A Strong Maximum Principle for Reaction-Diffusion Systems and a Weak Convergence Scheme for Reflected Stochastic Differential Equations

by

Lawrence Christopher Evans

Bachelor of Arts, University of California, June 2001

Submitted to the Department of Mathematics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2010

Author ....
Department of Mathematics
April 23, 2010

Certified by ....
Daniel W. Stroock
Professor of Mathematics
Thesis Supervisor

Accepted by ....
Bjorn Poonen
Chairman, Department Committee on Graduate Students

A Strong Maximum Principle for Reaction-Diffusion Systems and a Weak Convergence Scheme for Reflected Stochastic Differential Equations

by

Lawrence Christopher Evans

Submitted to the Department of Mathematics
on April 23, 2010, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

This thesis consists of two results.

The first result is a strong maximum principle for certain parabolic systems of equations, which, for illustrative purposes, I consider as reaction-diffusion systems. Using the theory of viscosity solutions, I give a proof which extends the previous theorem to no longer require any regularity assumptions on the boundary of the convex set in which the system takes its values.

The second result is an approximation scheme for reflected stochastic differential equations (SDE) of the Stratonovich type. This is a joint result with Professor Daniel W. Stroock. We show that the distribution of the solution to such a reflected SDE is the weak limit of the distribution of the solutions of the reflected SDEs one gets by replacing the driving Brownian motion by its N-dyadic linear interpolation. In particular, we can infer geometric properties of the solutions to a Stratonovich reflected SDE from those of the solutions to the approximating reflected SDE.

Thesis Supervisor: Daniel W. Stroock
Title: Professor of Mathematics

Acknowledgments

First and foremost, I thank my advisor Daniel Stroock. Professor Stroock has shared his mathematical expertise to help me in my research and has devoted an inordinate amount of time to my mathematical education, including having as many as three advisor meetings a week! Without him this thesis would not have been possible.

I would also like to thank my committee members, David Jerison and Scott Sheffield, for reading my thesis and offering me suggestions. In particular I thank Professor Jerison for serving as a temporary advisor when my advisor was away and for his advice concerning the first result of this thesis.

I would like to thank my family for their constant support throughout my time as a graduate student. In particular, I thank my dad for his useful advice, be it about specific math problems or the life of a mathematician in general.

I would also like to thank my fellow math graduate students. Whenever I needed advice about math concepts or had questions about graduate life, they were there for me.

Last but not least, I would like to thank my friends at MIT. Hanging out with my friends at night provided a nice balance to the world of math I spent my days in and helped me keep my sanity.

Contents

Chapter 1

First Result: Introduction

My first result consists of proving a version of the strong maximum principle for solutions of vector valued diffusion equations. The classical version of the weak and strong maximum principles can be summarized as follows: A solution u to an (unspecified) partial differential equation is said to satisfy the *weak maximum principle* if it has the property that the maximum value achieved by u is achieved on the boundary of the domain of the PDE. u is said to satisfy the *strong maximum principle* if, in addition, it has the property that if u achieves its maximum in the interior of its domain, it is then the case that u is constant.

When u takes values in $\mathbb{R}^k$, it no longer makes sense to talk about the "maximum of u" and so we must modify the definitions of the weak and strong maximum principles as follows: Given a convex set $K \subset \mathbb{R}^k$, a solution u to an (unspecified) partial differential equation is said to satisfy the *weak maximum principle with respect to* K if it has the property that when its boundary data takes values in K then u takes values in K at all points in its domain. u is said to satisfy the *strong maximum principle with respect to* K if, in addition, it has the property that if u takes a value in ∂K at some point (x_0, t_0) in the interior of its domain, it is then the case that $u \in \partial K$ at all points in the closure of its domain such that $t \leq t_0$.

The main result of the first half of this thesis concerns the weak and strong maximum principle for the following PDE:

$$\begin{cases} u_t = D(x,t,u)\sum_{i,j} a_{ij}(x,t)u_{x_ix_j} + \sum_i M_i(x,t,u)u_{x_i} + \phi(x,t,u), \text{ in } \Omega \\ u = g, \text{ on } \partial\Omega \end{cases} \tag{1.1}$$

where $\Omega = X \times (0,\infty)$ for some open and connected $X \subset \mathbb{R}^n$, the a_{ij} are real-valued, ϕ takes values in $\mathbb{R}^k$, and $D(x,t,z)$ and each of the $M_i(x,t,z)$ take values in the space of $k \times k$ matrices. We also assume that D and the matrix $\{a_{ij}\}$ take symmetric non-negative definite values. The system of PDEs (1.1) can be thought of as a reaction-diffusion system, and I adopt this view in this thesis to aid in intuition.

Proving that solutions to (1.1) satisfy the weak and strong maximum principle is not as simple as just applying the classical versions to the functions $\ell(u)$ for supporting hyperplanes ℓ of K. An example which shows why this approach cannot work is the following: Let $X = \mathbb{R}$, let K be the unit circle in R^2, let $\phi(x,t,u) = \phi(u) = (-u_2, u_1)$ and let $u(x,0) = g(x) \equiv (1,0)$. Then it is clear that $u(x,t) = (\cos(t), \sin(t))$ is a solution to (1.1) which visits and leaves every supporting hyperplane many times (This example is fleshed out in Chapter 3).

In his 1975 paper [17], H. Weinberger considers the case where $D = I$ and the M_i are real-valued functions. In this case, he proves that, under mild regularity assumptions, any solution to (1.1) satisfies the weak maximum principle with respect to any set K for which the vector field ϕ is "inward pointing" in the following sense: At each $v \in \partial K$, the inner product of $\phi(x,t,v)$ with $u - v$ is non-negative for all $u \in K$. Weinberger goes on to prove that any solution to (1.1) which satisfies the weak maximum principle with respect to a convex set K satisfying the inward pointing condition above also satisfies the strong maximum principle with respect to K under the additional assumption that ∂K satisfies what Weinberger refers to as the "slab condition".

In their 1977 paper [3], K. Chueh, C. Conley, and J. Smoller prove that, under mild regularity assumptions, any solution to (1.1) satisfies the weak maximum principle with respect to any set K for which the vector field ϕ is inward pointing and such that for every point $v \in \partial K$, we have that every normal vector to the boundary (there may be more than one where the boundary is not differentiable) at v is a left eigenvector of each of $D(x,t,v)$ and the $M_i(x,t,v)$ for all x and t.

In his 1990 article [16], X. Wang extends Weinberger's proof of the strong maximum principle to the case of general D and M_i, under the eigenvector condition of Chueh, Conley, and Smoller. For his proof, Wang also makes the additional assumption that the boundary of K is C^2. Under this assumption, the distance function, d, to the boundary of K is a C^2 function near the boundary of K, and Wang's argument is to show that $d(u)$ satisfies the PDE:

$$\frac{\partial}{\partial t} d(u(x,t)) \geq \mathcal{L} d(u(x,t)) - C(x,t) d(u(x,t)) \tag{1.2}$$

(here $C(x,t)$ depends on the Lipschitz constant of ϕ) and apply the classical strong maximum principle to this PDE.

In my paper [8], I have extended Wang's argument to apply to *any* convex set K satisfying the inward pointing and eigenvector conditions. This is a tight result as it is easy to provide counterexamples when the convexity of K or either condition is relaxed. For my proof, I show that while, in general, $d(u(x,t))$ may not be twice differentiable, $d(u(x,t))$ is still a super solution to a certain parabolic differential equation in the viscosity sense. I then invoke a strong maximum principle for viscosity solutions (provided by F. Da Lio in [6]) to achieve the desired result.

In Chapter 2, I give an overview of reaction diffusion systems. In particular I introduce the "blob picture" way of viewing reaction-diffusion systems.

In Chapter 3, I give an overview of various maximum principles starting from the classical ones and progressing to the maximum principles for systems. In particular I give some examples which motivate the theorems for the various maximum principles.

In Chapter 4, I give an outline of my proof from[8] for the simple system of PDE $u_t = \Delta u + \phi u$. This system of PDE is simpler than the more general system (1.1) but the key points of the proof remain the same. In particular, in this simple case it is easy to see how the theory of viscosity solutions is used.

In Chapter 5, I present my proof from[8] in its entirety.

Finally, for completeness, in Chapter 6, I present DaLio's proof of the strong maximum principle for viscosity solutions from [6], but in the simple case of the linear PDE we are concerned with.

Also for completeness, I have included in Appendix A the basic theory and results of viscosity solution theory that are needed in this thesis.

Chapter 2

Reaction-Diffusion Systems

As the main result of the first half of this thesis is a strong maximum principle for reaction-diffusion systems, we will begin by giving a brief overview of reaction-diffusion systems.

2.1 Diffusion Equations

We begin with the standard model for diffusion. Suppose we have a substance (e.g. heat or some chemical) which diffuses within a connected open region $X \subset \mathbb{R}^d$. We let $u(x,t)$ represent the density of the substance at position x and time t. Letting $\Omega = X \times (0,\infty)$ we have that $u : \Omega \to \mathbb{R}$ and we describe the diffusion via a PDE which u solves. The most familiar diffusion PDE is the heat equation which models the flow of heat:

$$\frac{\partial u}{\partial t} = \Delta u \tag{2.1}$$

More generally, a diffusion PDE is of the form

$$\frac{\partial u}{\partial t} = \mathcal{L}u \tag{2.2}$$

where

$$\mathcal{L} = \sum_{i,j} a_{ij}(x,t)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_i b_i(x,t)\frac{\partial}{\partial x_i} \tag{2.3}$$

where $\{a_{ij}\}$ is locally uniformly positive definite. For simplicity, we will consider the case that $\mathcal{L} = \Delta$.

Suppose we have the following PDE which represents, say, the heat in a rod of length L whose ends are held at a constant temperature 0: Let X be the interval $[0, L]$ and let u satisfy the diffusion PDE (with Dirichlet boundary conditions)

$$\begin{cases} \frac{\partial u}{\partial t} = \Delta u, \text{ in } \Omega \\ u(x,0) = 1, x \in X \\ u(0,t) = 0, t > 0 \\ u(L,t) = 0, t > 0 \end{cases}$$

There are two standard ways in which to present the solution $u(x,t)$ graphically. The first is a three dimensional graph with x, t, and u as axes (See Figure 2-1).

