(t)$ for all $t > 0$. Also, we know that given ODEs

$$\dot{x_1} = -\alpha x_1 \quad \text{with} \quad x_1(0) = a \tag{6.13}$$

and

$$\dot{x_2} = -\beta x_2 \quad \text{with} \quad x_2(0) = a, \tag{6.14}$$

where $0 < \alpha < \beta$, then the solutions have the property that $x_2(t) < x_1(t)$ for all $t > 0$. Thus, it is clear that we can evolve the system using the boundary functions f_m and f^M because all trajectories resulting from functions in the interval $[f_m, f^M]$ are in the cone enclosed by the trajectories resulting from the boundary functions. The way this cone is constructed is through evolving each point in the set $\mathcal{I}_0$ according to f_m and f^M. Refer to figure (6.1) for a conceptual sketch of this phenomena. Each point in $\mathcal{I}_0$ evolves in to interval. The intervals at time $T > 0$ are contained in the set $\mathcal{I}_T$ and are of the form $((a_i)_T, (b_i)_T)$ where the subscript T indicates the time index,

$$(b_i)_T = f^M(x_i), \tag{6.15}$$

$$(a_i)_T = f_m(x_i), \tag{6.16}$$

and $x_i \in \mathcal{I}_0$. The subscript notation i indicates that the interval $((a_i)_T, (b_i)_T)$ corresponds to the initial point x_i.

We take $f^M(x_i)$ to be b_i and we take $f_m(x_i)$ to be a_i since, by definition, f_m and f^M are of the form given in (6.7) with $\alpha_m < \alpha^M$. For $i, j \in \mathbb{N}$, $i \neq j$, the intervals

$(a_i, b_i)_T$ and $(a_j, b_j)_T$ are not necessarily disjoint.

Now, we must define the probability density $u_{X_T}(x)$ for some time $T > 0$. We define $u_{X_T}(x)$ as follows:

$$u_{X_T}(x) := \sum_i u^{(i)}_{X_T}(x), \tag{6.17}$$

where

$$u^{(i)}_{X_T} = \begin{cases} \frac{\beta_i}{(b_i)_T - (a_i)_T} & \text{if } x \in ((a_i)_T, (b_i)_T), \\ 0 & \text{otherwise} \end{cases} \tag{6.18}$$

6.3.1 Construction of Probability Space

Now, for the sake of simplicity we assume that the initial point-set contains only two elements. Define the initial point set to be

$$\mathcal{I}_0 := \{x_1, x_2\}. \tag{6.19}$$

The theory developed can be extended to a countable set. The initial density function $u_{X_0}(x)$ over the initial set is presented in figure (6.2). We also have that $\beta_1 + \beta_2 = 1$.

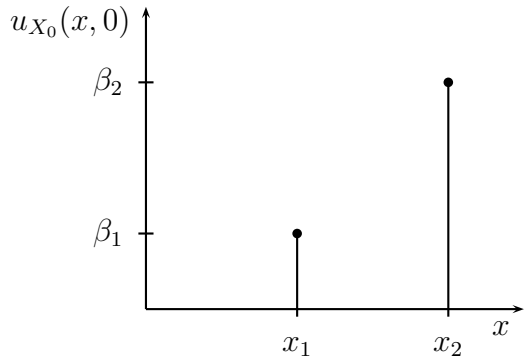


Figure 6.2: Initial Information: Density function for X_0.

The elements in the initial point-set evolve according to the system dynamics. $X_T^{(1)}$ and $X_T^{(2)}$ are random variables corresponding to the state of the system at time T given that the initial state x_0 is x_1 or x_2 respectively. Using the dynamics, we determine the interval at time T given the initial condition x_1 to be (a_T, b_T) where $a_T = f_m(x_1)$ and $b_T = f^M(x_1)$. For the initial condition x_2, the interval at time T is given by (c_T, d_T) is similarly determined. We define pdfs for $X_T^{(1)}$ and $X_T^{(2)}$ to be uniform density functions by distributing the initial probability over the interval. Figures (6.3) and (6.4) are examples of functions $u_{X_T^{(1)}}$ and $u_{X_T^{(2)}}$ corresponding to $X_T^{(1)}$ and $X_T^{(2)}$ respectively. Their sum is the density function such that it is consistent with the construction given in equations (6.17) and (6.18) and it satisfies all the usual properties of a density function, i.e.

1. Integration to unity:

$$\int_{\Omega_T} u_{X_T}(x)dx = \int 1_{\Omega_T} u_{X_T}(x)dx = 1,$$

where 1_{Ω_T} is the indicator function defined as

$$1_{\Omega_T}(x) = \begin{cases} 1 & \text{if } x \in \Omega_T \\ 0 & \text{otherwise} \end{cases} \tag{6.20}$$

2. $u_{X_T}(x) \geq 0 \ \ \forall x \in \mathbb{R}$.

Now, we must construct the probability structure at time T which satisfies all the desired properties of a probability space and is consistent with the initial probability

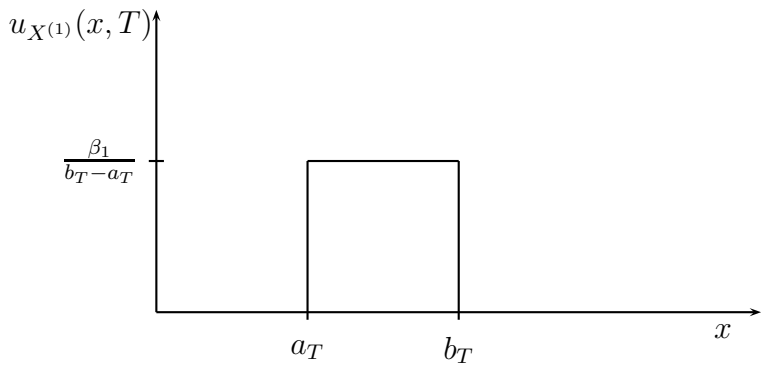


Figure 6.3: Density function for X_1 at time T_2.

structure and system dynamics. At time T, we define

$$u_{X_T}(x) := u_{X_T^{(1)}}(x) + u_{X_T^{(2)}}(x) \tag{6.21}$$

so that

$$P_T(\Omega_T) = \sum_{n=1}^{2} \int_{\Omega_T} u_{X_T^{(i)}}(x)dx = \sum_{n=1}^{2} \int 1_{\Omega_T} u_{X_T^{(i)}}(x)dx = 1, \tag{6.22}$$

where P_T is the probability measure at time T and $(\Omega_T, \mathcal{B}_T, P_T)$ is the probability space. For $T > 0$ and for each i, we define $\Omega_T^{(i)}$ to be $\mathbb{R}$ since $x \in \mathbb{R}$ where x is the state of the system. Also, the dynamics are defined so that $\mathbb{R} \mapsto \mathbb{R}$. Now, for time $T > 0$ and for each i, we take $\mathcal{B}_T$ to be the Borel σ-algebra defined on $\mathbb{R}$ since we desire that our evolution operator be a measurable function. In order for the operator to be measure preserving, i.e. consistent with the conservation principle, it needs to be a measurable function. Since the inverse of any Borel set is measurable in $\mathcal{L}$, the event space, the appropriate σ-algebra to take at time T is the Borel σ-algebra $\mathcal{B}_T$.

Definition 6.3.1. *(Measurable Function) A real-valued function is said to be a mea-*

surable function if the preimage of each Borel set is measurable.

We define the probability measure P_T as follows:

Definition 6.3.2. *Given a set $A \in \mathcal{B}_T$,*

$$P_T(A) := \int_A u_{X_T}(x)dx. \tag{6.23}$$

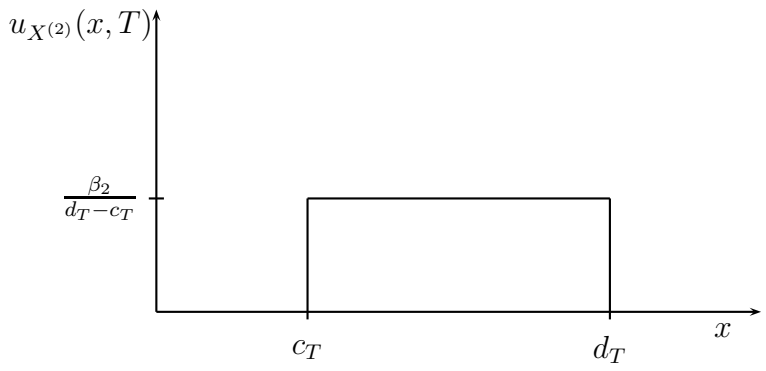


Figure 6.4: Density function for X_2 at time T_2.

In order for $(\Omega_T, \mathcal{B}_T, P_T)$ to be a proper probability space we must prove that P_T is a probability measure and that is is defined on the appropriate σ-algebra. Lemma (6.3.1) allows us to claim that the σ-algebras corresponding to the elements in the initial point-set are equivalent to one another.

Lemma 6.3.1. *For a fixed time T, the σ-algebra corresponding to the probability measure P_i^T and the σ-algebra corresponding to the probability measure P_j^T are equivalent.*

Proof. The proof of Lemma (6.3.1) follows directly from how we defined the probability space at time T. Recall that at time T, the σ-algebra corresponding to each point in the initial point-set is $\mathcal{B}_T$. □

Lemma (6.3.1) can be applied to any countable set $\mathcal{I}$. Now, we show that P_T is a probability measure.

Theorem 6.3.1. *Given the probability space $(\Omega_T, \mathcal{B}_T, P_T)$ at time T, the measurable function, P_T, is a probability measure satisfying all the properties of a probability measure.*

1. *$P_T(A) \geq 0$ for all $A \in \mathcal{B}_T$.*

2. *$P_T(\Omega_T) = 1$ and $P_T(\emptyset) = 0$.*

3. *If A_1 and A_2 are disjoint events in $\mathcal{B}_T$, then*

$$P_T(A_1 \cup A_2) = P_T(A_1) + P_T(A_2). \tag{6.24}$$

Proof. 1. Pick an arbitrary element $A \in \mathcal{B}_T$. Since $u_{X_T}(x) \geq 0$ for all $x \in \mathbb{R}$,

$$P_T(A) = \int_A u_{X_T}(x)dx = \int 1_A u_{X_T}(x)dx \geq 0. \tag{6.25}$$

2. $P_T(\Omega_T) = \int_{\Omega_T} u_{X_T}(x)dx = 1$ and $P_T(\emptyset) = \int_{\emptyset} u_{X_T}(x)dx = 0$. Note that the complement of the whole space is the empty set, i.e. $(\Omega_T)^c = \emptyset$, so that $P_T(\emptyset) = 1 - P_T(\Omega_T) = 0$.

3. Pick arbitrary disjoint sets $A_1, A_2 \in \mathcal{B}_T$.

$$P_T(A_1 \cup A_2) = \int_{A_1 \cup A_2} u_{X_T}(x)dx \tag{6.26}$$

$$= \int_{A_1} u_{X_T}(x)dx + \int_{A_2} u_{X_T}(x)dx - \int_{A_1 \cap A_2} u_{X_T}(x)dx \tag{6.27}$$

$$= \int_{A_1} u_{X_T}(x)dx + \int_{A_2} u_{X_T}(x)dx \quad \text{since } A_1 \cap A_2 = \emptyset \tag{6.28}$$

$$= P_T(A_1) + P_T(A_2). \tag{6.29}$$

□

Now, we can extend the above theorem and lemma to differential inclusions where the initial set is a countable set.

Theorem 6.3.2. *Given the probability space $(\Omega_T, \mathcal{B}_T, P_T)$ at time T, P_T is σ-additive: If $\{A_n, n \geq 1\}$ are events in $\mathcal{B}_T$ that are disjoint, then*

$$P_T\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P_T(A_n). \tag{6.30}$$

Proof. Pick an arbitrary set of disjoint events $\{A_n, n \geq 1\}$ in $\mathcal{B}_T$. By the properties of the indicator function for disjoint joint sets we have

$$1_{\dot{\cup}_{n=1}^{\infty} A_n} = 1_{A_1} + 1_{A_2} + \cdots \tag{6.31}$$

Thus,

$$P_T\left(\bigcup_{n=1}^{\infty} A_n\right) = \int 1_{\cup_{n=1}^{\infty} A_n} u_{X_T}(x)dx \tag{6.32}$$

$$= \int \left[\left(1_{A_1} + 1_{A_2} + \cdots\right) u_{X_T}(x)\right] dx \tag{6.33}$$

$$= \int 1_{A_1} u_{X_T}(x)dx + \int 1_{A_2} u_{X_T}(x)dx + \cdots \tag{6.34}$$

$$= \sum_{n=1}^{\infty} \int 1_{A_n} u_{X_T}(x)dx \tag{6.35}$$

$$= \sum_{n=1}^{\infty} P_T(A_n). \tag{6.36}$$

□

We have constructed our probability space $(\Omega_T, \mathcal{B}_T, P_T)$ and show that it satisfies all the necessary properties including that P_T is a proper probability measure. In fact, it is clear that by the construction the event space and the σ-algebra for each time T are equivalent (refer to Lemma (6.3.1), and thus, for all time $t > 0$ we denote the probability space to be simply $(\Omega, \mathcal{B}, P_T)$, dropping the subscript T on the event space Ω and the σ-algebra.

Using the probability structure and the dynamics, an uncertain initial condition, given as a probability distribution over a set of points, can be propagated to future times. This means that for some interval (x_1, x_2) we can determine the probability that the state x will be in this interval at any future time t by calculating $P(x_1, x_2) = P_T(x_1, x_2)$ where P_T is defined in definition (6.3.2).

6.4 Conservation Form

The problem of uncertainty propagation in differential inclusions of the form studied in this thesis can be written in conservation form. Consider the interval $(x, x+\Delta x)$ at time T. We want to determine the change in probability mass in the interval $(x, x + \Delta x)$ for the time step Δt. The change in probability mass in the interval $(x, x + \Delta x)$ is then

$$P_{T+\Delta t}(x, x + \Delta x) - P_T(x, x + \Delta x) = \int_{(x,x+\Delta x)} u_{X_{T+\Delta T}}(x)dx - \int_{(x,x+\Delta x)} u_{X_T}(x)dx \tag{6.37}$$

This will be equal to the flow into of the interval $(x, x + \Delta x)$ subtract the flow out the interval $(x, x + \Delta x)$. First, we denote the evolution operator as $g(\cdot)$. $g(\cdot)$ maps intervals on $\mathbb{R}$ to intervals on $\mathbb{R}$. The flow into $(x, x + \Delta x)$ is given by

$$P_{T-\Delta T}(g^{-1}(x, x + \Delta x)) = \int_{g^{-1}(x,x+\Delta x)} u_{X_{T-\Delta T}}(x)dx. \tag{6.38}$$

The flow out is given by

$$P_{T+\Delta T}(x, x + \Delta x) = \int_{(x,x+\Delta x)} u_{X_{T+\Delta T}}(x)dx. \tag{6.39}$$

Thus, the conservation form is

$$\int_{(x,x+\Delta x)} u_{X_{T+\Delta T}}(x)dx - \int_{(x,x+\Delta x)} u_{X_T}(x)dx$$

$$= \int_{g^{-1}(x,x+\Delta x)} u_{X_{T-\Delta T}}(x)dx - \int_{(x,x+\Delta x)} u_{X_{T+\Delta T}}(x)dx. \tag{6.40}$$

6.5 Summary

In this chapter we construct the theory for propagation of uncertainty in the initial condition for a particular class of differential inclusions. We present the case where the initial condition is a random variable with a discrete density function defined over an initial point-set. Although the class of functions we study here is only a small portion of all possible differential inclusions, it is an important one since the ODEs that generate the differential inclusions in this class are common in many applications including electronic circuits and population growth modeling. Using probability theory, we present a method for uncertainty propagation in systems of this type. The work presented in this chapter is all original work and its contribution is the development of a previously nonexistent theory for uncertainty propagation in differential inclusions defined by an important class of differential equations.

CHAPTER 7

STOCHASTIC DIFFERENTIAL EQUATIONS

In this chapter we review the concept of stochastic differential equations and how they are used in mathematical modeling. We also develop the theory for how to propagate uncertainty in the initial condition through SDEs where this initial uncertainty is given within a probability structure. The method for propagating uncertainty is based on the conservation principle. We use the Fokker-Planck equation (FPE) as the method to propagate uncertainty through SDEs. The contribution of this chapter is presentation of the Fokker-Planck equation as a conservation based method of uncertainty propagation in SDEs.

7.1 What are Stochastic Differential Equations?

SDEs are mathematical models in which the dynamics themselves have uncertainty. SDEs are important in mathematical modeling of financial systems, weather modeling, and many other physical systems. In order to understand where the motivation for SDEs comes from let us consider the following system of ODEs:

$$\begin{cases} \dot{x}(t) = f(x(t)) \quad \forall t > 0 \\ x(0) = x_0, \end{cases} \tag{7.1}$$

where the dynamics $f : \mathbb{R}^n \to \mathbb{R}^n$ is a smooth function and the solution is the trajectory of the state, $x(\cdot) : \mathbb{R}^+ \to \mathbb{R}^n$. In many applications, the experimentally measured trajectories of the system modeled by a system of ODEs do not behave as predicted and instead we have a trajectory similar to the one in figure (7.1) [4]. In

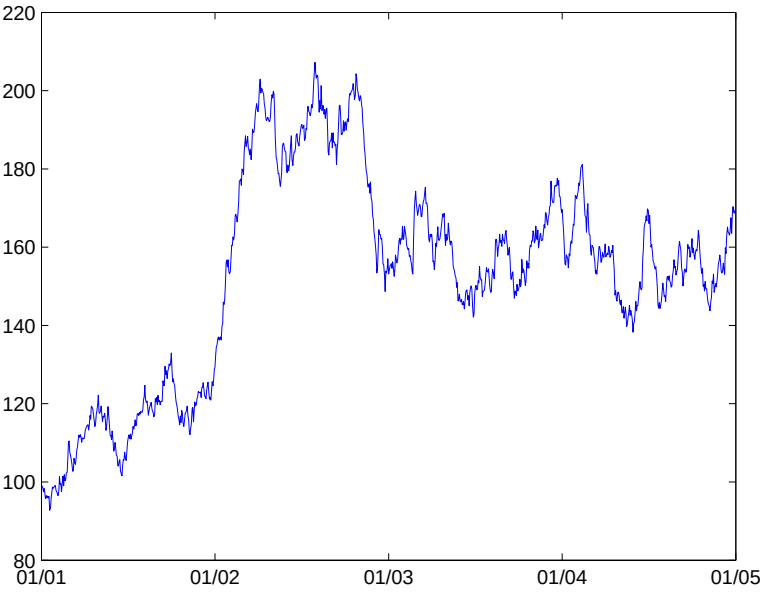


Figure 7.1: Path generated from exponential Brownian motion.

order to adapt for this in the mathematical model, we must modify the system in (7.1) to include the possibility of random perturbations in the system. Thus, we have the following:

$$\begin{cases} \dot{X}(t) = f(X(t)) + B(X(t))\xi(t) \quad \forall t > 0 \\ X(0) = x_0, \end{cases} \tag{7.2}$$

where X is a random variable, B maps $\mathbb{R}^n$ to a space of matrices denoted by $\mathbb{M}^{n\times m}$ and $\xi(\cdot) :=$ m-dimensional 'white noise'. We call equation (7.2) a SDE.

Let us consider the case where $m = n$, $f \equiv 0$, B is the $n \times n$ identity matrix which we denote as $I_{n\times n}$, and the initial condition is $x_0 = 0$. In this case, the solution to equation (7.2) is Brownian motion. The solution is also a n-dimensional Wiener process. We denote a Wiener process by $W(\cdot)$. The details of the Wiener process will be discussed in later sections. Now, we claim that the noise term, $\xi(t)$ is the time

derivative of the Wiener process,

$$dW_t = \xi(t). \tag{7.3}$$

We can now rewrite equation (7.2) in the form which we introduced in chapter 3. Recall

$$\begin{cases} dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dW_t \\ X(0) = x_0 \end{cases} \tag{7.4}$$

where X is a random variable, the subscript t indicates time dependence, $\sigma(\cdot, \cdot)$ and $\mu(\cdot, \cdot)$ are given functions, and dW_t is the noise term. We call $X(\cdot)$ a solution to (7.4) if it solves

$$X(t) = x_0 + \int_0^t \mu(X(s))ds + \int_0^t \sigma(X(s))dW \quad \forall t > 0. \tag{7.5}$$

Existence and uniqueness of solutions are guaranteed only under certain conditions. The following theorem provides the details for existence of solutions. The theorem was adapted from [31].

Theorem 7.1.1. *Let $\mu(\cdot, \cdot) : \mathbb{R}^n \times \mathbb{R}^+ \to \mathbb{R}^n$, $\sigma(\cdot, \cdot) : \mathbb{R}^n \times \mathbb{R}^+ \to \mathbb{R}^{n \times m}$ be measurable functions satisfying*

$$|\mu(x, t)| + |\sigma(x, t)| \leq C(1 + |x|) \quad x \in \mathbb{R}^n,\ t \in \mathbb{R}^+ \tag{7.6}$$

for some constant C such that

$$|\mu(x, t) - \mu(y, t)| + |\sigma(x, t) - \sigma(y, t)| \leq D \cdot |x - y| \quad x, y \in \mathbb{R}^n,\ t \in \mathbb{R}^+ \tag{7.7}$$

for some constant D. Note that

$$|\sigma|^2 = \sum_{i,j=1}^{n} |\sigma_{ij}|^2 \tag{7.8}$$

Let X_0 be a random variable such that

$$E\left[|X_0|^2\right] < \infty, \tag{7.9}$$

and such that X_0 is independent of the σ-algebra

$$\mathcal{F}_{W_s} = W_s^{-1}(\mathcal{B}) := \{W_s^{-1}(B) : B \in \mathcal{B}\}, \tag{7.10}$$

where $\mathcal{B}$ is the Borel σ-algebra on $\mathbb{R}$ and $s \geq 0$. Then the SDE in equation (7.4) has a unique solution X_t that is continuous in t.

Proof. The proof for this theorem can be found in chapter 5 of [31]. □

7.2 Propagation of Uncertainty through SDEs

We are given a SDE of the form

$$dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dW_t \tag{7.11}$$

where the initial condition, $X(0) = X_0$, is unknown but its statistical properties are known through its pdf $u_{X_0}(x)$, which is given within the probability structure $(\Omega, \mathcal{F}, P)$. There is uncertainty in the system dynamics, hence they are stochastic

in nature, and there is uncertainty in the initial condition. We desire a method for determining the pdf of the state, X, at future times. The method for doing this is the application of the FPE. The FPE can be used to compute the pdf of the state of a dynamic system described by a SDE. If we consider the Itô SDE

$$dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dW_t, \tag{7.12}$$

and the initial distribution is $u_{X_0}(x)$, then the pdf $u_{X_t}(x)$ of the state X_t is given by the FPE. In order to develop the theory for uncertainty propagation using the FPE, we must introduce some important background information. The material in the following sections on Itô calculus and Itô's formula is adapted from [4].

7.3 Itô Calculus

Itô calculus allows for the ideas and concepts in calculus to be extended to stochastic processes such as the Wiener process. A Wiener process is a continuous-time stochastic process characterized by

1. $W(0) = 0$ a.s.,

2. $W(t) - W(s) \sim \mathcal{N}(0, t-s) \;\; \forall t \geq s \geq 0$,

3. $\forall t_i$ such that $0 < t_1 < t_2 < \ldots < t_n$, the random variables $W(t_1), W(t_2) - W(t_1), \ldots, W(t_n) - W(t_{n-1})$ are independent.

Also, it is important to note that

$$E(W(t)) = 0 \ \ \forall t \geq 0 \tag{7.13}$$

and

$$E(W^2(t)) = t \ \ \forall t \geq 0, \tag{7.14}$$

where $E(X)$ is the expected value. Formally, if X is a random variable defined on a probability space $(\Omega, \mathcal{F}, P)$, then the expected value of X, denoted $E(X)$, is defined as

$$E(X) := \int_{\Omega} X dP, \tag{7.15}$$

where the integral is a Lebesgue integral, Ω is the sample space, $\mathcal{F}$ is the σ-algebra of subsets of Ω, and P is the measure on $(\Omega, \mathcal{F})$ such that $P(\Omega) = 1$. These definitions become important later in the derivation of the FPE. Now that we have defined a Wiener process, we will develop Itô's formula.