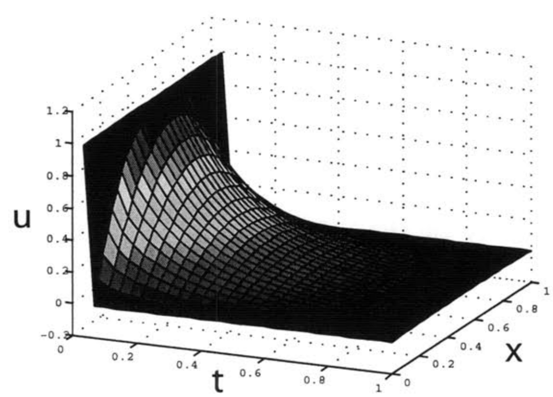

Figure 2-1:

This presentation lets us see the evolution of the heat flow in the rod all at once. The second is to give a series of snapshots of the heat in the rod at different times (See Figure 2-2).

For the purposes of this thesis, it will be convenient to present the diffusion in a

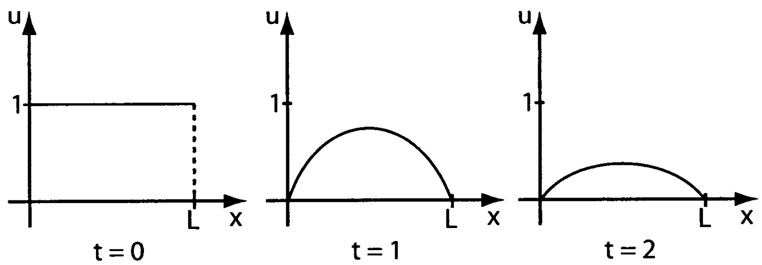

Figure 2-2:

third way which I will henceforth call the "blob picture". This corresponds to a series of snapshots at different times, but with the x axis removed (See Figure 2-3).

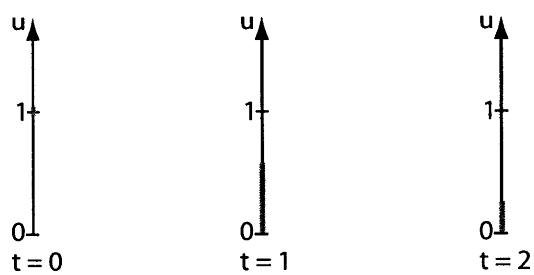

Figure 2-3:

The obvious deficiency of this presentation is that for a given time t, we only see the values that u takes and not where in X it takes those values. However the blob picture will be ideal for the purposes of this thesis for two reasons: First, in a graph we can have at most three axes and in the blob picture we can use every axis for the components of u (We will see this when we look at systems of diffusions where u is vector valued). Second, the blob picture is naturally connected to the maximum principles we will see below.

2.2 System of Diffusions

Now suppose instead of one diffusing substance we have n. We label their densities as $u_1(x,t),\dots,u_n(x,t)$. Suppose further that each substance diffuses in the same way, i.e. $\frac{\partial u_k}{\partial t} = \mathcal{L}u_k$ for some $\mathcal{L}$ of the form (2.3). Then letting $\mathbf{u}(x,t) = (u_1(x,t),\dots,u_n(x,t))$ we can write this more compactly as

$$\frac{\partial \mathbf{u}}{\partial t} = \mathcal{L}\mathbf{u} \tag{2.4}$$

For example, suppose that $X = \{(x_1,x_2) : x_1^2 + x_2^2 < 1\}$ is the open unit disk in $\mathbb{R}^2$ and $\mathbf{u} = (u_1,u_2)$ satisfies the diffusion PDE with Neumann boundary condition:

$$\begin{cases} \frac{\partial \mathbf{u}}{\partial t} = \Delta\mathbf{u}, \text{ in } \Omega \\ u_1(x,0) = x_1 + 10, x \in X \\ u_2(x,0) = x_2 + 10, x \in X \\ \frac{\partial \mathbf{u}}{\partial \nu} = 0, \text{ on } \partial X \times (0,\infty) \end{cases}$$

where $\nu(x)$ represents the unit normal to the boundary for $x \in \partial X$. We can view $u_1(x,t)$ and $u_2(x,t)$ as being the concentration of two different chemicals in an insulated circular region, each of which diffusing as heat would. We present the dynamics of this system with our blob picture (see Figure 2-4).

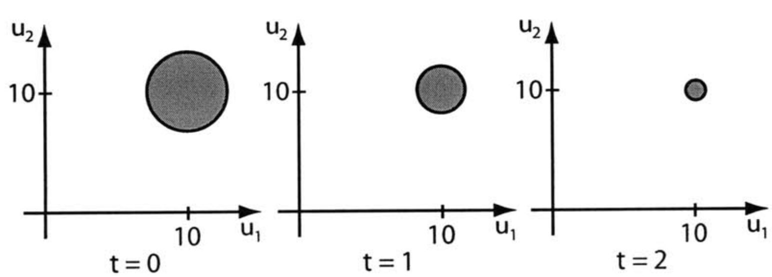

Figure 2-4:

In the blob picture, our blob contracts! This is not surprising since as each chemical diffuses its concentration averages out. (A note of caution is in order. You cannot tell just by looking at the blob how it will contract... after all, the dynamics of the diffusion depend on where in X the various values of $\mathbf{u}$ are coming from).

The fact that, for a system of diffusions with Neumann boundary condition, the blob in its blob picture contracts is closely related to the weak and strong maximum principles we will define in the next overview.

2.3 Reaction Systems

We will now let the substances we are modeling interact. The word "reaction" in reaction-diffusion systems comes from the case in which we model the concentration of different chemicals which react with one another. However, since I know very little about chemistry, we shall instead focus on biological models of different animal species which interact with each other.

A familiar model, often seen in an introductory ODE course, is the Lotka-Volterra model for predator-prey interaction. Suppose $u_1(t)$ and $u_2(t)$ represent the deer and wolf populations in a given area over time. Suppose u_1 and u_2 satisfy the reaction system:

$$\frac{du_1}{dt} = au_1 + bu_1u_2 \tag{2.5}$$

$$\frac{du_2}{dt} = cu_1u_2 + du_2 \tag{2.6}$$

where $a, c > 0$ and $b, d < 0$ (i.e. deer thrive in the absence of wolves and wolves starve in the absence of deer). In general, we consider a reaction system modeling n species of the form

$$\frac{d\mathbf{u}}{dt} = \phi(\mathbf{u}) \tag{2.7}$$

for some Lipschitz continuous vector field ϕ (We take ϕ be Lipschitz to ensure that (2.7) has a unique solution, but we will see later that the Lipschitz continuity of ϕ

also plays a strong role in proving various strong maximum principles). We can view the dynamics of our reaction system as a dot flowing along the vector field ϕ. For our deer-wolf system we have the following picture (see Figure 2-5).

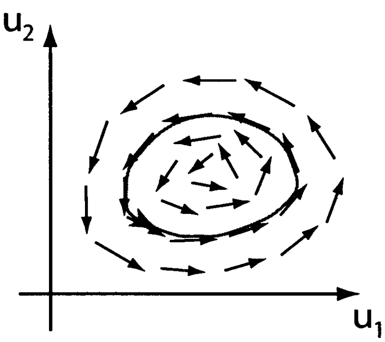

Figure 2-5:

2.4 Reaction-Diffusion Systems

Reaction-diffusion systems are a combination of our previous two models. That is, we model the concentration of n substances $u_1(x,t), \ldots, u_n(x,t)$ which not only diffuse but also react. Mathematically, we consider the following reaction diffusion system:

$$\frac{\partial \mathbf{u}}{\partial t} = \mathcal{L}\mathbf{u} + \phi(\mathbf{u}) \tag{2.8}$$

for a diffusion operator $\mathcal{L}$ of the form (2.3) and a Lipschitz continuous vector field ϕ. As an example, we consider the dynamics of a deer and wolf population confined to the disk(perhaps an island) $X = \{(x_1, x_2) : x_1^2 + x_2^2 < 1\}$ and beginning with initial population densities $f(x)$ and $g(x)$ respectively:

$$\begin{cases} \frac{\partial u_1}{\partial t} = \Delta u_1 + au_1 + bu_1u_2, \text{ in } \Omega \\ \frac{\partial u_2}{\partial t} = \Delta u_2 + cu_1u_2 + du_2, \text{ in } \Omega \\ u_1(x,0) = f(x), x \in X \\ u_2(x,0) = g(x), x \in X \\ \frac{\partial u_1}{\partial \nu} = 0, \text{ on } \partial X \times (0,\infty) \\ \frac{\partial u_2}{\partial \nu} = 0, \text{ on } \partial X \times (0,\infty) \end{cases} \tag{2.9}$$

where $a, c > 0$ and $b, d < 0$. We have the following blob picture for the dynamics of the deer-wolf system. We overlay the vector field $\phi(\mathbf{u}) = \begin{pmatrix} au_1 + bu_2 \\ cu_1 + du_2 \end{pmatrix}$ to get a better feel for what is going on (See Figure 2-6).

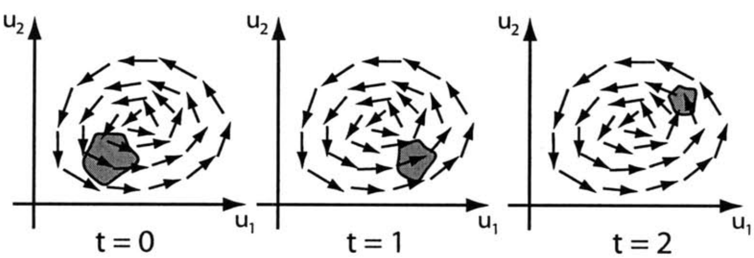

Figure 2-6:

Not only is the blob contracting, but it is being pushed by the vector field ϕ! To see this mathematically, note that each point (v_1, v_2) in the blob at time $\bar{t}$ corresponds to at least one point $\bar{x} \in X$ such that $u_1(\bar{x}, \bar{t}) = v_1$ and $u_2(\bar{x}, \bar{t}) = v_2$. How are the densities u_1 and u_2 instantaneously changing at the point $(\bar{x}, \bar{t})$? We have

$$\frac{\partial \mathbf{u}}{\partial t}(\bar{x}, \bar{t}) = \Delta \mathbf{u}(\bar{x}, \bar{t}) + \phi(\mathbf{u}(\bar{x}, \bar{t}))$$

and we can think of the first term as the "contraction force" and the second term as the "vector field force".

2.5 More Complex Reaction-Diffusion Systems

Up until now we have been assuming our different substances/species diffuse in the same way. But we could, of course, have a reaction-diffusion system of the form:

$$\begin{cases} \frac{\partial u_1}{\partial t} = \xi \Delta u_1 + \phi_1(\mathbf{u}) \\ \frac{\partial u_2}{\partial t} = \eta \Delta u_2 + \phi_2(\mathbf{u}) \end{cases}$$

where the diffusion coefficients ξ and η differ. In the case that $\xi > \eta$ we have that the first substance diffuses more quickly than the second (perhaps deer are faster than wolves). But we will consider more complex models than this.

The most general model for reaction-diffusion systems we will consider is the following:

$$\frac{\partial \mathbf{u}}{\partial t} = D(x,t,\mathbf{u}) \sum_{i,j} a_{ij}(x,t) \frac{\partial^2 \mathbf{u}}{\partial x_i \partial x_j} + \sum_i M_i(x,t,\mathbf{u}) \frac{\partial \mathbf{u}}{\partial x_i} + \phi(x,t,\mathbf{u}), \tag{2.10}$$

where D and the M_i are matrix valued, D and $\{a_{ij}\}$ are locally uniformly positive definite, and ϕ is Lipschitz continuous. The case where $D \equiv I$ and the M_i are diagonal reduces to the simpler reaction diffusion system (2.8). The case where D and the M_i are diagonal corresponds to the system (2.8) with different diffusion rates.

I am sorry to say that I don't have a good idea for what (2.10) represents in general. However, we can examine the dynamics of the simpler system of diffusions

$$\frac{\partial \mathbf{u}}{\partial t} = D(\mathbf{u}) \Delta \mathbf{u}$$

to get a feel for what the matrix-valued D does in our blob picture.

First, let's recall how an $n \times n$ symmetric positive definite matrix A acts on a vector $\mathbf{v}$. We know that the eigenvalues of A are positive and real and that A has a full set of n eigenvectors which are mutually orthogonal (see Figure 2-7).

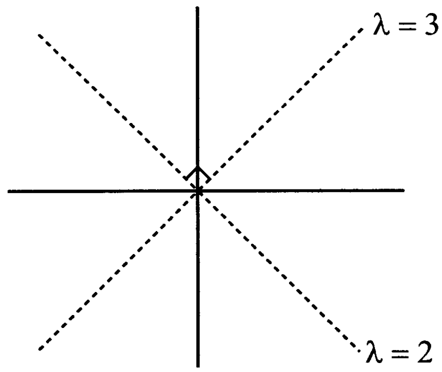

Figure 2-7:

Suppose we graph $\mathbf{v}$ and $A\mathbf{v}$ for various choices of $\mathbf{v}$ (see Figure 2-8).

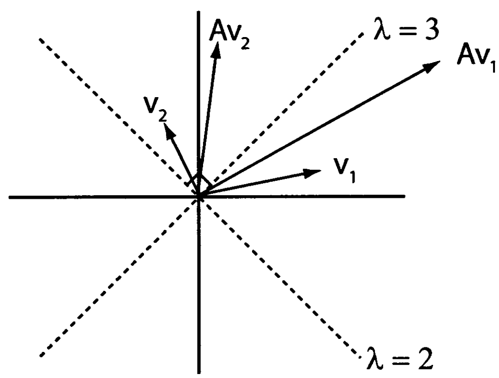

Figure 2-8:

Since all of its eigenvalues are positive and real, we see that A stretches or compresses $\mathbf{v}$ along each of its eigenvector axes. It is clear therefore that $\mathbf{v}$ and $A\mathbf{v}$ lie in the same eigen-orthant.

So at each $\mathbf{u}$, $D(\mathbf{u})$ is a symmetric positive definite matrix which perturbs the vector $\Delta\mathbf{u}$ within its eigen-orthant. Graphically we overlay an "eigen-axes field" over our blob picture (See Figure 2-9).

Then, at a given (x, t), we can draw the vectors $\Delta\mathbf{u}(x, t)$ and $D(\mathbf{u})\Delta\mathbf{u}(x, t)$ at the point $\mathbf{u}(x, t)$ (see Figure 2-10).

Since $\frac{\partial \mathbf{u}}{\partial t} = D(\mathbf{u})\Delta\mathbf{u}$, this latter vector shows how the value of $\mathbf{u}$ at (x, t) changes instantaneously. In the blob picture, this vector shows where the image of the point (x, t) moves in the blob. We see that our blob contracts as before, but now its

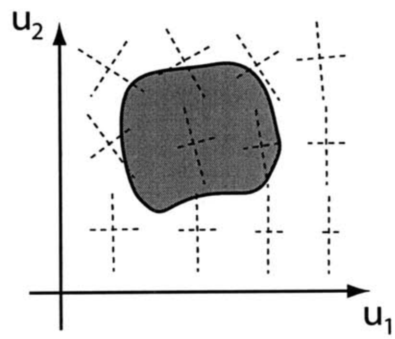

Figure 2-9:

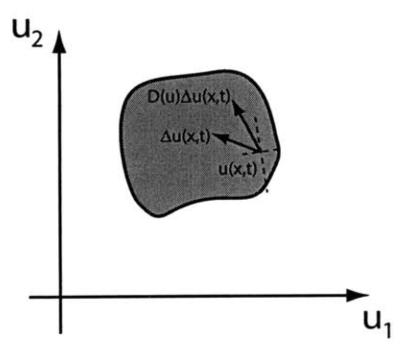

Figure 2-10:

contraction at each point is perturbed by the eigenbasis at that point.

Chapter 3

Maximum Principles

3.1 Diffusion Equations

In this section of the overview we will explain the various extensions of the classical weak and strong maximum principles. As before, let X be a connected open subset of $\mathbb{R}^d$ and let $\Omega = X \times (0, \infty)$. We begin by recalling the classical weak and strong maximum principles for a diffusion system of the form (2.2).

Definition 3.1.1. A function $u : \Omega \longrightarrow \mathbb{R}$ is said to satisfy the *weak maximum principle* if it has the property that the maximum value of u in $\bar{\Omega}$ is achieved on the boundary of its domain, $\partial\Omega$.

A function $u : \Omega \longrightarrow \mathbb{R}$ is said to satisfy the *strong maximum principle* if, in addition, it has the property that if the maximum value of u in $\bar{\Omega}$ is achieved at an interior point $(x_0, t_0) \in \Omega$, then $u(x, t) \equiv u(x_0, t_0)$(i.e. u is constant and equal to this maximum value) for $(x, t) \in \bar{\Omega}$ with $t \leq t_0$.

The weak and strong minimum principles are defined analogously.

Under mild regularity assumptions on the domain X, the operator $\mathcal{L}$, and the solution u, we have that solutions of (2.2) satisfy the weak and strong maximum/minimum principles.

A classic example where these principles can be seen is in the diffusion of heat through a rod of length L whose ends are first held at a constant temperature 1 and then at a constant temperature 0. Let $X = [0, L]$, $\Omega = X \times [0, \infty)$ and suppose u satisfies the diffusion PDE

$$\begin{cases} \frac{\partial u}{\partial t} = \Delta u, \text{ in } \Omega \\ u(x,0) = 1, 0 \leq x \leq L \\ u(0,t) = 1, 0 < t \leq 1 \\ u(L,t) = 1, 0 < t \leq 1 \\ u(0,t) = 0, t > 1 \\ u(L,t) = 0, t > 1 \end{cases}$$

We graph the solution to this diffusion in Figure 3-1.

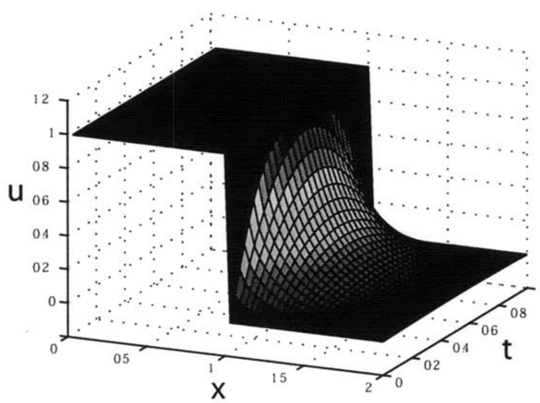

Figure 3-1:

Notice that the weak maximum principle is satisfied: The maximum value u attains is 1 and it does so on the boundary. Notice that the strong maximum principle is also satisfied: Wherever u attains the value 1 in the interior, it is constant up until that time.

Consider next the heat diffusion in a rod of length L whose ends are insulated:

$$\begin{cases} \frac{\partial u}{\partial t} = \Delta u, \text{ in } \Omega \\ u(x,0) = f(x), \text{ on } X \times \{0\} \\ \frac{\partial u}{\partial \nu} = 0, \text{ on } \partial X \times (0, \infty) \end{cases}$$

Notice that the weak maximum principle is satisfied: u attains its maximum value on the boundary. Furthermore, as u never attains it maximum value in the interior, the strong maximum principle holds by default.

Let's examine this diffusion via the blob picture (let $a = \inf_{x\in[0,L]} f(x)$, $b = \sup_{x\in[0,L]} f(x)$) (See Figure 3-2).

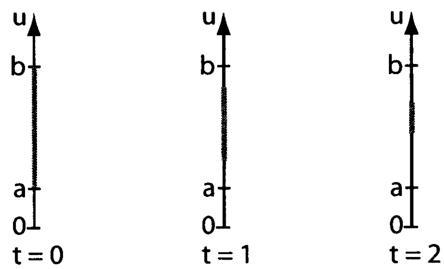

Figure 3-2:

Sure enough the blob contracts. In particular we see that the minimum value that u achieves is a and the maximum value that u achieves is b. And so u achieves its maximum and minimum on the boundary. This suggests a new formulation of the weak maximum/minimum principle (we will henceforth drop the term "minimum" and refer only to maximum principles):

Definition 3.1.2. We say that u satisfies the *weak maximum principle (version 2)* if the following holds for each closed interval $[a, b]$: If $u(x, t) \in [a, b]$ at all boundary points $(x, t) \in \partial\Omega$, then $u(x, t) \in [a, b]$ for all points $(x, t) \in \bar{\Omega}$.