7.3.1 Itô's Formula

Given the SDE

$$dX = b(X)dt + dW_t, \tag{7.16}$$

where W_t is a one-dimensional Wiener process. Assume $X(\cdot)$ to be a solution to (7.16). Let $f : \mathbf{R} \to \mathbf{R}$ be a smooth function and define

$$Y(t) := f(X(t)) \quad (t \geq 0). \tag{7.17}$$

We want to examine the SDE that Y solves. First, we will assume that

$$dW \approx \sqrt{(}dt). \tag{7.18}$$

Thus, if we compute the derivative of Y, keeping all terms of the order dt and $dt^{\frac{1}{2}}$, then we get the following:

$$\begin{aligned} dY &= f'dX + \frac{1}{2}f''(dX)^2 + \dots && (7.19)\\ &= f'(bdt + dW) + \frac{1}{2}f''(bdt + dW)^2 + \dots && (7.20)\\ &= \left(f'b + \frac{1}{2}f''\right) + f'dW + \{\text{higher order terms}\}. && (7.21) \end{aligned}$$

Thus, we have

$$dY = \left(f'b + \frac{1}{2}f''\right)dt + f'dW. \tag{7.22}$$

There is an extra term, $\frac{1}{2}f''dt$, which is not present in ordinary calculus [4]. The above is a heuristic derivation which shows the existence of additional terms in computing the derivative in Itô calculus. This term is a direct result of equation (7.18), i.e. $(dW)^2 = dt$.

Another important piece of information from Itô calculus is Itô's chain rule which follows from Itô's formula.

Theorem 7.3.1. *Itô's Chain Rule Let $\mu \in L^1(\mathbb{R}^+)$ and $\sigma \in L^2(\mathbb{R}^+)$. Suppose that $X(\cdot)$ has the stochastic differential*

$$dX_t = \mu dt + \sigma dW_t. \tag{7.23}$$

Assume $f : \mathbb{R} \times \mathbb{R}^+ \to \mathbb{R}$, $\frac{\partial f}{\partial t}$, $\frac{\partial f}{\partial x}$, *and* $\frac{\partial^2 f}{\partial x^2}$ *exist and are continuous. Define*

$$Y(t) := f(X(t), t). \tag{7.24}$$

Then $Y(t)$ *has the stochastic differential*

$$dY_t = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dX_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\sigma^2 dt \tag{7.25}$$

$$= (\frac{\partial f}{\partial t} + \frac{\partial f}{\partial x}\mu + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\sigma^2)dt + \frac{\partial f}{\partial x}\sigma dW. \tag{7.26}$$

Let us consider an example. This example is adapted from a similar example in [4].

Example 7.3.1.

$$\begin{cases} dX = XdW \\ \\ X_0 = 1 \end{cases} \tag{7.27}$$

Applying Itô's formula, we find that the solution to (7.27) is $X(t) := \exp(W_t - \frac{t}{2})$.

This might be surprising since normally we would expect the solution to be $X(t) := \exp(W_t)$. Itô's formula tells us that we need to keep the additional terms to be consistent with $(dW)^2 = dt$. Using Itô's formula and conservation principles as applied before, we can derive the pdf for X_t to be the solution to the FPE.

7.4 Derivation of the Fokker-Planck Equation

In this section we will develop the FPE using ideas from Itô calculus. The derivation of the FPE as presented here follows the derivation in [5]. The FPE describes

the temporal evolution of the pdf of the state of a system. In order to derive the FPE, we start with the following SDE:

$$dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dW_t. \tag{7.28}$$

Let χ_A be the characteristic function defined by

$$\chi_{A(x)} := \begin{cases} 1 \text{ if } x \in A \\ \\ 0 \text{ if } x \notin A \end{cases} \tag{7.29}$$

where A is an element of the σ-algebra set of a probability space $(\Omega, \mathcal{F}, P)$. The following portion of the derivation was adapted from [5]. We assume χ is approximated by a smooth function. We can now take its derivative. We pick any Borel set $B \in \mathcal{F}$ so that

$$P(X_t \in B) = E[\chi_{B(X_t)}] = \int_B p(x, t)dx. \tag{7.30}$$

Taking the derivative of both sides of (7.30) and applying Itô's formula, we get

$$dP(X_t \in B) = E[\chi_{B(X_t)}]' \tag{7.31}$$

$$= E[\chi'_{B(X_t)}dX_t] + \frac{1}{2}E[\chi''_{B(X_t)}(dX_t)^2]. \tag{7.32}$$

Plugging in equation (7.28) for dX_t, the first term in the left-hand side of (7.32) we get

$$E[\chi'_{B(X_t)}dX_t] = E[\chi'_{B(X_t)}(\mu(X_t,t)dt + \sigma(X_t,t)dW_t)] \tag{7.33}$$

$$= E[\chi'_{B(X_t)}\mu(X_t,t)]dt, \tag{7.34}$$

since from the properties of the Wiener process given in (7.13) we have

$$E[\chi'_{B(X_t)}\sigma(X_t,t)dW_t] = 0. \tag{7.35}$$

The second term in the left-hand side of (7.32) becomes

$$\frac{1}{2}E[\chi''_{B(X_t)}(dX_t)^2] = \frac{1}{2}E[\chi''_{B(X_t)}(\mu(X_t,t)dt + \sigma(X_t,t)dW_t)^2] \tag{7.36}$$

$$= \frac{1}{2}E[\chi''_{B(X_t)}(\mu^2dt^2 + 2\mu\sigma dtdW_t + \sigma^2dW_t^2)] \tag{7.37}$$

$$= \frac{1}{2}E[\chi''_{B(X_t)}(\mu^2dt^2 + 2\mu\sigma dt^{3/2} + \sigma^2dt)]. \tag{7.38}$$

Dropping the higher order terms, i.e. $(dt)^{3/2}$ and higher, we get

$$\frac{1}{2}E[\chi''_{B(X_t)}(dX_t)^2] = \frac{1}{2}E[\chi''_{B(X_t)}\sigma^2dt]. \tag{7.39}$$

Plugging equations (7.34) and (7.39) into equation (7.32), we get

$$dP(X_t \in B) = E[\chi_{B(X_t)}\mu(X_t,t)]dt + \frac{1}{2}E[\chi''_{B(X_t)}\sigma^2(X_t,t)]dt \tag{7.40}$$

$$= \int \chi'_{B(x)}\mu(x,t)p(x,t)dx + \frac{1}{2}\int \chi''_{B(x)}\sigma^2(x,t)p(x,t)dx. \tag{7.41}$$

Integrating by parts, we can rewrite the first integral in the right-hand side of (7.41) as

$$\int \chi'(x \in B)\mu(x,t)p(x,t)dx \tag{7.42}$$

$$= \int \left\{ \frac{\partial}{\partial x}[\chi(x \in B)\mu(x,t)p(x,t)] - \chi(x \in B)\frac{\partial}{\partial x}[\mu(x,t)p(x,t)] \right\} dx \tag{7.43}$$

$$= -\int_B \frac{\partial}{\partial x}[\mu(x,t)p(x,t)]dx, \tag{7.44}$$

since $\chi(x \in B)\mu(x,t)p(x,t)$ vanishes at the boundary and $\chi_B = 1$ for $(x \in B)$. Similarly,

$$\int \chi''(x \in B)\sigma^2(x,t)p(x,t)dx \tag{7.45}$$

$$= \int \left\{ \frac{\partial}{\partial x}\left[\frac{\partial}{\partial x}[\chi(x \in B)]\sigma^2(x,t)p(x,t) \right] - \frac{\partial}{\partial x}[\chi(x \in B)]\frac{\partial}{\partial x}[\sigma^2(x,t)p(x,t)] \right\} dx \tag{7.46}$$

$$= -\int \frac{\partial}{\partial x}[\chi(x \in B)]\frac{\partial}{\partial x}\sigma^2(x,t)p(x,t)dx \tag{7.47}$$

$$= -\int \left\{ \frac{\partial}{\partial x}\left[\chi(x \in B)\frac{\partial}{\partial x}\sigma^2(x,t)p(x,t) \right] - \chi(x \in B)\frac{\partial^2}{\partial x^2}[\sigma^2(x,t)p(x,t)] \right\} dx \tag{7.48}$$

$$= \int_B \frac{\partial^2}{\partial x^2}[\sigma^2(x,t)p(x,t)]dx. \tag{7.49}$$

Thus, we have

$$\frac{d}{dt}P(X_t \in B) = \int_B \frac{\partial}{\partial t}p(x,t) = \int_B \left\{ -\frac{\partial}{\partial x}[\mu(x,t)p(x,t)] + \frac{\partial^2}{\partial x^2}[\sigma^2(x,t)p(x,t)] \right\} dx. \tag{7.50}$$

We know that if we have $\int_B f(x)dx = \int_B g(x)dx$ for any Borel set B, then f=g a.e. Therefore, since B is arbitrary we get the FPE. In one dimension the FPE is

$$\frac{\partial}{\partial t}u(x,t) = -\frac{\partial}{\partial x}[D_1(x,t)u(x,t)] + \frac{\partial^2}{\partial x^2}[D_2(x,t)u(x,t)], \tag{7.51}$$

where $u(x,t)$ is the probability density function for the state X_t and $D_1(x,t)$ and $D_2(x,t)$ are drift and diffusion terms respectively. The general form of the FPE is

$$\frac{\partial u}{\partial t} = -\sum_{i=1}^{N} \frac{\partial}{\partial x_i}\left[D_i^1(x_1,\ldots,x_N)u\right] + \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{\partial^2}{\partial x_i \partial x_j}\left[D_{ij}^2(x_1,\ldots,x_N)u\right], \tag{7.52}$$

where D^1 and D^2 are the drift vector and the diffusion tensor respectively. Since the FPE is a partial differential equation, it can only be solved analytically in a few cases. In order to find solutions, we must solve the FPE numerically.

7.5 Conservation Form

The FPE can be written as a conservation law. We will write the one-dimensional FPE in conservation form by defining

$$J := D_1(x,t)u(x,t) - \frac{\partial}{\partial x}D_2(x,t)u(x,t). \tag{7.53}$$

Often J is called the probability current. Thus, the conservation form is

$$\frac{\partial}{\partial t}u(x,t) + \frac{\partial}{\partial x}J = 0. \tag{7.54}$$

The conservation form is direct result of the global conservation of probability. By this we mean

$$\int_{\Omega} u(x,t)dP = 1. \tag{7.55}$$

7.6 Summary

In summary, the FPE is a partial differential equation which allows us to propagate uncertainty in the initial condition through a SDE. The FPE has both diffusion and drift terms so that the initial pdf both shifts in space and diffuses. We can use the FPE to find the pdf of the state of the system, X_t. The contribution of this chapter is the presentation of the FPE as a conservation based method for uncertainty propagation in SDEs.

CHAPTER 8

MARKOV CHAINS

In this chapter we review the theory of Markov chains. We also present a method for propagating uncertainty through Markov chains. The contribution of this chapter is the application of transition probabilities to the problem of uncertainty propagation in dynamic systems characterized by Markov chains.

8.1 What are Markov Chains?

Recall from chapter one that a system is said to satisfy the Markov property if it has the property that given the present state, the past states have no influence on the future [33]. Markov chains satisfy the Markov property. In mathematical form, this is

$$P(X_{n+1} = x_{n+1} | X_0 = x_0, \ldots, X_n = x_n) = P(X_{n+1} = x_{n+1} | X_n = x_n), \tag{8.1}$$

where $P(\cdot)$ is the probability function and X_k is the state of the system at time k. In chapter one we did not formally define Markov chains, and thus, we do that here.

Let us consider a system satisfying the properties listed below [40].

1. The system can occupy a finite or countably infinite number of states. Let those states be denoted as $x^{(1)}, x^{(2)}, \ldots$

2. Let the random variable X_k, where k represents discrete time steps and takes values $k = 0, 1, \ldots$, represent the state of the system. Then the evolution of the

system in time is described by the transitions

$$X_0 \to X_1 \to X_2 \to \cdots \tag{8.2}$$

3. The probability that the state of the system is $x^{(i)}$ at time $k = 0$ is given by

$$p_i^0 = p_i(0) = \mathbf{P}(X_0 = x^{(i)}), \quad i = 1, 2, \ldots \tag{8.3}$$

4. Given that the system is in state $x^{(i)}$ at time n, the probability that the system is in state $x^{(j)}$ at time $n + 1$ is called the transition probability and is given by

$$p_{ij} = \mathbf{P}(X_{n+1} = x^{(j)} | X_n = x^{(i)}), \quad i, j = 1, 2, \ldots \tag{8.4}$$

A system satisfying these properties is a Markov chain.

8.2 Transition Probabilities

Transition probabilities are an important concept in probability theory. They are also important for understanding how uncertainty propagates in Markov chains. Using transition probabilities, we can determine, mathematically, the probability that the state of the system is $x^{(j)}$ after n time steps. We denote this probability as

$$p_j(n) = \mathbf{P}(X_n = x^{(j)}). \tag{8.5}$$

Now, since the system cannot occupy two states simultaneously and we know that the system had to be in some state $x^{(i)}$ where $i = 1, 2, \ldots$ at time $n-1$, it is clear that the set composed of all events $[X_n = x^{(i)}]$, $i = 1, 2, \ldots$ is a collectively exhaustive and mutually exclusive set of events. Thus, by the total probability formula,

$$\mathbf{P}(A) = \sum_i \mathbf{P}(A|B_i)\mathbf{P}(B_i) \tag{8.6}$$

where A is some event and $\{B_i\}$ is a collectively exhaustive and mutually exclusive set of events, we can determine the probability $p_j(n)$.

$$p_j(n) = \mathbf{P}(X_n = x^{(j)}) = \sum_i \mathbf{P}(X_n = x^{(j)}|X_{n-1} = x^{(i)})\mathbf{P}(X_{n-1} = x^{(i)}). \tag{8.7}$$

Transition probabilities allow us to determine the probabilities of future states of the system.

8.3 Uncertainty Propagation in Markov Chains

Given a system which can be described mathematically with a Markov chain and an initial probability distribution over the states of the system at time $n = 0$, we can determine the future probability density function using transition probabilities. Let us consider a simple example in order to understand the details.

Example 8.3.1. *Consider a finite state machine (FSM) with the state space* $\mathcal{L} = \{A, B, C, D\}$ *and the initial distribution given in figure (8.1). Let* X *be the random variable corresponding to the state of the system. Notice that* $\sum_{i=1}^{4} P(X = x^{(i)}) = 1$

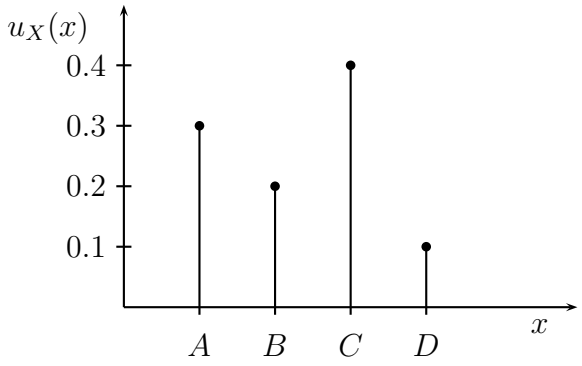


Figure 8.1: Initial distribution for the FSM

where $x^{(i)} \in \mathcal{L}$*. At time* $n = 1$*, using the transition probabilities, we can determine*

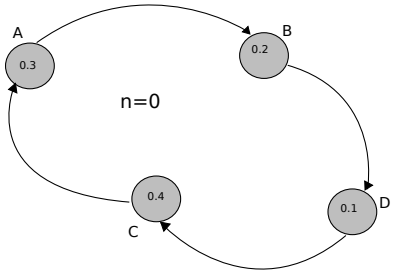


Figure 8.2: FSM at time $n = 0$.

the probability that the state will be $X_1 = 1$ *by calculating*

$$P(X_1 = A) = \sum_{i=1}^{4} P(X_0 = x^{(i)}) P(X_1 = A | X_0 = x^{(i)}) = 0.4. \tag{8.8}$$

Equivalently, we can calculate the probabilities $P(X_1 = B) = 0.3$*,* $P(X_1 = C) = 0.1$*, and* $P(X_1 = D) = 0.2$*.*

The previous example shows us that we can use the transition probabilities to propagate the initial uncertainty, namely the initial pdf, through a system described by a Markov chain using an iterative processes that depends on the transition probabilities.

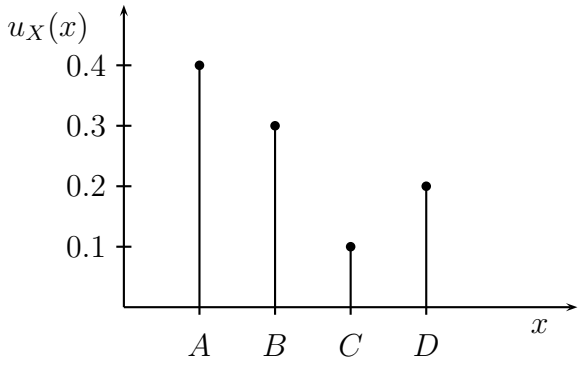


Figure 8.3: Distribution at $n = 1$.

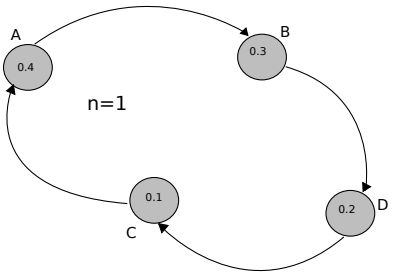


Figure 8.4: FSM at time $n = 1$.

8.4 Conservation Form

We can write the method for uncertainty propagation in Markov chains in conservation form by examining the change in the probability of a state, x, in one time step and equating that with the difference in the flow of 'mass' leaving x and the flow of 'mass' into the the state x. Consider state A in example (8.3.1). The conservation form is as follows:

$$P(X_{k+1} = A) - P(X_k = A) = \sum_x P(X_k = x)P(X_{k+1} = A|X_k = x)$$

$$- \sum_x P(X_{k-1} = x)P(X_k = A|X_{k-1} = x), \tag{8.9}$$

where $x \in \mathcal{L}$. This can be rearranged to

$$P(X_{k+1} = A) + \sum_x P(X_{k-1} = x)P(X_k = A|X_{k-1} = x)$$

$$- \left[P(X_k = A) + \sum_x P(X_k = x)P(X_{k+1} = A|X_k = x) \right] = 0. \qquad (8.10)$$

As in the case with SDEs, the conservation form follows directly from the global conservation of probability.

8.5 Summary

In this thesis we only present a simple example of a FSM in order to convey the concept of uncertainty propagation in systems which have the Markov property where the uncertainty is a given probability distribution over the states of the system and defined with in a probability structure. It is easy to see how this concept can be extended to more general Markov chains and in future work these details will be explored. The main idea here is that we can determine the future probability of the states, a posterior, of a system given some initial distribution, a prior, and that this method is consistent with conservation laws. In this sense, the contribution of this chapter is the application of the theory of Markov chains, specifically transition probabilities, to the problem of uncertainty propagation in dynamic systems which satisfy the Markov property and can be defined as Markov chains.

CHAPTER 9

UNCERTAINTY PROPAGATION IN BURGERS' EQUATION

In this chapter we look at how to propagate uncertainty in the initial condition through Burgers' equation. The uncertainty is stochastic and the is given as a uniform density function. We examine the uncertainty in the velocity variable at future times by evaluating the first and second moments. The contribution of this chapter is the development of the theory for uncertainty propagation in Burgers' equation.

9.1 Burgers' Equation

Burgers' equation is a quasilinear hyperbolic partial differential equation given by

$$\rho_t + \left(\frac{1}{2}\rho^2\right)_x = 0. \tag{9.1}$$

We define the flux function f as follows

$$f(\rho) = \left(\frac{1}{2}\rho^2\right). \tag{9.2}$$

The LWR model for traffic flow as presented in Chapter 3 gives a scalar conservation law with a quadratic flux function. Burgers' equation is a simpler version of the scalar conservation law. Solutions to Burgers' equation have the same basic structure as solutions to the traffic flow problem. We must allow for weak solutions because the nature of the partial differential equation is such that it has shock wave formation. At the point of discontinuity, a shock wave forms and travels forward in the case of Burgers' equation. Figure (9.2) dipicts the shock traveling. In general, Burgers'

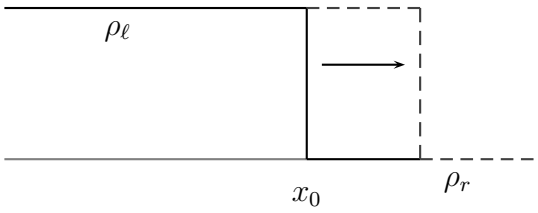


Figure 9.1: Solution after some time.

equation is not useful in practical applications, but a study of Burgers' equation informs many problems. For instance in [11], Pettersson, et al. state that Burgers' equation is of limited practical use in fluid mechanics applications, but due to the fact that it is highly nonlinear, the results of a study on Burgers' equation can be extended to other hyperbolic systems such as the Euler equations. Further information on Burgers' equation and its uses can be found in [41] and [42].

9.2 Uncertainty Propagation in Burgers' Equation

In this study, we consider Burgers' equation with piecewise constant initial data dependent on the uncertain parameter θ. The stochastic initial condition is then given by

$$\rho(x,0,\theta) = \begin{cases} \rho_\ell = a + b\theta & \text{if } x < x_0 \\ \rho_r = -a + b\theta & \text{if } x > x_0 \end{cases} \tag{9.3}$$

where x_0 is the location of the initial shock, $a > 0$ is a constant, and $b > 0$ is a constant. θ is the random variable representing the uncertainty parameter. The pdf

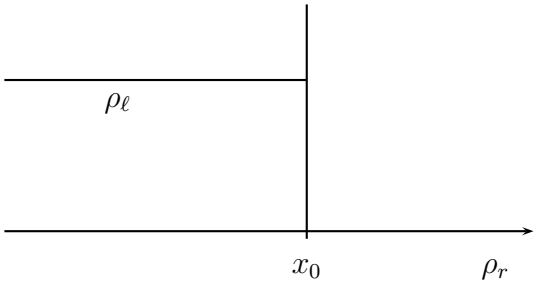


Figure 9.2: Initial Data.

for θ is a uniform density function and is given by

$$f_\theta(\theta) = \begin{cases} \frac{1}{B-A} & \text{if } A \leq \theta \leq B \\ 0 & \text{otherwise} \end{cases} \tag{9.4}$$

For a fixed θ, the initial discontinuity will travel with a shock speed of

$$s = \frac{f(\rho_\ell) - f(\rho_r)}{\rho_\ell - \rho_r} = \frac{1}{2}\left(\frac{(a+b\theta)^2 - (-a+b\theta)^2}{a+b\theta-(-a+b\theta)}\right) = b\theta. \tag{9.5}$$

For any given shock location x_s at time t_s, there exists a unique θ_s and it is given by

$$\theta_s(x,t) = \frac{x_s - x_0}{b(t_s - t_0)}. \tag{9.6}$$

For simplicity, we take $x_0 = 0$ and $t_0 = 0$. Thus, equation (9.6) becomes

$$\theta_s(x,t) = \frac{x_s}{b(t_s)}. \tag{9.7}$$

Using the relation in equation (9.6), we determine the analytical solution to Burgers' equation with the initial data in equation (9.3) to be

$$\rho(x,t,\theta) = \begin{cases} \rho_\ell = a + b\theta & \text{if } \theta > \theta_s \\ \rho_r = -a + b\theta & \text{if } \theta < \theta_s \end{cases} \tag{9.8}$$

Let us consider graphically the characteristics. θ is linearly related to the shock speed by the constant b, $s = \theta b$. The shock speed is related to the change in x over the change in t. Figure (9.3) shows the characteristics. As θ_s varies between A and B it is clear that the shock speed will vary accordingly.