And this can be further restated as

Definition 3.1.3. We say that u satisfies the *weak maximum principle (version 3)* if the following holds for each closed convex set K:

If $u(x, t) \in K$ at all boundary points $(x, t) \in \partial\Omega$, then $u(x, t) \in K$ for all points $(x, t) \in \bar{\Omega}$.

Some thought shows that the three versions of the weak maximum principle are in fact the same. It is this third version which will extend to the case of systems of PDE and so it is this third version that we will focus on.

What should the restatement in terms of closed convex sets K look like for the strong maximum principle? In the case when $K = [a,b]$ and $u(x,t) \in [a,b]$ for $(x,t) \in \bar{\Omega}$, we want to say that if $u(x_0,t_0) = b$ for some $(x_0,t_0) \in \Omega$ then $u(x,t) \equiv b$ for $t \le t_0$, and the same for a. We can think of a and b as either the boundary or the extreme points of the closed convex set $[a,b]$ (Recall that an extreme point of a convex set is a point that cannot be expressed as a convex combination of two other points in the set). These two interpretations yield the following two restatements of the strong maximum principle:

Definition 3.1.4. We say that u which satisfies the weak maximum principle (version 3) also satisfies the *strong maximum principle (boundary version)* if the following holds for each closed convex set K:

If $u(x_0,t_0) \in \partial K$ at an interior point $(x_0,t_0) \in \Omega$, then $u(x,t) \in \partial K$ for all points $(x,t) \in \bar{\Omega}$ with $t \le t_0$.

Definition 3.1.5. We say that u which satisfies the weak maximum principle (version 3) also satisfies the *strong maximum principle (extreme point version)* if the following holds for each closed convex set K:

If for some extreme point $v \in K$, $u(x_0,t_0) = v$ at an interior point $(x_0,t_0) \in \Omega$, then $u(x,t) \equiv v$ for all points $(x,t) \in \bar{\Omega}$ with $t \le t_0$.

3.2 System of Diffusions

We now consider the system of diffusions (2.4). How should we define the weak and strong maximum principles for such a system? The classical formulation in terms of maximum and minimum doesn't make sense when $\mathbf{u}$ is vector valued, but our reformulation in terms of closed convex sets carries over to higher dimensions. And so in the case where $\mathbf{u}$ is vector valued, we again define what it means for $\mathbf{u}$ to satisfy the weak and strong maximum principles by Definitions 3.1.3, 3.1.4, and 3.1.5.

Again, under mild regularity assumptions on Ω, $\mathcal{L}$, and $\mathbf{u}$, we have that solutions to (2.4) satisfy these three maximum principles.

It is instructive to note the connection between the maximum principles and our blob picture. In the blob picture, a system of diffusions (with Neumann boundary conditions) can be viewed as a contracting blob. This insight makes it easy to find examples which illustrate the importance of the convexity assumption for K in Definition 3.1.3. Indeed, if K were not convex, it is easy to construct examples where the blob contracts outside of K (See Figure 3-3).

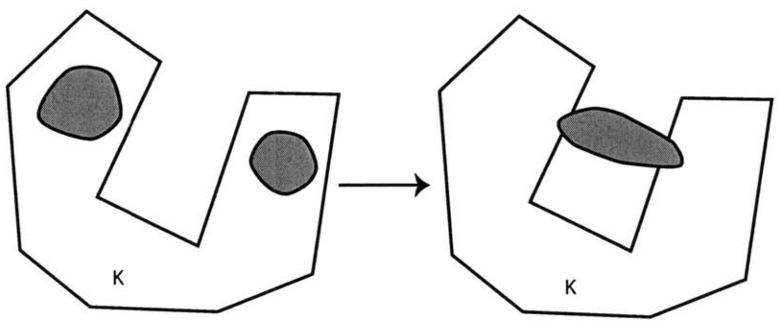

Figure 3-3:

3.3 Reaction-Diffusion Systems

We first consider reaction-diffusion systems of the form (2.8). To make our analysis simpler, we consider the specific reaction-diffusion system:

$$\frac{\partial \mathbf{u}}{\partial t} = \Delta \mathbf{u} + \phi(\mathbf{u})$$

where, as usual, ϕ is Lipschitz continuous.

Recall that in our blob picture, we can view the dynamics of this reaction-diffusion system as a blob which is simultaneously contracting and being pushed by the vector field ϕ. What should the weak and strong maximum principles look like in this case?

Consider as an example the reaction-diffusion system (2.9) (Deer-Wolf Island).

Suppose we fix the following convex set K in our blob picture at time $t = 0$ (See Figure 3-4).

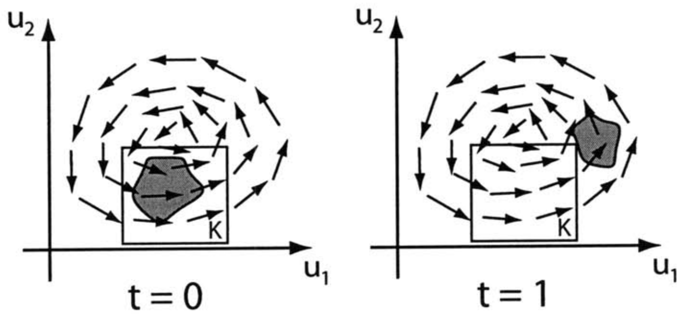

Figure 3-4:

If the weak maximum principle stated in Definition 3.1.3 were to hold, the blob should remain in K. However, at time $t = 1$, the blob will in fact have left K (See Figure 3-4).

On the other hand, if we took K to be the circle seen below, the blob will stay in K for all time (see Figure 3-5).

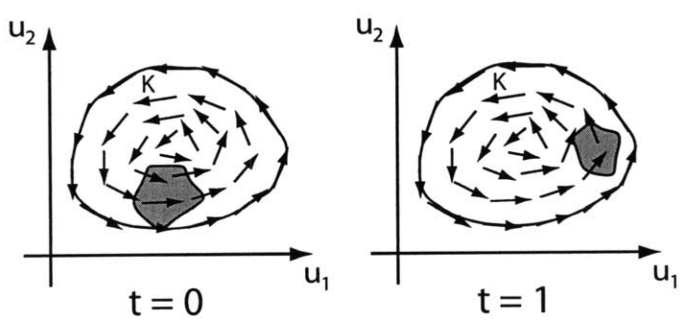

Figure 3-5:

So the weak maximum principle only holds for some K. We need that K be such that the vector field ϕ "points inward" on (including parallel to) the boundary of K. Whether or not the weak maximum principle holds now depends on the choice of K and so we include this in our new definition of the weak maximum principle:

Definition 3.3.1. For a given closed convex set K, we say that $\mathbf{u}$ satisfies the *weak maximum principle with respect to* K if the following holds:

If $\mathbf{u}(x,t) \in K$ at all boundary points $(x,t) \in \partial\Omega$, then $\mathbf{u}(x,t) \in K$ for all points $(x,t) \in \bar{\Omega}$.

In Theorem 4.4 of [7], S.D. Eidel'man shows that under mild regularity assumptions on Ω, $\mathcal{L}$, and $\mathbf{u}$, every solution $\mathbf{u}$ of (2.8) satisfies the weak maximum principle with respect to K for every closed convex set K such that ϕ points inward on the boundary of K. This is exactly what one would expect from the blob picture!

Note the importance of our assumption that ϕ be Lipschitz continuous: When ϕ is not Lipschitz continuous, the weak maximum principle with respect to K may not hold for solutions to (2.8) even when ϕ points inward on the boundary.

For example, consider the vector field given by

$$\phi(\mathbf{u}) = \begin{pmatrix} 1 \\ -sgn(u_2)\sqrt{|u_2|} \end{pmatrix}$$

and let $K = \{\mathbf{u} : u_2 \geq 0\}$. Standard ODE theory shows that there are integral curves of ϕ which begin in K and exit K despite the fact that ϕ points inward on the boundary (See Figure 3-6).

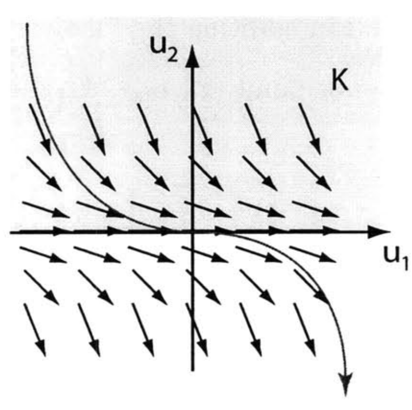

Figure 3-6:

With this in mind, we can construct the following counter example: Let $\mathbf{u}(x,t)$ be a solution to (2.8) which is independent of x. Then $\mathcal{L}u = 0$ and so $\mathbf{u}(x,t) = \mathbf{u}(t)$

solves $\frac{\partial \mathbf{u}}{\partial t} = \phi(\mathbf{u})$. As $\mathbf{u}$ is independent of x, it is clear that in our blob picture our blob is now just a point which follows an integral curve of ϕ. Taking ϕ and K as above we see that the weak maximum principle with respect to K doesn't hold for $\mathbf{u}$ (See Figure 3-7).

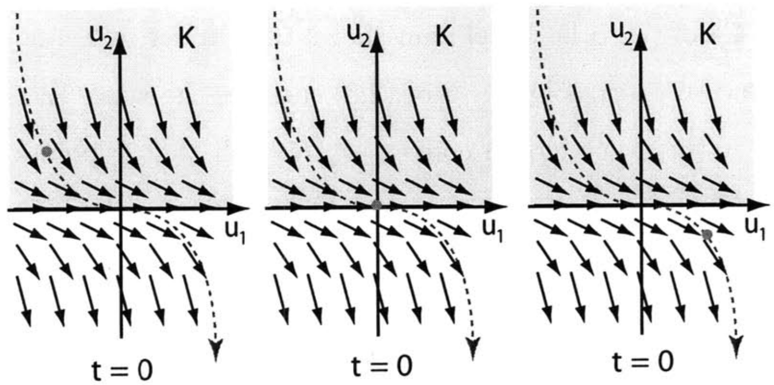

Figure 3-7:

This trick of considering a point-blob along with ODE theory is useful for constructing counter examples and we will use it again below.