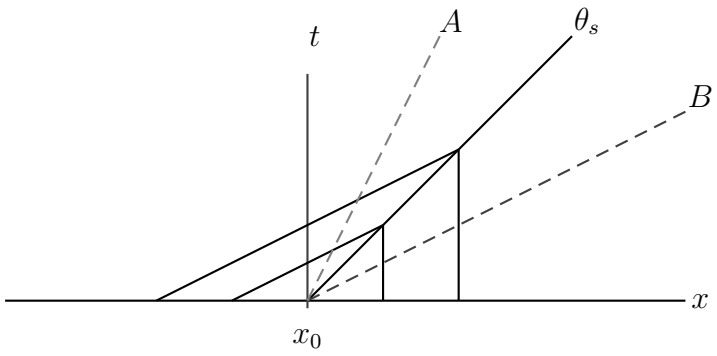


Figure 9.3: Characteristics

Now, we study uncertainty propagation through Burgers' equation in this context by calculating the expectation of $\rho(x,t)$. The domain in x is divided into three different regions or intervals. Figure (9.4) shows how we divide the regions. In order to calculate the expectation of $\rho(\cdot)$ we must consider the behavior in the following intervals: $(-\infty, x_A]$, $[x_A, x_B]$ and $[x_B, \infty)$ where x_A corresponds to $\theta_s = A$ and x_B corresponds to $\theta_s = B$.

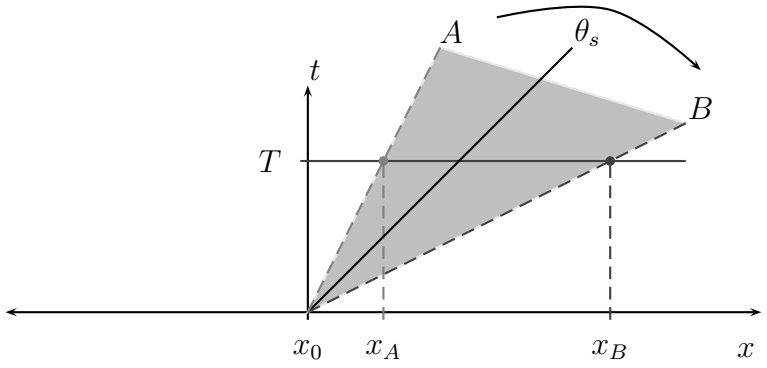


Figure 9.4: Intervals Created by θ.

Recall that expectation of a function $g(\cdot)$ is given by the integral

$$E[g(x,\theta)] = \int_{-\infty}^{\infty} f_\theta(\theta) g(x,\theta) d\theta. \tag{9.9}$$

Thus, for the interval $(-\infty, x_A]$ the expectation of $\rho(\cdot)$ is given by

$$E[\rho(x,t,\theta)]_{x\in(-\infty,x_A]} = \int_A^B \frac{a+b\theta}{B-A} d\theta = a + \frac{b}{2}\frac{B^2-A^2}{B-A} = c_1, \tag{9.10}$$

where c_1 is a constant. Similarly, the expectation of $\rho(\cdot)$ in the interval $[x_B, \infty)$ is given by

$$E[\rho(x,t,\theta)]_{x\in[x_B,\infty)} = \int_A^B \frac{-a+b\theta}{B-A} d\theta = -a + \frac{b}{2}\frac{B^2-A^2}{B-A} = c_2, \tag{9.11}$$

where c_2 is a constant. Notice that since $a > 0$ and $b > 0$, we have $c_2 < c_1$. Now, if we fix x and t and vary θ_s in the cone created by rotating θ_s between A and B, then we see that if $\theta_s \in [A, \frac{x}{tb}]$, then $\rho = \rho_r$ and if $\theta_s \in [\frac{x}{tb}, B]$, then $\rho = \rho_\ell$. Thus, for the

interval $[x_A, x_B]$, the expectation is given by

$$\begin{aligned} E[\rho(x,t,\theta)]_{x\in[x_A,x_B]} &= \int_A^{x/(tb)} \frac{-a+b\theta}{B-A}d\theta + \int_{x/(tb)}^{B} \frac{a+b\theta}{B-A}d\theta \\ &= \frac{1}{B-A}\left(\left[-a\theta+\frac{b}{2}\theta^2\right]_A^{x/(tb)} + \left[a\theta+\frac{b}{2}\theta^2\right]_{x/(tb)}^{B}\right) \\ &= \frac{1}{B-A}\left[-2a\frac{x}{tb}+a(A+B)+\frac{b}{2}(B^2-A^2)\right]. \end{aligned} \tag{9.12}$$

Notice, that the expectation $E[\rho(x,t,\theta)]$ where $x \in [x_A, x_B]$ is a linear relation dependent on x and t. To verify that the expectations are consistent with one another at the end points, consider $x_B = tbB$. Plug $x = x_B$ into equation (9.13). We get

$$\begin{aligned} E[\rho(x_B,t,\theta)]_{x\in[x_A,x_B]} &= \frac{1}{B-A}\left[2a\frac{x_B}{tb}-a(A+B)+\frac{b}{2}(B^2-A^2)\right] \\ &= \frac{1}{B-A}\left[-2a\frac{tbB}{tb}+a(A+B)+\frac{b}{2}(B^2-A^2)\right] \\ &= \frac{1}{B-A}\left[-2aB+a(A+B)+\frac{b}{2}(B^2-A^2)\right] \\ &= \frac{1}{B-A}\left[-a(B-A)+\frac{b}{2}(B^2-A^2)\right] \\ &= \left[-a+\frac{b}{2}\frac{(B^2-A^2)}{B-A}\right] \\ &= c_2 \end{aligned} \tag{9.13}$$

Now, consider $x_A = tbA$. Plug $x = x_A$ into equation (9.13). We get

$$\begin{aligned}
E[\rho(x_A, t, \theta)]_{x\in[x_A,x_B]} &= \frac{1}{B-A}\left[2a\frac{x_A}{tb} - a(A+B) - \frac{b}{2}(B^2 - A^2)\right] \\
&= \frac{1}{B-A}\left[-2a\frac{tbA}{tb} + a(A+B) + \frac{b}{2}(B^2 - A^2)\right] \\
&= \frac{1}{B-A}\left[-2aA + a(A+B) + \frac{b}{2}(B^2 - A^2)\right] \qquad (9.14) \\
&= \frac{1}{B-A}\left[a(B-A) + \frac{b}{2}(B^2 - A^2)\right] \\
&= \left[a + \frac{b}{2}\frac{(B^2 - A^2)}{B-A}\right] \\
&= c_1
\end{aligned}$$

Figure (9.5) shows the expectation of $\rho(x, t, \theta)$ for fixed t in the three regions. As expected the expectation is a continuous function, constant in the intervals $(-\infty, x_A]$ and $[x_B, \infty)$ and linear in the region $[x_A, x_B]$.

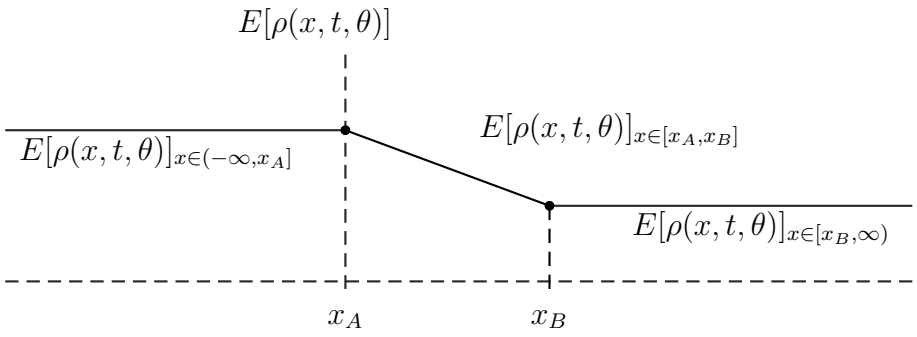


Figure 9.5: Expectation of $\rho(x, t, \theta)$.

Now, we consider the second moment. Recall the second moment of a function

$g(\cdot)$ is given by the integral

$$E[g^2(x,\theta)] = \int_{-\infty}^{\infty} f_\theta(\theta) g^2(x,\theta) d\theta. \tag{9.15}$$

In order to calculate the second moment of $\rho(\cdot)$ we must consider the behavior in the following intervals: $(-\infty, x_A]$, $[x_A, x_B]$ and $[x_B, \infty)$ where x_A corresponds to $\theta_s = A$ and x_B corresponds to $\theta_s = B$. For the interval $(-\infty, x_A]$, the second moment of $\rho(\cdot)$ is given by

$$\begin{aligned} E[\rho^2(x,t,\theta)]_{x\in(-\infty,x_A]} &= \int_A^B \frac{(a+b\theta)^2}{B-A} d\theta \\ &= a^2 + \frac{1}{B-A}\left[ab(B^2-A^2) + \frac{b^2}{3}(B^3-A^3)\right] = c_3, \end{aligned} \tag{9.16}$$

where c_3 is constant. Similarly, the second moment of $\rho(\cdot)$ in the interval $[x_B, \infty)$ is given by

$$\begin{aligned} E[\rho^2(x,t,\theta)]_{x\in(x_B,\infty]} &= \int_A^B \frac{(-a+b\theta)^2}{B-A} d\theta \\ &= a^2 - \frac{1}{B-A}\left[ab(B^2-A^2) - \frac{b^2}{3}(B^3-A^3)\right] = c_4, \end{aligned} \tag{9.17}$$

where c_4 is a constant. Notice that $a > 0$ and $b > 0$, we have $c_4 < c_3$. Similar to the calculation of expectation, if we fix x and t and vary θ_s in the cone created by rotating θ_s between A and B, then we see that if $\theta_s \in [A, \frac{x}{tb}]$, then $\rho = \rho_r$ and if $\theta_s \in [\frac{x}{tb}, B]$, then $\rho = \rho_l$. Thus, for the interval $[x_A, x_B]$, the second moment is given

by

$$
\begin{aligned}
E[\rho^2(x,t,\theta)]_{x\in[x_A,x_B]} &= \int_A^{x/(tb)} \frac{(-a+b\theta)^2}{B-A}d\theta + \int_{x/(tb)}^{B} \frac{(a+b\theta)^2}{B-A}d\theta \qquad (9.18)\\
&= \frac{1}{B-A}\left(\left[\frac{b^2}{3}\theta^3 - ab\theta^2 + a^2\theta\right]_A^{x/(tb)} + \left[a^2\theta + ab\theta^2 + \frac{b^2}{3}\theta^3\right]_{x/(tb)}^{B}\right)\\
&= a^2 + \frac{1}{B-A}\left[-2ab\left(\frac{x}{tb}\right)^2 + \frac{b^2}{3}(B^3-A^3) + ab(A^2+B^2)\right].
\end{aligned}
$$

To verify that the second moments in the intervals are consistent with one another at the end points, consider $x_B = tbB$. Plug $x = x_B$ into equation (9.19).

$$
\begin{aligned}
E[\rho^2(x_B,t,\theta)]_{x\in[x_A,x_B]} &= a^2 + \frac{1}{B-A}\left[-2ab\left(\frac{x_B}{tb}\right)^2 + \frac{b^2}{3}(B^3-A^3) + ab(A^2+B^2)\right]\\
&= a^2 + \frac{1}{B-A}\left[-2abB^2 + \frac{b^2}{3}(B^3-A^3) + ab(A^2+B^2)\right] \qquad (9.19)\\
&= a^2 - \frac{1}{B-A}\left[ab(B^2-A^2) - \frac{b^2}{3}(B^3-A^3)\right] = c_4.
\end{aligned}
$$

Now, consider $x_A = tbA$. Plug $x = x_A$ into equation (9.19).

$$
\begin{aligned}
E[\rho^2(x_A,t,\theta)]_{x\in[x_A,x_B]} &= a^2 + \frac{1}{B-A}\left[-2ab\left(\frac{x_A}{tb}\right)^2 + \frac{b^2}{3}(B^3-A^3) + ab(A^2+B^2)\right]\\
&= a^2 + \frac{1}{B-A}\left[-2abA^2 + \frac{b^2}{3}(B^3-A^3) + ab(A^2+B^2)\right] \qquad (9.20)\\
&= a^2 + \frac{1}{B-A}\left[ab(B^2-A^2) + \frac{b^2}{3}(B^3-A^3)\right] = c_3.
\end{aligned}
$$

Figure (9.6) shows the second moment of $\rho(x,t,\theta)$ for fixed t in the three re-

gions. As expected the expectation is a continuous function, constant in the intervals $(-\infty, x_A]$ and $[x_B, \infty)$ and linear in the region $[x_A, x_B]$.

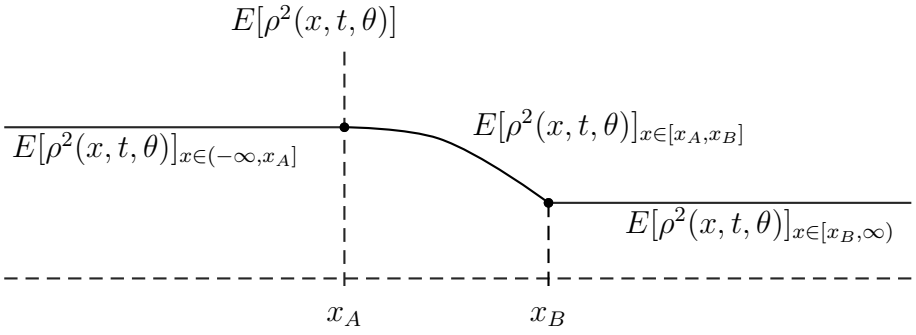


Figure 9.6: Second Moment of $\rho(x,t,\theta)$.

9.3 Summary

In this chapter we have provided the theory for how to propagate uncertainty in the initial condition through Burgers' equation. In future work, a conservation form will be developed for this method of propagating uncertainty in the initial condition. The work presented here is original, and thus, is a contribution.

CHAPTER 10

NUMERICAL SOLUTION TO THE LIOUVILLE EQUATION

In this chapter we apply the Liouville equation to examples of ODEs including various linear ODEs. The resulting Liouville equations that we solve are constant advection and variable advection equations. We used MATLAB to numerically solve the Liouville equation. The numerical method we use is the Godunov method, which is a common method for numerically solving hyperbolic partial differential equations like the Liouville equation. The Godunov method utilizes a Riemann solver. The results are presented in this chapter and the code is presented in Appendix A.

10.1 Simulations for Constant Advection Equation

The first example of an ODE that we explore is of the form

$$\dot{x} = -5, \tag{10.1}$$

with initial condition given as a normally distributed random variable. The pdf for the random variable corresponding to the initial condition was taken to be

$$u_{X_0}(x) = \frac{1}{\sqrt{5\pi}} e^{\frac{-x}{5}}. \tag{10.2}$$

The LE then becomes

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}[-5u] = 0. \tag{10.3}$$

Figure (10.1) confirms that as the system evolves the geometry of the pdf remains the same but it advects or drifts. The diffusion that is present in figure (10.1) is a

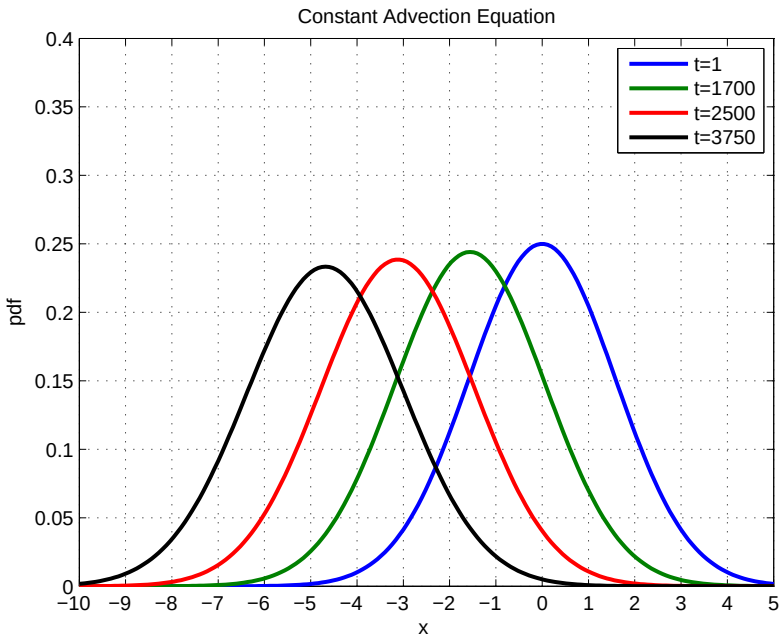


Figure 10.1: Constant Advection Equation Results

result of the numerical method and not a result of the application of the LE to the problem of uncertainty propagation.

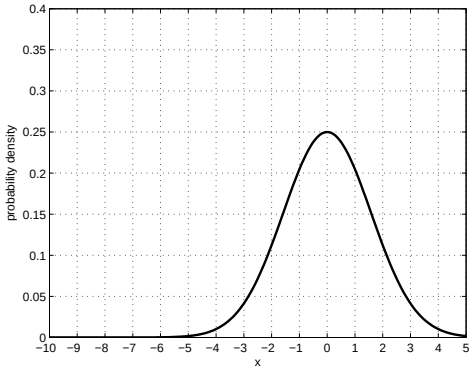
10.2 Simulations for Variable Advection Equation

The second example of an ODE that we explore is of the form

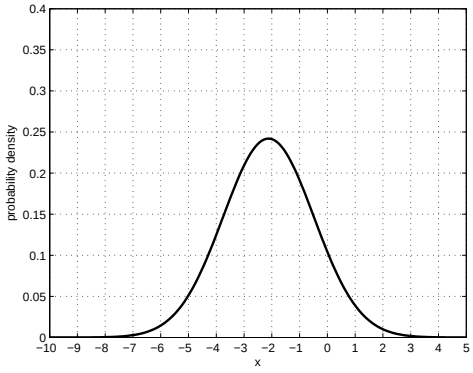
$$\dot{x} = -5x, \tag{10.4}$$

with initial condition given as a normally distributed random variable. The pdf for the random variable corresponding to the initial condition was taken to be

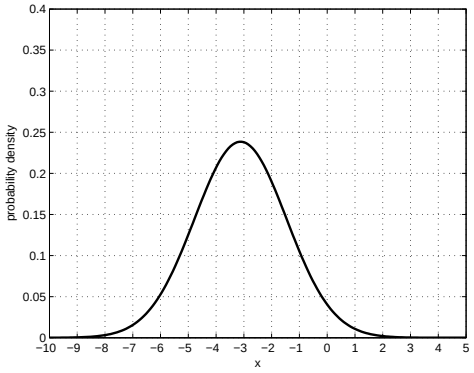
$$u_{X_0}(x) = \frac{1}{\sqrt{5\pi}} e^{\frac{-x}{5}}. \tag{10.5}$$



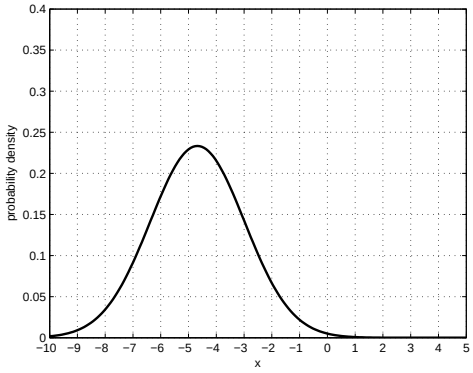
(a) Time 1: Initial



(b) Time 2



(c) Time 3



(d) Time 4

Figure 10.2: Uncertainty Propagation in ODE $y' = -5$

The LE then becomes

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}[-5xu] = 0. \tag{10.6}$$

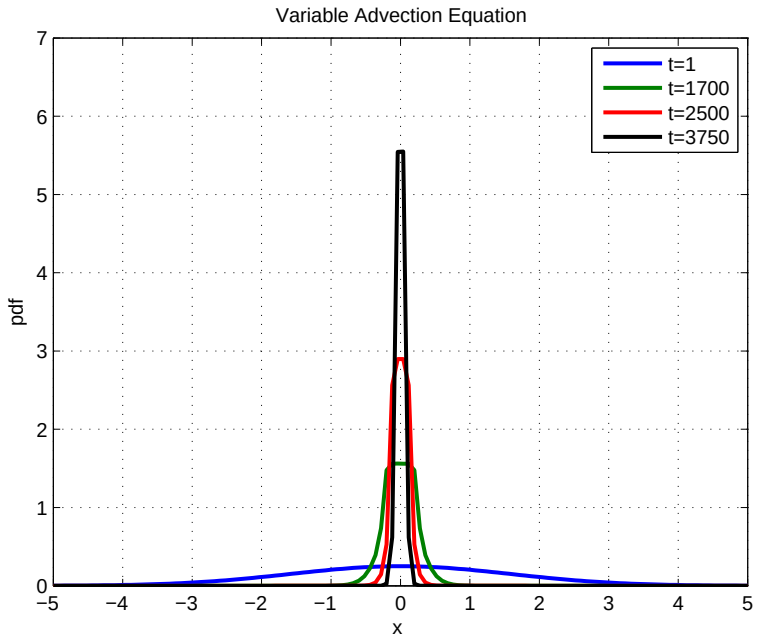
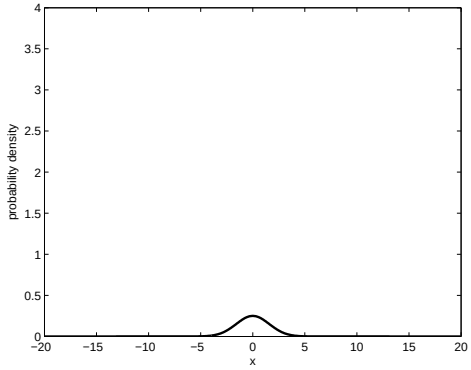


Figure 10.3: Variable Advection Equation Results

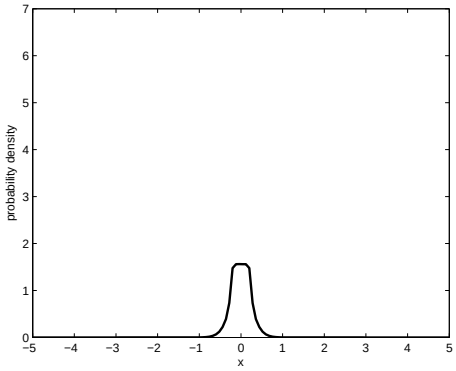
We expect that with a stable ODE such as the one chosen for this example, we should observe the geometry of the pdf approach that of a Dirac delta distribution centered around the mean. Figures(10.4) and (10.2) confirm that as the system evolves the geometry of the pdf approaches that of the Dirac delta distribution.

10.3 Summary

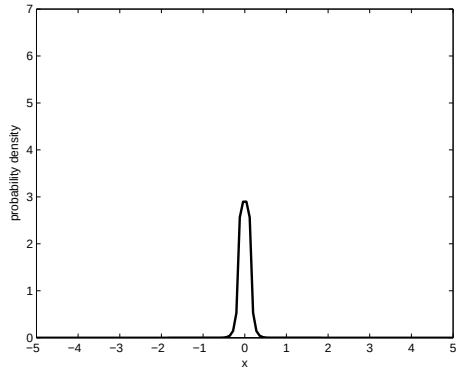
In this chapter we present several examples of propagation of uncertainty through ODEs. MATLAB code was written to solve the LE which is the method for propagation of uncertainty through ODEs. Plots of the pdf as it evolves have been included.