We can extend our strong maximum principles similarly:

Definition 3.3.2. For a given closed convex set K, we say that $\mathbf{u}$ satisfies the *boundary strong maximum principle with respect to* K if $\mathbf{u}$ satisfies the weak maximum principle with respect to K and in addition the following holds:

If $\mathbf{u}(x_0, t_0) \in \partial K$ at an interior point $(x_0, t_0) \in \Omega$, then $\mathbf{u}(x, t) \in \partial K$ for all points $(x, t) \in \bar{\Omega}$ with $t \leq t_0$.

Definition 3.3.3. For a given closed convex set K, we say that $\mathbf{u}$ satisfies the *extreme point strong maximum principle with respect to* K if $\mathbf{u}$ satisfies the weak maximum principle with respect to K and in addition the following holds:

If for some extreme point $v \in K$, $u(x_0, t_0) = v$ at an interior point $(x_0, t_0) \in \Omega$, then $u(x, t) \equiv v$ for all points $(x, t) \in \bar{\Omega}$ with $t \leq t_0$.

The main result of this thesis is the following theorem.

Theorem 3.3.4. *For each closed convex set K such that ϕ points inward on the boundary, every solution $\boldsymbol{u}$ of (2.8) which satisfies the weak maximum principle with respect to K also satisfies the boundary strong maximum principle with respect to K.*

Combined with Theorem 4.4 of [7], we have the following corollary.

Corollary 3.3.5. *Under mild regularity conditions on Ω, $\mathcal{L}$, and $\boldsymbol{u}$, every solution $\boldsymbol{u}$ of (2.8) satisfies the boundary strong maximum principle with respect to K for every closed convex set K such that ϕ points inward on the boundary.*

H. Weinberger first stated the boundary strong maximum principle for reaction-diffusion systems in his 1975 paper [17]. Weinberger proved Theorem 3.3.4 under an additional regularity condition on K which he called the "slab condition". In his 1990 paper [16], X. Wang, gives a geometric proof of Theorem 3.3.4, following Weinberger's arguments, in the case that the boundary of K is C^2. In this case, the distance function to the boundary of K, d, is C^2 in K (at least near the boundary), and so the boundary strong maximum principle is proved by applying the classical strong maximum principle to $d(\mathbf{u}(x,t))$. I have removed the regularity assumptions on K imposed by Weinberger and Chen by proving (3.3.4) using the techniques of viscosity solutions (see next section).

It is interesting to note that an analogous theorem to Theorem 3.3.4 for the extreme point strong maximum principle does *not* hold. We construct a simple counter example via the point-blob technique described above:

Let

$$\phi(\mathbf{u}) = \begin{pmatrix} -u_2 \\ u_1 \end{pmatrix}$$

and let $K = \mathbf{u} : u_1^2 + u_2^2 \leq 1$ be the closed unit disk. Then

$$\mathbf{u}(x,t) = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}$$

is a solution to (2.8) which is independent of x. We have the following blob picture (See Figure 3-8).

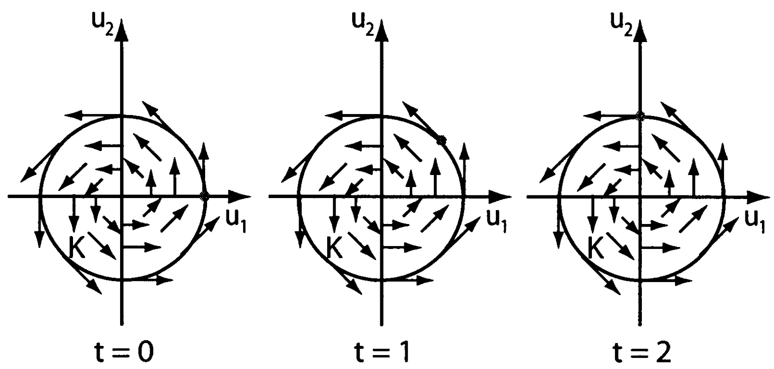

Figure 3-8:

Our blob is a point-blob which travels counter-clockwise around the unit circle. As every point on the unit circle is an extreme point of K, $\mathbf{u}$ takes values at extreme points of K in the interior of Ω, yet $\mathbf{u}$ is not constant at earlier times. And so we cannot hope to have a theorem like Theorem 3.3.4 for the extreme point strong maximum principle. Therefore, we conclude that the boundary point strong maximum principle is the true strong maximum principle for reaction-diffusion systems and henceforth it will be the only strong maximum principle we consider.

3.4 More Complex Reaction-Diffusion Systems

Finally, we would like to state the weak and strong maximum principles for the most general reaction diffusion system (2.10). To get a feel for things, let's look at the simpler equation

$$\frac{\partial \mathbf{u}}{\partial t} = D(\mathbf{u})\Delta \mathbf{u}$$

and in fact let's first look at the special case where $D(\mathbf{u})$ is constant and diagonal. In two dimensions we then have

$$\frac{\partial u_1}{\partial t} = a\Delta u_1 \tag{3.1}$$

$$\frac{\partial u_2}{\partial t} = b\Delta u_2 \tag{3.2}$$

$$\tag{3.3}$$

We can view u_1 and u_2 as the densities of two non-reacting chemicals. Suppose $b > a$. Then the second chemical diffuses faster than the first, and our blob picture might look like (See Figure 3-9).

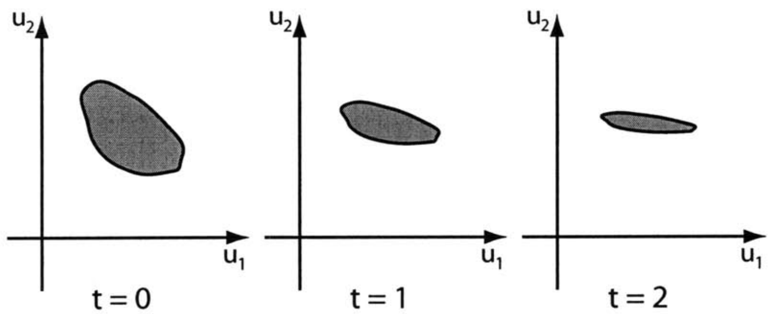

Figure 3-9:

That is, our blob contracts vertically faster than it does horizontally. Notice what happens when we overlay the following closed convex set K (See Figure 3-10).

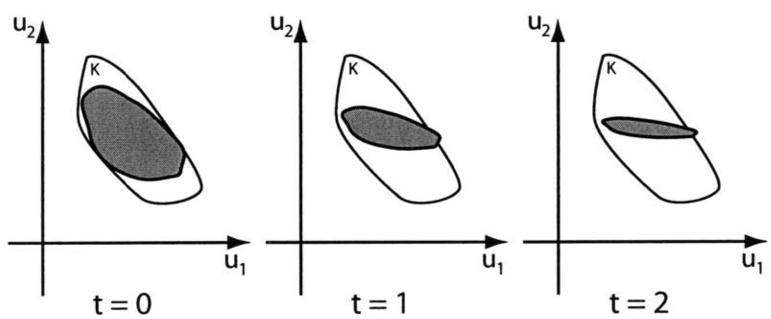

Figure 3-10:

Even without a vector field ϕ to push it, our blob has contracted outside of K! So we are going to need some additional conditions on K in order for a solution $\mathbf{u}$ of (2.10) to satisfy the weak maximum principle with respect to K.

Recall that we can view the dynamics of (3.1) in the blob picture by thinking of $D(\mathbf{u})$ as an eigen-axes field which perturbs the vectors $\Delta\mathbf{u}$. In the case where $D(\mathbf{u}) = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$, the eigen-axes field looks like Figure 3-11.

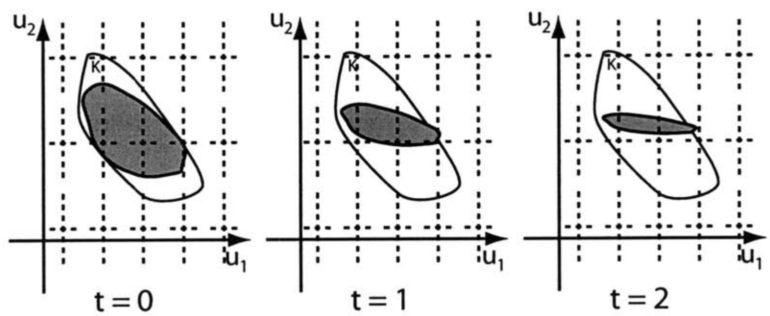

Figure 3-11:

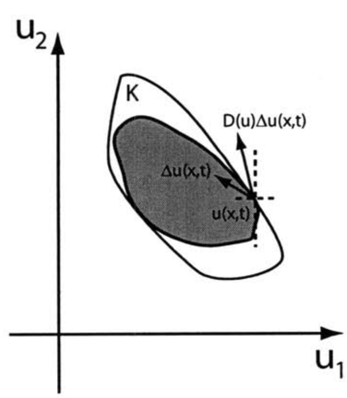

Figure 3-12:

Now let's see what is happening that causes the blob to escape K.

Notice in Figure 3-12 that at the boundary point of K, the vector $\Delta\mathbf{u}$ is perturbed by $D(\mathbf{u})$ to point outside the set K! How can we ensure that at each boundary point of K, the vector $D(\mathbf{u})\Delta\mathbf{u}$ points inward? We know that at each boundary point of K, $\Delta\mathbf{u}$ points inward and we also know that $D(\mathbf{u})$ only perturbs vectors within its eigen-orthants. Therefore, the key is to consider sets K which are aligned with $D(\mathbf{u})$ in the sense that for all $v \in \partial K$, the normal vector to K at v is an eigenvector of $D(\mathbf{u})$. This will ensure that at each $v \in \partial K$, the eigen-axes of $D(v)$ are tangent to ∂K and so $\Delta\mathbf{u}$ will not be perturbed to point outside of K. See Figure 3-13.