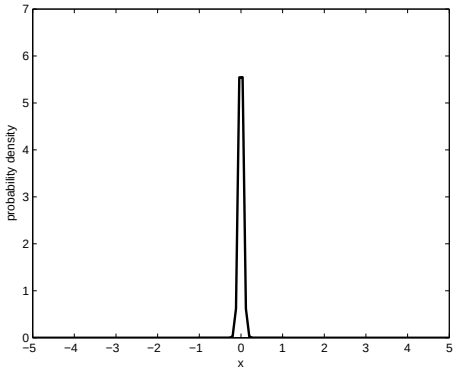
(a) Time 1: Initial



(b) Time 2



(c) Time 3



(d) Time 4

Figure 10.4: Uncertainty Propagation in ODE $y' = -5x$

CHAPTER 11

NUMERICAL SOLUTION FOR UNCERTAINTY PROPAGATION IN FDES

In this chapter we apply the theory of functions of random variables to basic finite difference equations. Taking simple finite difference equations, we use MATLAB to propagate the uncertainty from the initial condition.

11.1 Simulations

Consider a FDE of the form

$$x[k+1] = ax[k] + b \tag{11.1}$$

with an initial condition $x[0] = X_0$ given as a normally distributed random variable. Using functions of random variables, code was written in MATLAB to simulate the propagation of uncertainty in a system of this type.

11.1.1 FDE Example One

The example presented in this section is uncertainty propagation through the FDE given by

$$x[k+1] = 0.5x[k] \tag{11.2}$$

with the pdf of the initial condition given by

$$u_{X_0}(x) = \frac{2}{25} e^{\frac{-x^2}{50}}. \tag{11.3}$$

The code used to propagate the initial uncertainty is given in Appendix A.

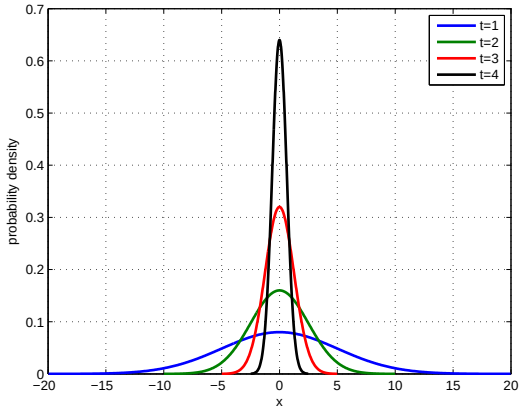


Figure 11.1: Uncertainty Propagation in FDE $x[k+1] = 0.5x[k]$

We can see from figure (11.1) that the pdf is approaching the Dirac delta distribution since the constant $a = 0.5$ is less than one. The mean of the pdf remains at zero due to the fact that there is no 'drift' term associated with the FDE.

11.1.2 FDE Example Two

The example presented in this section uncertainty propagation through the FDE given by

$$x[k+1] = 2x[k] + 6 \tag{11.4}$$

with the pdf of the initial condition given by

$$u_{X_0}(x) = \frac{2}{25} e^{\frac{-x^2}{50}}. \tag{11.5}$$

The code used to propagate the initial uncertainty is given in Appendix A.

We can see from figure (11.2) that the pdf is approaching is diffusing since the constant $a = 2$ is greater than one. The mean of the pdf shifting due to the fact that

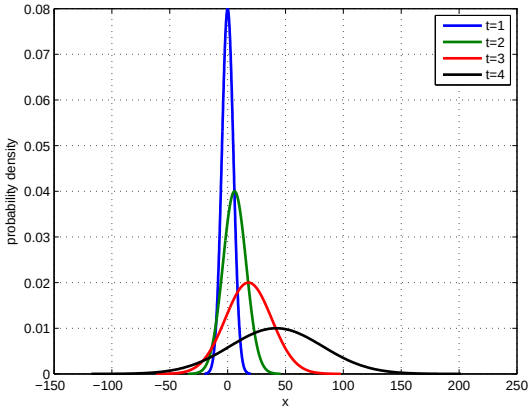


Figure 11.2: Uncertainty Propagation in FDE $x[k+1] = 2x[k] + 6$

there is a 'drift' term associated with the FDE, i.e. $b = 6$.

The examples presented in this chapter provide an basic understanding of how we propagate uncertainty through FDEs using functions of random variables.

11.2 Summary

In this chapter we present several examples of propagation of uncertainty through FDEs. MATLAB code was written to simulate the propagation of uncertainty from the initial condition in FDEs. Plots of the pdf as it evolves have been included.

CHAPTER 12
CONCLUSIONS

A fundamental principle of science is that no real measurement or model is infinitely precise. It has an inherent degree of uncertainty that we must accept and learn to process and incorporate into our models and our designs. The utility of a model is determined by its predictive ability. To consider uncertainty in the modeling process allows for increased model utility. Uncertainty can be in the model parameters, boundary conditions, initial conditions and the dynamics themselves. The work presented in this thesis attempts to address the issue of uncertainty modeling in dynamic systems. In particular, we present a conservation based method of propagating uncertainty in the initial condition through the system given that the initial condition is defined within some probability structure. The conservation based method for uncertainty propagation follows directly from the global conservation law of probability,

$$\int_{\Omega} u_X(x,t)dP = 1, \tag{12.1}$$

which is a fundamental principle in probability theory. It allows for the probability function to be a measurable function and allows for the theory developed in this thesis to be consistent with the probability structure on the problem of uncertainty propagation from the initial condition. The mathematical integrity of the problem of uncertainty propagation heavily depends on this.

The main idea presented in this thesis is that given an initial state space defined on a probability structure, system dynamics, and an evolution operator that maps the initial state space to a state space at time $t > 0$, we define a probability structure

on the state at time $t > 0$ such that the evolution operator is measure preserving. This idea is the fundamental concept which is central to the theory of conservation based uncertainty propagation in the dynamic systems presented in this thesis.

12.1 Ordinary Differential Equations

The method for propagating uncertainty in the initial condition through ODEs and systems of ODEs is the Liouville equation. The work in this thesis presents the derivation of the Liouville Equation and its solution. The Liouville equation is derived based on conservation of probability mass and is consistent with preserving the probability measure. The Liouville equation as derived for propagation of uncertainty in ODEs is a semilinear hyperbolic partial differential equation, and thus, there are only a few cases in which we can solve the Liouville equation analytically. Numerical method must be employed to propagate uncertainty in ODEs in a practice.

12.2 Finite Difference Equations

FDEs are the discrete time analog to ODEs. The method for propagating uncertainty in the initial condition in FDEs is through the theory of functions of random variables. The work in this thesis presents the development of the theory of functions of random variables are it pertains to propagation of uncertainty in FDEs. Using functions of random variables we are able to preserve the probability measure defined in the probability space given as part of the initial condition through defining the state of the system as a random variable and considering the dynamics to be simply a function of random variable(s). Thus, this method is consistent with the conservation

principle.

12.3 Differential Inequalities and Inclusions

Differential inequalities and inclusions are the most abstract mathematical model under consideration. In this thesis we present a method for uncertainty propagation only in a small class of differential inclusions. The method for propagation is through defining a probability structure on the initial set of possible values the dynamics can take. Over each interval in the initial set, we define a scaled uniform density. The sum the sum of these functions over the entire initial set is a density function with the appropriate properties. For each interval, we propagate this initial distribution according to the dynamics by taking the supremum of the upper bound function over the interval and the infimum of the lower bound function over all possible values in the interval thereby creating a new interval from these values. The distribution over the initial interval is uniformly spread out over the new interval. This method of uncertainty propagation is also conservation based and is consistent with our defined probability structure.

12.4 Stochastic Differential Equations

The method for propagating uncertainty in the initial condition in SDEs is through the Fokker-Planck equation. The work in this thesis presents the derivation of the FPE and how it is used to solve for the pdf of the states of a system of SDEs at future time. We also show that the FPE can be written in conservation form. The FPE is a partial differential equation, and thus, numerical methods must be use to solve for

the pdf at future times. It is very interesting, mathematically speaking, to note that uncertainty propagation from the initial condition in a SDE, which is a model that accounts for randomness, can be converted to a partial differential equation which is deterministic. This means an inherently stochastic problem becomes deterministic when the appropriate probability structure is defined.

12.5 Markov Chains

Markov chains are an interesting class of dynamics systems which are memoryless. Future states only depend on the current state of a system. Thus, for uncertainty propagation, they are interesting to study. In this thesis we review what Markov property and what constitutes a Markov chain. We also show that the problem of uncertainty propagation in Markov chains can be solved using probabilistic inference. This means that by applying the transition probabilities iteratively, posterior distributions for the states of the system can be determined.

12.6 Uncertainty Propagation in Burgers' Equation

Burgers' equation is a nonlinear hyperbolic partial differential equation. In this thesis, we present a method for propagating uncertainty in the initial condition through Burgers' equation. We assume that the initial data is piecewise constant function with a random parameter that renders the initial data to be stochastic. We also assume the the random parameter is uniformly distributed. The method presented uses only the dynamics of the system and the initial data. Propagation of uncertainty is done using the characteristics of the system and evaluating the expec-

tation of the velocity variable.

12.7 Future Work

The work presented in this thesis is only a scratch in the surface. The range of the term dynamic systems is very broad and this study covers only a small subset of what we call dynamic systems. In the future, we hope to expand this study to include a study of a more general and abstract set of dynamic systems and a study of qualitative properties such as stability analysis and the conditions under which the problems become well posed. In addition, we hope to apply some of the results presented to various practical systems such as transportation networks and financial markets.

APPENDIX A

MATLAB CODE

A.1 MATLAB Code for Liouville Equation

This section of the appendix includes the MATLAB code used to propagate uncertainty through ODEs.

A.1.1 Constant Advection Equation

```
a = -20; b = 20; T = 1; M = 500; rhom = 0.2;

n = 100; drho = rhom/n; rho = 0:drho:rhom;

h = (b-a) / M;

xticks = transpose(a+h/2:h:b-h/2);frho = fluxadvec(xticks);

[qmax,s] = max(frho);

rhostar = rho(s);


% CFL Condition

lambda = max(abs(frho(2:n)-frho(1:n-1)))/drho;


ka = 0.5 * h / lambda;

N = ceil(T/ka) ;

k = T / N;


% Initialise:

U = zeros(M,N+1);

U(:,1) = initialadvec(xticks);
```

```
figure(1);

plot(xticks,U(:,1),'k','LineWidth',2)

axis([a b 0 4])

ylabel('probability density')

xlabel('x')


% Algorithm

for j=1:N

rhol = U(1:M-1,j);

rhor = U(2:M,j);

qval=-5*U(:,j);

ql = qval(1:M-1);

qr = qval(2:M);

case1 = rhol <= rhor;

case2 = rhol > rhostar & rhostar > rhor;

case3 = not(case1 | case2);

Q = case1 .* min(ql,qr) + case2 * qmax + case3 .* max(ql,qr);


% Specifying Boundary Flows

d=transpose(fluxadvec(initialadvec(a)));

Q = [-5*initialadvec(a); Q; qval(M)];

U(:,j+1)=U(:,j) + (k/h) * (Q(1:M) - Q(2:M+1));

end
```

```
% initialadvec.m

function init=initialadvec(x)

init=1/4 * exp(-x.^2/5);


% fluxadvec.m

function fl=fluxadvec(rho)

fl = -5*rho;


% fluxadvec2.m

function fl=fluxadvec2(u)

for i = 1:500

fl(i)=-5*u(i);

end
```

A.1.2 Variable Advection Equation

```
%System Variables

%Domain: [a,b]

a = -20;

b = 20;

%Stopping Time = 1

T = 1;

M = 500;

%x-vector and dx

h = (b-a) / M;

xticks = transpose(a+h/2:h:b-h/2);
```

```
%Initial Condition

u0=initialadvec(xticks);

%Determine rho*

[umax,m]=max(u0);

du=umax/n;

frho = fluxadvec(xticks);

[qmax,s] = max(frho);

rhostar = xticks(s);

% CFL Condition

lambda = max(abs(frho(2:n)-frho(1:n-1)))/du;

ka = 0.5 * h / lambda;

N = ceil(T/ka) ;

k = T / N;

% Initialize:

U = zeros(M,N+1);

U(:,1) = initialadvec(xticks);

% Algorithm

for j=1:N

rhol = U(1:M-1,j);

rhor = U(2:M,j);

qval = transpose(fluxadvec2(xticks,U(:,j)));

ql = qval(1:M-1);

qr = qval(2:M);
```

```
case1 = rhol <= rhor;

case2 = rhol > rhostar & rhostar > rhor;

case3 = not(case1 | case2);

Q = case1 .* min(ql,qr) + case2 * qmax + case3 .* max(ql,qr);

% Specifying Boundary Flows

d=transpose(fluxadvec(initialadvec(a)));

Q = [d*U(1,j); Q; qval(M)];

%Determine U at t=j+1

U(:,j+1)=U(:,j) + (k/h) * (Q(1:M) - Q(2:M+1));

end


% fluxvaradvec.m

function fl=fluxvaradvec(x,u)

fl1 = -5*x;

for i = 1:500

fl(i)=fl1(i)*u(i);

end
```

A.2 MATLAB Code for Finite Difference Equations

The code in this section is for the simulation of uncertainty propagation in FDEs.

A.2.1 FDE Example 1

```
%Initialize Variables

a=0.5; b= 0; N=500;M=100;
```

```
y1=-20; y2=20; dy = (y2-y1)/N;

y= transpose(y1+dy/2:dy:y2-dy/2);

fx= zeros(N,M+1);

fx(:,1) = 0.08 * exp(-y.^2/50);

yold1=y1;

yold2=y2;

yv(:,1)=y;

%Algorithm

for j = 1: 4

   yend1=yold1*a+b;

   yend2=yold2*a+b;

   delta=(yend2-yend1)/N;

   yvec=transpose(yend1+delta/2:delta:yend2-delta/2);

   yv(:,j+1)=yvec;

   sum = 0;

   for m = 1: j

       sum = sum +b/a^m;

    end

    for i = 1:N

        ynew(i) = yvec(i)/a^(j)-sum;

    end

    f = 0.08 * exp(-ynew.^2/50);

    fx(:,j+1)=(1/(abs(a))^(j))*f;
```

```
    yold1=yend1;
    yold2=yend2;
end
%Plot
figure(1);
plot(yv(:,1),fx(:,1),yv(:,2),fx(:,2),yv(:,3)...
    ,fx(:,3),yv(:,4),fx(:,4),'k','LineWidth',2)
grid on
legend('t=1','t=2','t=3', 't=4')
ylabel('probability density')
xlabel('x')
```

A.2.2 FDE Example 2

```
%Initialize Variables
a=2; b= 6; N=500;M=100;
y1=-20; y2=20; dy = (y2-y1)/N;
y= transpose(y1+dy/2:dy:y2-dy/2);
fx= zeros(N,M+1);
fx(:,1) = 0.08 * exp(-y.^2/50);
yold1=y1;
yold2=y2;
yv(:,1)=y;
%Algorithm
```

```
for j = 1: 4
   yend1=yold1*a+b;
   yend2=yold2*a+b;
   delta=(yend2-yend1)/N;
   yvec=transpose(yend1+delta/2:delta:yend2-delta/2);
   yv(:,j+1)=yvec;
   sum = 0;
   for m = 1: j
       sum = sum +b/a^m;
    end
    for i = 1:N
        ynew(i) = yvec(i)/a^(j)-sum;
    end
    f = 0.08 * exp(-ynew.^2/50);
    fx(:,j+1)=(1/(abs(a))^(j))*f;
    yold1=yend1;
    yold2=yend2;
end
%Plot
figure(1);
plot(yv(:,1),fx(:,1),yv(:,2),fx(:,2),yv(:,3)...
    ,fx(:,3),yv(:,4),fx(:,4),'k','LineWidth',2)
grid on
```

```
legend('t=1','t=2','t=3', 't=4')

ylabel('probability density')

xlabel('x')
```

BIBLIOGRAPHY

[1] George L. Shackle. *Decision, order, and time in human affairs.* Cambridge University Press, 1961.

[2] Frederic Y. Bois and Cheikh T. Diack. *Recent Advances in Quantitative Methods in Cancer and Human Health Risk Assessment*, chapter 15, pages 255–266. Wiley, 2005.

[3] Athanasios Papoulis. *Probability, Random Variables and Stochastic Processes.* McGraw-Hill Book Company, 1965.

[4] Lawrence C Evans. An Introduction to Stochastic Differential Equations Version 1.2, 2008.

[5] Hisa Tanaka. Quick derivation of fokker-planck equation. htanaka2@wisc.edu, Nov. 2005.

[6] Alex Bielajew. Monte carlo modeling in external electronbeam radiotherapy-why leave it to chance? National Research Council, Ottawa Canada.

[7] Norbert Wiener. The homogeneous chaos. *American Journal of Mathematics*, 60(4):897–936, 1938.

[8] Adrian Sandu, Corina Sandu, and Mehdi Ahmadian. Modeling multibody systems with uncertainties. part 1: Theoretical and computational aspects. *Multibody System Dynamics*, 15(4):369–391, May 2006.

[9] Teems E. Lovett, Ferdinanda Ponci, and Antonello Monti. A polynomial chaos approach to measurement uncertainty. *IEEE Transactions on Intrumentation and Measurement*, 55(3):723–736, June 2006.

[10] Teems E. Lovett, Antonello Monti, and Ferdinanda Ponci. Automatic synthesis of uncertain models from linear circuit simulation: A polynomial chaos theory approach. *Simulation Modelling Practice and Theory*, 16:796–816, 2008.

[11] Per Pettersson, Gianluca Iaccarino, and Jan Nordstrom. Numerical analysis of the burger's equation in the presence of uncertainty. *Journal of Computational Physics*, 2009.

[12] R. H. Cameron and W. T. Martin. The orthogonal development of non-linear functionals in series of fourier-hermite functionals. *Annals of Mathematics*, 48(2):385–392, April 1946.

[13] Dongbin Xiu and George Em Karniadakis. The wiener-askey polynomial chaos for stochastic differential equations. *SIAM Journal on Scientific Computing*, 24(2):619–644, 2002.

[14] Pradya Prempraneerach. *Uncertainty Analysis in Shipboard Integrated Power System using Multi-Element Polynomial Chaos*. PhD thesis, Massachusetts Institute of Technology, 2007.

[15] Willian A. Link and Richard J. Barker. *Bayesian Inference with Ecological Applications*. Elsevier, 2010.

[16] S.F. Su and C.S.G. Lee. Uncertainty manipulation and propagation and verification of applicability of actions in assembly tasks. In *Proceedings of the 1991 IEEE International Conference on Robotics and Automation*, Sacramento, California, April 1991.

[17] L.A.F. Fernandes, M. Oliveira, and R. da Silva. Uncertainty propagation: Avoiding the expensive sampling process for real-time image-based measurements. *Journal of Computational Statistics and Data Analysis*, 52:3852–3876, 2007.

[18] B. Ruggeri. Chemicals exposure: Scoring procedure and uncertainty propagation in scenario selection for risk analysis. *Chemosphere*, 77:330–338, 2009.

[19] A. Abusam, K.J. Keesman, and G. van Straten. Forward and backward uncertainty propagation: an oxidation ditch modeling example. *Water Research*, 37:429–435, 2003.

[20] P. Baraldi and E. Zio. A combined monte carlo and possibilistic approach to uncertianty propagation in event tree analysis. *Risk Analysis*, 28:1309–1325, 2008.

[21] S. Blasone, A. Vrugt, H. Madsen, D. Rosbjerg, B. Robinson, and G. Zyvoloski. Generalized likelihood uncertainty estimation (glue) using adaptive markov chain monte carlo sampling. *Advances in Water Resources*, 31:630–648, 2008.

[22] L. Cristaldi, M. Faifer, and F. Ponci. Uncertainty propagation in diagnostic systems based on soft computing techniques. In *Proceedings of International*

Workshop on Advanced Methods for Uncertianty Estimation in Measurement, Sardagna, Trento, Italy, July 2007.

[23] Martin Ehrendorfer. The liouville equation and atomospheric predictability. In *Predictability of Weather and Climate*, pages 59–98. Cambridge University Press, 2006.

[24] S.H. Lee and W. Chen. A comparative study of uncertainty propgation methods for black-box type problems. *Structural and Multidisciplinary Optimization*, 37(3):239–253, January 2009.

[25] T. Lee, M. Leok, and N.H. McClamroch. Global symplectic uncertainty propagation on so(3). In *Proceedings of the 47th IEEE Conference on Decision and Control*, Cancun, Mexico, December 2008.

[26] L. Mari. A computational system for uncertainty propagation of measurement results. *Measurement*, 42:844–855, 2009.

[27] A. Mencattini, S. Marcello, and R. Lojacono. On generalized t-norm for the representation of uncertianty propagation in statistically correlated measurements by means of fuzzy variables. In *Proceedings of International Workshop on Advanced Methods for Uncertianty Estimation in Measurement*, Sardagna, Trento, Italy, December 2007.

[28] A. Sandu, C. Sandu, and M. Ahmadian. Modeling multibody systems with uncertainties. part 2: Numerical applications. *Multibody System Dynamics*, 15(4):241–262, May 2006.

[29] G. Terejanu, P. Singla, T. Singh, and P. Scott. Uncertianty propagation for nonlinear dynamic systems using gaussian mixture models. *Journal of Guidance, Control and Dynamics*, 31:1623–1633, 2008.

[30] Anthony N. Michel, Kaining Wang, and Bo Hu. *Qualitative Theory of Dynamical Systems: The Role of Stability Preserving Mappings*. Marcel Dekker, Inc., 2nd edition, 2001.

[31] Bernt Oksendal. *Stochastic Differential Equations*. Springer-Verlag, 3 edition, 1992.

[32] Zdzislaw Brezezniak and Tomasz Zastawniak. *Basic Stochastic Processes*. Springer, 1999.

[33] Paul G. Hoel, Sidney C. Port, and Charles J. Stone. *Introduction to Stochastic Processes*. Houghton Mifflin Company, 1972.

[34] Richard Bronson and Gabriel B. Costa. *Schaum's Outline of Differential Equations*. McGraw-Hill Book Company, 2006.

[35] Martin Ehrendorfer. The Liouville Equation and Its Potential Usefulness for the Prediction of Forecast Skill. Part I: Theory. *Monthly Weather Review*, 122:703–713, May 1993.

[36] Martin Ehrendorfer. The Liouville Equation and Its Potential Usefulness for the Prediction of Forecast Skill. Part II: Applications. *Monthly Weather Review*, 122:714–728, May 1993.

[37] Patricia Mellodge and Pushkin Kachroo. *Model Abstraction in Dynamical Systems: Application to Mobile Robot Control*. Lecture Notes In Control and Information Sciences. Springer, 2008.

[38] David Richard Wolf. Information and correlation in statistical mechanical systems. University of Leeds, 1996.

[39] Jean-Pierre Aubin. *Set-Valued Analysis*. Birkhauser, 1990.

[40] Y.A. Rozanov. *Probability Theory: A Concise Course*. Dover Publications, 1969.

[41] Randall J. LeVeque. *Numerical Methods for Conservation Laws*. Birkhauser, 1992.

[42] Randall J. LeVeque. *Finite Volume Methods for Hyperbolic Problems*. Cambridge University Press, 1995.

VITA

Graduate College
University of Nevada, Las Vegas

Lillian J. Ratliff

Degrees:

Bachelors of Science, Electrical Engineering, 2008
University of Nevada, Las Vegas

Bachelors of Science, Mathematics, 2008
University of Nevada, Las Vegas

Special Honors and Awards:

National Science Foundation Graduate Research Fellowship 2009

Publications:

Cook, D. P., Y. Chen, L. J. Ratliff, and J. Ma. 2006. 'Numerical Modeling of EM Pump Efficiency'. Proceedings of the 2006 American Society for Mechanical Engineering Annual Conference & Exposition.

Ratliff, L. and P. Kachroo. 2010. 'Validating Numerically Consistent Macroscopic Traffic Models Using Microscopic Data'. Transportation Research Board Annual Meeting.

Thesis Title:

Conservation Based Uncertainty Propagation in Dynamic Systems

Thesis Examination Committee:

Chairperson, Pushkin Kachroo, Ph.D.
Committee Member, Mehdi Ahmadian, Ph.D.
Committee Member, Rama Venkat, Ph.D.
Committee Member, Yahia Baghzouz, Ph.D.
Graduate Faculty Representative, Mohammad Trabia, Ph.D.