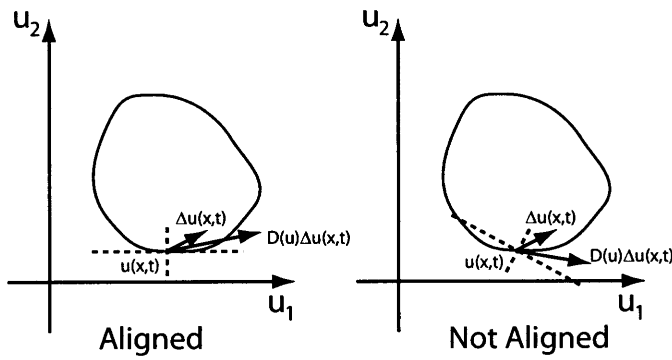

Figure 3-13:

So for a solution $\mathbf{u}$ of (2.10) to satisfy the weak maximum principle with respect to a closed convex set K, we need that ϕ points inward on K and that D is aligned with K in the sense explained above. It turns out we need a similar condition for each of the matrices M_i and we will want the normal vectors at points $v \in \partial K$ to be left-eigenvectors of the $M_i(v)$ (recall that as D is symmetric its left and right eigenvectors are the same and so this fits with our analysis above). We give the following definition:

Definition 3.4.1. A closed, convex set K is *compatible* with ϕ, D, and the M_i if ϕ points inward on K and for all $(x,t) \in \Omega$, $v \in \partial K$, and each vector ν normal to ∂K at v we have that ν is a left eigenvector of $D(x,t,v)$ and each of the $M_i(x,t,v)$.

In their 1977 paper [3], K. Chueh, C. Conley, and J. Smoller prove that, under mild regularity assumptions on Ω, $\mathcal{L}$, and $\mathbf{u}$, solutions $\mathbf{u}$ of (2.10) satisfy the weak maximum principle with respect to K for every K compatible with ϕ, D, and the M_i.

The main result of this thesis is the following theorem:

Theorem 3.4.2. *For each closed convex set K compatible with ϕ, D, and the M_i, every solution $\mathbf{u}$ of (2.10) which satisfies the weak maximum principle with respect to K also satisfies the boundary strong maximum principle with respect to K.*

Combined with the result of Chueh, Conley, and Smoller, we have the following corollary.

Corollary 3.4.3. *Under mild regularity conditions on Ω, $\mathcal{L}$, and $\mathbf{u}$, every solution $\mathbf{u}$ of (2.10) satisfies the boundary strong maximum principle with respect to K for every closed convex set K compatible with ϕ, D, and the M_i.*

In his 1990 paper [16], X. Wang, gives a geometric proof of Theorem 3.4.2 in the case that the boundary of K is C^2. I have removed the regularity assumptions on K imposed by Chen by proving (3.4.2) using the techniques of viscosity solutions (see next section).

Chapter 4

Overview of the Argument

In this section, we prove Theorem 3.3.4 for the simple reaction-diffusion system

$$\frac{\partial \mathbf{u}}{\partial t} = \Delta \mathbf{u} + \phi(\mathbf{u}) \tag{4.1}$$

The proof of the more general Theorem 3.4.2 for general reaction-diffusion systems 2.10 follows the same outline as the proof in this simple case and is given in the next section.

We assume that $\mathbf{u} \in C^{2,1}(\Omega; \mathbb{R}^k) \cap C(\bar{\Omega}; \mathbb{R}^k)$ satisfies (4.1) and that $\mathbf{u}(x,t) \in K$ for $(x,t) \in \partial\Omega$ (and hence for all $(x,t) \in \Omega$ by the weak maximum principle), where ϕ is Lipschitz and satisfies the inward pointing condition for K i.e.

$$\phi(z) \cdot \nu \geq 0$$

for all $z \in \partial K$ and inward pointing normal vectors ν to the boundary of K at z. Here we say that ν at $w \in \partial K$ is *inward pointing* if there exists a supporting hyperplane ℓ of K which touches K at w and has ν as its "inward pointing" normal.

Suppose also that $\mathbf{u}(x_0, t_0) \in \partial K$ for $(x_0, t_0) \in \Omega$. Our goal is then to show that $\mathbf{u}(x,t) \in \partial K$ for all $(x,t) \in X \times [0, t_0]$. As our proof extends X. Wang's proof from [16], we first sketch that proof of the boundary strong maximum principle for the case that the boundary ∂K is C^2 (and much of our notation is from [16] as well). For

$z \in K$, let

$$d(z) = \inf_{w \in \partial K} |z - w|$$

be the distance function to the boundary of K. As ∂K is C^2, it can be shown that $d(z)$ is also C^2 (at least near the boundary of K).

We know that $\mathbf{u}(x,t)$ is C^2 in the interior of Ω, and so we have that $d(\mathbf{u}(x,t))$ is C^2 (at least for (x,t) such that $\mathbf{u}(x,t)$ is near the boundary of K). If we let $\bar{d}(x,t) = d(\mathbf{u}(x,t))$, it is easy enough to show that (using $(\cdot,\cdot)$ to denote the Euclidian inner product)

$$\begin{aligned}\bar{d}_t(x,t) =& \Delta_x \bar{d}(x,t) + \nabla_z d(\mathbf{u}(x,t)) \cdot \phi(\mathbf{u}(x,t)) - \sum_i ([\nabla_z^2 d(\mathbf{u}(x,t))]\mathbf{u}_{x_i}(x,t), \mathbf{u}_{x_i}(x,t)) \\ \geq& \Delta_x \bar{d}(x,t) + \nabla_z d(\mathbf{u}(x,t)) \cdot \phi(\mathbf{u}(x,t))\end{aligned}$$

where the inequality holds because $d(z)$ is a concave function in K. Next, for $z \in K$, let $y(z)$ be the point on ∂K closest to z. By our inward pointing condition we have that for any $z \in K$,

$$\nabla_z d(y(z)) \cdot \phi(y(z)) \geq 0$$

and so combining this with our previous inequality we get that

$$\bar{d}_t(x,t) \ \geq \ \Delta_x \bar{d}(x,t) + \nabla_z d(\mathbf{u}(x,t)) \cdot \phi(\mathbf{u}(x,t)) - \nabla_z d(y(\mathbf{u}(x,t))) \cdot \phi(y(\mathbf{u}(x,t)))$$

We next note that $\nabla_z d(z) = \nabla_z d(y(z))$ and that $|\nabla_z d(z)| = 1$ for all $z \in K$. And so we have that

$$\begin{aligned}\bar{d}_t(x,t) &\geq \Delta_x \bar{d}(x,t) + \nabla_z d(\mathbf{u}(x,t)) \cdot \phi(\mathbf{u}(x,t)) - \nabla_z d(\mathbf{u}(x,t)) \cdot \phi(y(\mathbf{u}(x,t))) \\ &= \Delta_x \bar{d}(x,t) + \nabla_z d(\mathbf{u}(x,t)) \cdot (\phi(\mathbf{u}(x,t)) - \phi(y(\mathbf{u}(x,t)))) \\ &\geq \Delta_x \bar{d}(x,t) - C|\mathbf{u}(x,t) - y(\mathbf{u}(x,t))| \\ &= \Delta_x \bar{d}(x,t) - C\bar{d}(x,t)\end{aligned}$$

where C is the Lipschitz constant for ϕ. We now apply the classical strong maximum principle to the partial differential inequality

$$\bar{d}_t(x,t) \geq \Delta_x \bar{d}(x,t) - C\bar{d}(x,t) \tag{4.2}$$

to see that if $\bar{d}(x_0, t_0) = 0$ at some (x_0, t_0) in the interior of Ω, we have that $\bar{d}(x,t) = 0$ for all $(x,t) \in X \times [0, t_0]$. This then clearly implies the boundary strong maximum principle since $\bar{d}(x,t) = 0$ iff $(x,t) \in \partial K$.

In order for this proof to work, however, we need that ∂K be C^2. Otherwise, we cannot know that the partial derivatives of $\bar{d}$ will exist and equation (4.2) will have no meaning in the classical sense.

We now show that a proof similar to this can be used for arbitrary convex sets K, by showing that in general, $\bar{d}$ is a *viscosity super solution* to the differential inequality (4.2). We will be done once we show this since an analogous strong maximum principle for viscosity solutions is known (see proof in chapter 6).

We argue that $\bar{d}(x,t)$ is a viscosity super solution to (4.2) as follows. First, we note that as K is a convex set,

$$d(z) = \inf_{\ell} \ell(z)$$

where ℓ ranges over all supporting affine functionals of K. We say that an affine function $\ell(z)$ is a *supporting affine functional* of K if $\ell(z)$ is the distance from z to a supporting hyperplane of K, which by abuse of notation we also call ℓ.

Next, we note that for each z, this infimum is actually achieved. That is, for each $z \in K$, there is at least one point $y(z) \in \partial K$ and supporting affine functional ℓ such that

$$|z - y(z)| = \ell(z) = d(z)$$

We display this graphically in Figure 4-1.

So in particular, letting $\bar{\ell}(x,t) = \ell(\mathbf{u}(x,t))$, we have that

$$\bar{d}(x,t) = d(\mathbf{u}(x,t)) = \inf_{\ell} \ell(\mathbf{u}(x,t)) = \inf_{\ell} \bar{\ell}(x,t)$$

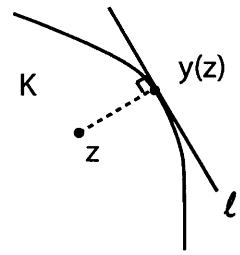

Figure 4-1: Picture of the ℓ which achieves the infimum.

and for each (x,t), this inf is achieved for some ℓ. We say that a supporting affine functional $\ell(z)$ is *active at* (x,t) if ℓ in fact achieves this infimum at (x,t). That is,

$$\ell(\mathbf{u}(x,t)) = \inf_{\ell} \ell(\mathbf{u}(x,t))$$

Next, we examine what differential equation $\bar{\ell}$ satisfies. Since $\ell(z)$ is affine, we have that $\nabla_z \ell(z) \equiv \nabla \ell$ is constant. We therefore compute:

$$\begin{aligned} \partial_t \bar{\ell}(x,t) &= \partial_t[\ell(\mathbf{u}(x,t))] = \nabla \ell \cdot \mathbf{u}_t(x,t) = \nabla \ell \cdot [\Delta \mathbf{u}(x,t) + \phi(\mathbf{u}(x,t))] \\ &= \Delta \ell(\mathbf{u}(x,t)) + \nabla \ell \cdot \phi(\mathbf{u}(x,t)) \end{aligned}$$

So for each supporting affine functional $\ell(z)$, we have that

$$\bar{\ell}_t(x,t) = \Delta_x \bar{\ell}(x,t) + \nabla \ell \cdot \phi(\mathbf{u}(x,t))$$

Furthermore, at points (x,t) where ℓ is active, by our inward pointing condition we have that $\nabla \ell \cdot \phi(y(\mathbf{u}(x,t))) \geq 0$, and so at these points we have that

$$\begin{aligned} \bar{\ell}_t(x,t) &= \Delta_x \bar{\ell}(x,t) + \nabla \ell \cdot \phi(\mathbf{u}(x,t)) \\ &\geq \Delta_x \bar{\ell}(x,t) + \nabla \ell \cdot \phi(\mathbf{u}(x,t)) - \nabla \ell \cdot \phi(y(\mathbf{u}(x,t))) \\ &\geq \Delta_x \bar{\ell}(x,t) - C|\mathbf{u}(x,t) - y(\mathbf{u}(x,t))| \\ &\geq \Delta_x \bar{\ell}(x,t) - C\bar{\ell}(x,t) \end{aligned}$$

where C is the Lipschitz constant for ϕ. Here we have used the fact that since ℓ is active, $\ell(\mathbf{u}(x,t)) = d(\mathbf{u}(x,t)) = |\mathbf{u}(x,t) - y(\mathbf{u}(x,t))|$.

So we have that

$$\bar{d}(x,t) = \inf_{\ell} \bar{\ell}(x,t)$$

and that

$$\partial_t \bar{\ell}(\hat{x},\hat{t}) \geq \Delta_x \bar{\ell}(\hat{x},\hat{t}) - C\bar{\ell}(\hat{x},\hat{t}) \tag{4.3}$$

for each pair of a point $(\hat{x},\hat{t})$ and supporting affine functional ℓ active at $(\hat{x},\hat{t})$. It therefore follows that

$$\partial_t \bar{d}(x,t) \geq \Delta_x \bar{d}(x,t) - C\bar{d}(x,t)$$

in the viscosity sense, because if $\psi(x,t)$ is any smooth function touching $\bar{d}(x,t)$ from below at the point $(\hat{x},\hat{t})$, we have that

$$\psi(\hat{x},\hat{t}) = \bar{d}(\hat{x},\hat{t}), \text{ and } \psi(x,t) \leq \bar{d}(x,t) \text{ in a neighborhood of } (\hat{x},\hat{t})$$

Taking ℓ to be the supporting affine functional active at $(\hat{x},\hat{t})$, we then have that

$$\psi(\hat{x},\hat{t}) = \bar{\ell}(\hat{x},\hat{t}), \text{ and } \psi(x,t) \leq \bar{\ell}(x,t) \text{ in a neighborhood of } (\hat{x},\hat{t})$$

Thus, $\psi(x,t)$ also touches the $C^{2,1}$ function $\bar{\ell}(x,t)$ from below at $(\hat{x},\hat{t})$, and so

$$\partial_t \psi = \partial_t \bar{\ell}(\hat{x},\hat{t}), \text{ and } \Delta_x \psi \leq \Delta_x \bar{\ell}(\hat{x},\hat{t})$$

from which, along with (4.3), it follows that

$$\partial_t \psi(\hat{x},\hat{t}) \geq \Delta_x \psi(\hat{x},\hat{t}) - C\psi(\hat{x},\hat{t})$$

Since this is true for all points $(\hat{x},\hat{t})$ and all smooth functions ψ touching $\bar{d}$ from below at $(\hat{x},\hat{t})$, we have shown that $\bar{d}(x,t)$ is a viscosity super solution to (4.2).

Chapter 5

Main Proof

This section contains the contents of my paper [8], and is devoted to proving Theorem 3.4.2. To be precise, let $X \subseteq \mathbb{R}^n$ be open and connected, and set $\Omega = X \times (0, \infty)$. Suppose $\mathbf{u} \in C^{2,1}(\Omega; \mathbb{R}^k) \cap C(\bar{\Omega}; \mathbb{R}^k)$ satisfies a parabolic system of equations of the form

$$\frac{\partial \mathbf{u}}{\partial t} = D(x,t,\mathbf{u}) \sum_{i,j} a_{ij}(x,t) \frac{\partial^2 \mathbf{u}}{\partial x_i \partial x_j} + \sum_i M_i(x,t,\mathbf{u}) \frac{\partial \mathbf{u}}{\partial x_i} + \phi(x,t,\mathbf{u}). \tag{5.1}$$

where the a_{ij} are real-valued, ϕ takes values in $\mathbb{R}^k$, and $D(x,t,z)$ and each of the $M_i(x,t,z)$ take values in the space of $k \times k$ matrices.

We make the following regularity assumptions: As a function of z, we assume that $\phi(x,t,z)$, $D(x,t,z)$ and each of the $M_i(x,t,z)$ are Lipschitz continuous uniformly for (x,t) in compact subsets of Ω. We assume also that each of the $a_{ij}(x,t)$ is locally bounded in Ω and that $D(x,t,z)$ and each of the $M_i(x,t,z)$ are locally bounded in $\Omega \times \mathbb{R}^k$. Finally, we assume that the matrix $\{a_{ij}(x,t)\}$ is symmetric and locally uniformly positive definite in Ω and the matrix $D(x,t,z)$ is locally uniformly positive definite in $\Omega \times \mathbb{R}^k$.

Next, suppose that K is a closed, convex subset of $\mathbb{R}^k$ which is *compatible* with ϕ, D, and the M_i in the following sense: For all $(x,t) \in \Omega$, $v \in \partial K$ and each vector ν which is inward pointing at v (See the next subsection for the definition of "inward pointing at v."), we have that $\phi(x,t,v) \cdot \nu \geq 0$ and that ν is a left eigenvector of

$D(x,t,v)$ and each of the $M_i(x,t,v)$.

We will show that if

$$\mathbf{u}(x,t) \in K \text{ for all } (x,t) \in \Omega, \tag{5.2}$$

then

$$\mathbf{u}(x_0,t_0) \in \partial K \text{ for some } (x_0,t_0) \in \Omega \implies \mathbf{u}(x,t) \in \partial K \text{ for all } (x,t) \in X \times [0,t_0], \tag{5.3}$$

and we will refer to the implication in (5.3) as **the strong maximum principle** with respect to K for solutions to (5.1). Under circumstances when solutions to (5.1) satisfy a **weak maximum principle** with respect to K of the form

$$\mathbf{u}(x,t) \in K \text{ for } (x,t) \in \partial\Omega \implies \mathbf{u}(x,t) \in K \text{ for } (x,t) \in \Omega, \tag{5.4}$$

it is obvious that (5.2) can be replaced by

$$\mathbf{u}(x,t) \in K \text{ for } (x,t) \in \partial\Omega.$$

5.1 Proof of (5.3)

We begin with the following definitions:

Definition 5.1.1. Given a convex set K and a boundary point $v \in \partial K$, a function $\ell : K \to \mathbb{R}$ is called a *supporting affine functional of K at v* if $\ell(z) \geq 0$ for all $z \in K$, $\ell(v) = 0$, and $|\nabla \ell(z)| \equiv |\nabla \ell| = 1$ (That is, $\ell(z)$ is the distance to a supporting hyperplane of K at v, which, by abusing notation, we also denote by ℓ).

We say that a function $\ell : K \to \mathbb{R}$ is a *supporting affine functional* of K if it is a supporting affine functional of K at some $v \in \partial K$.

Definition 5.1.2. For $v \in \partial K$, a vector ν is an *inward pointing vector* at $v \in \partial K$ if there exists a supporting affine functional $\ell(z)$ at v such that $\nabla \ell = \nu$.

Note that in our definitions we include the assumption that an inward pointing vector has unit length and that a supporting affine functional has unit-length gradient. For geometric reasons, these are nice assumptions to make as they allow for the intuitively appealing interpretation of $\ell(z)$ as the distance to a supporting hyperplane, which, by abusing notation, we also denote by ℓ.

Given $v \in \partial K$, we use L_v to denote the set of all the supporting affine functionals ℓ of K at v. By well known results (e.g., the Hahn-Banach Theorem), $L_v \neq \emptyset$ for every $v \in \partial K$. For $z \in K$ we let $d(z) = \inf\{|z - v| : v \in \partial K\}$ denote the distance from z to the boundary of K.

First we present a lemma from convex analysis.

Lemma 5.1.3. *For any convex set K,*

$$d(z) = \inf_{\ell} \ell(z),$$

where the infimum is taken over all supporting affine functionals ℓ. Furthermore, for each $z \in K$, there is a (possibly non-unique) point $v \in \partial K$ and an $\ell \in L_v$ such that

$$d(z) = |z - v| = \ell(z).$$

In fact, L_v consists of just this one supporting affine functional ℓ.

Notice that there are two points, v_1 and v_2, on the boundary which are closest to z and that neither is at a corner of ∂K (where L_v has more than one element).

Proof. It is easy to see geometrically that $d(z) \leq \ell(z)$ for each ℓ: Given ℓ, $\ell(z) = |z - \bar{z}|$, where $\bar{z}$ is the projection of z onto the supporting hyperplane determined by ℓ. The line from z to $\bar{z}$ intersects ∂K at some point w. Thus

$$d(z) \leq |z - w| \leq |z - \bar{z}| = \ell(z).$$

Next we show that the infimum over ℓ is in fact achieved. Using $B(z, r)$ to denote the closed ball with radius r centered at z, we have that $d(z) = \inf\{|z - w| : w \in$

$\partial K\} = \inf\{|z-w| : w \in \partial K \cap B(z, 2d(z))\}$. By compactness, this infimum is achieved and we have $d(z) = |z - v|$ for some $v \in \partial K$.

Take $\ell \in L_v$. We have shown that $d(z) \leq \ell(z)$. Since $v \in \ell$, we also have that $\ell(z) = \mathrm{dist}(z, \ell) \leq |z - v| = d(z)$. And so it follows that $d(z) = \ell(z)$.

So for all $\ell \in L_v$ we have that $\ell(z) = |z - v|$ which means that the line from z to v is normal to the hyperplane ℓ. As there is only one hyperplane with this normal which touches v, we have that L_v is just the singleton $\{\ell\}$.

□

In the remainder of the paper we will use the notation $\bar{d}(x,t)$ for $d(\mathbf{u}(x,t))$ and $\bar{\ell}(x,t)$ for $\ell(\mathbf{u}(x,t))$. The following is the key result proved in this section.

Theorem 5.1.4. *$\bar{d}(x,t)$ satisfies, in the viscosity sense, a parabolic equation of the form*

$$\frac{\partial \bar{d}}{\partial t} - \sum_{i,j} \alpha_{ij}(x,t) \frac{\partial^2 \bar{d}}{\partial x_i \partial x_j} - \sum_i \beta_i(x,t) \frac{\partial \bar{d}}{\partial x_i} + \gamma(x,t)\bar{d} \geq 0,$$

where $\gamma(x,t) \geq 0$ in Ω, each of the α_{ij}, β_i, and γ are locally bounded in Ω, and $\{\alpha_{ij}\}$ is locally uniformly positive definite in Ω.

Proof. Call a quadruple (x, t, v, ℓ) *nice* if

1. $(x,t) \in \Omega$, $v \in \partial K$, and ℓ is a supporting affine functional of K at v.

2. $d(\mathbf{u}(x,t)) = |\mathbf{u}(x,t) - v|$.

By the previous lemma, we know that for each (x,t) we can find at least one $v \in \partial K$ satisfying the second condition and this in turn will determine a unique choice of ℓ satisfying the first condition. We then also have that $\ell(\mathbf{u}(x,t)) = |\mathbf{u}(x,t) - v|$.

For any nice quadruple we have that

$$
\begin{aligned}
\bar{\ell}_t &= \nabla\ell \cdot \mathbf{u}_t \\
&= \nabla\ell \cdot \Big[D(x,t,\mathbf{u})\sum_{i,j} a_{ij}(x,t)\mathbf{u}_{x_ix_j} + \sum_i M_i(x,t,\mathbf{u})\mathbf{u}_{x_i} + \phi(x,t,\mathbf{u})\Big] \\
&= \nabla\ell \cdot \Big[D(x,t,\mathbf{u})\sum_{i,j} a_{ij}(x,t)\mathbf{u}_{x_ix_j} + \sum_i M_i(x,t,\mathbf{u})\mathbf{u}_{x_i} + \phi(x,t,\mathbf{u})\Big] \\
&\quad -\nabla\ell \cdot \Big[D(x,t,v)\sum_{i,j} a_{ij}(x,t)\mathbf{u}_{x_ix_j} + \sum_i M_i(x,t,v)\mathbf{u}_{x_i} + \phi(x,t,v)\Big] \\
&\quad +\nabla\ell \cdot \Big[D(x,t,v)\sum_{i,j} a_{ij}(x,t)\mathbf{u}_{x_ix_j} + \sum_i M_i(x,t,v)\mathbf{u}_{x_i} + \phi(x,t,v)\Big] \\
&\geq -\left[c(x,t)\Big\|\sum_{i,j} a_{ij}(x,t)\mathbf{u}_{x_ix_j}\Big\| + \sum_i m_i(x,t)\left\|\mathbf{u}_{x_i}\right\| + p(x,t)\right] |\mathbf{u}-v| \\
&\quad +\mu(x,t,v,\nabla\ell)\sum_{i,j} a_{ij}(x,t)\left(\nabla\ell\cdot\mathbf{u}_{x_ix_j}\right) + \sum_i \lambda_i(x,t,v,\nabla\ell)\left(\nabla\ell\cdot\mathbf{u}_{x_i}\right) + 0 \\
&= -\left[c(x,t)\Big\|\sum_{i,j} a_{ij}(x,t)\mathbf{u}_{x_ix_j}\Big\| + \sum_i m_i(x,t)\left\|\mathbf{u}_{x_i}\right\| + p(x,t)\right] \bar{\ell} \\
&\quad +\mu(x,t,v,\nabla\ell)\sum_{i,j} a_{ij}(x,t)\bar{\ell}_{x_ix_j} + \sum_i \lambda_i(x,t,v,\nabla\ell)\bar{\ell}_{x_i},
\end{aligned}
$$

where $c(x,t)$, $m_i(x,t)$, and $p(x,t)$ are the Lipschitz constants for $D(x,t,\cdot)$, $M_i(x,t,\cdot)$, and $\phi(x,t,\cdot)$, respectively, and μ and λ_i are the eigenvalues for D and M_i at v corresponding to the left eigenvector $\nabla\ell$ (recall that $\nabla\ell$ is inward pointing at v). Note that in the inequality step we have used the fact that $\phi(x,t,v)\cdot\nabla\ell \geq 0$ which follows from the fact that $\nabla\ell$ is inward pointing at v and from our compatibility assumption on ϕ and K.

Next, let $\gamma(x,t) = \left[c(x,t)\|\sum_{i,j} a_{ij}(x,t)\mathbf{u}_{x_ix_j}\| + \sum_i m_i(x,t)\left\|\mathbf{u}_{x_i}\right\| + p(x,t)\right]$. It follows from our regularity assumptions on D, M_i, and ϕ that $c(x,t)$, $m_i(x,t)$, and $p(x,t)$ are locally bounded. Since $\mathbf{u} \in C^{2,1}(\Omega)$, its first and second spatial derivatives are locally bounded, and so, since the $a_{ij}(x,t)$ are locally bounded, it follows that $\gamma(x,t)$ is locally bounded. As each of the Lipschitz constants is positive, it is clear that $\gamma(x,t) \geq 0$.

Thus we have shown that for each nice quadruple (x, t, v, ℓ),

$$\bar{\ell}_t - \mu(x, t, v, \nabla \ell) \sum_{i,j} a_{ij}(x, t) \bar{\ell}_{x_i x_j} - \sum_i \lambda_i(x, t, v, \nabla \ell) \bar{\ell}_{x_i} + \gamma(x, t) \bar{\ell} \geq 0. \qquad (5.5)$$

As remarked at the outset of this proof, given $(x, t) \in \Omega$ we can always find a $v \in \partial K$ and ℓ to form a nice quadruple (x, t, v, ℓ). In general there will be more than one way to extend (x, t) into a nice quadruple, but for our purposes we only care that an extension is possible.

To avoid the axiom of choice, we now describe a method of choosing an extension: Given $(x, t) \in \Omega$ we need that $v \in \partial K$ be such that $|\mathbf{u}(x, t) - v| = d(\mathbf{u}(x, t))$. The set of v which satisfy this relation is a closed and bounded and hence compact subset of $\mathbb{R}^k$. We first look at v in this set with smallest first component. If there is a unique such v we choose it. Otherwise, among those v we look for the one with smallest second component and we continue this algorithm until we find a unique v. We denote this v by $v_{(x,t)}$ to make clear its dependence on (x, t). Lastly, once we have v, ℓ is uniquely determined (see Lemma 5.1.3) and we denote it by $\ell_{(x,t)}$.

Next we let $\tilde{\mu}(x, t) = \mu(x, t, v_{(x,t)}, \nabla \ell_{(x,t)})$ and $\tilde{\lambda}_i(x, t) = \lambda_i(x, t, v_{(x,t)}, \nabla \ell_{(x,t)})$. We claim that $\tilde{\mu}(x, t)$ and the $\tilde{\lambda}_i(x, t)$ are locally bounded. This is true as for any compact set $C \subset \Omega$,

$$\begin{aligned} \sup_{(x,t) \in C} \tilde{\mu}(x, t) &\leq \sup_{(x,t) \in C} \sup_{\{v : |\mathbf{u} - v| = d(\mathbf{u})\}} \mu(x, t, v, \nabla \ell) \\ &\leq \sup_{(x,t) \in C} \sup_{\{v : |\mathbf{u} - v| = d(\mathbf{u})\}} \|D(x, t, v)\| < \infty, \end{aligned}$$

where $\|D\|$ denotes the operator norm of D which is locally bounded by our regularity assumption on D. The same argument works for the $\tilde{\lambda}_i$.

We next claim that $\tilde{\mu}(x, t)$ is uniformly bounded away from 0 on compact sets. This is true as for any compact set $C \subset \Omega$,

$$\begin{aligned} \inf_{(x,t) \in C} \tilde{\mu}(x, t) &\geq \inf_{(x,t) \in C} \inf_{\{v : |\mathbf{u} - v| = d(\mathbf{u})\}} \mu(x, t, v, \nabla \ell) \\ &\geq \inf_{(x,t) \in C} \inf_{\{v : |\mathbf{u} - v| = d(\mathbf{u})\}} \Lambda_1(D(x, t, v)) > 0, \end{aligned}$$

where we denote by $\Lambda_1(D)$ the smallest eigenvalue of the positive definite matrix D. Here we have used our assumption that $D(x,t,v)$ is uniformly positive definite on compact sets.

Finally we reach the crux of our argument. We claim that $\bar{d}$ solves

$$\bar{d}_t - \tilde{\mu}(x,t)\sum_{i,j} a_{ij}(x,t)\bar{d}_{x_i x_j} - \sum_i \tilde{\lambda}_i(x,t)\bar{d}_{x_i} + \gamma(x,t) \geq 0$$

in the viscosity sense.

Suppose that $\psi \in C^\infty(\Omega)$ touches $\bar{d}$ from below at the point $(\hat{x},\hat{t})$, i.e.

$$\psi(\hat{x},\hat{t}) = \bar{d}(\hat{x},\hat{t}), \text{ and } \psi(x,t) \leq \bar{d}(x,t) \text{ in a neighborhood of } (\hat{x},\hat{t}).$$

We extend the point $(\hat{x},\hat{t})$ to the nice quadruple $(\hat{x},\hat{t},v_{(\hat{x},\hat{t})},\ell_{(\hat{x},\hat{t})})$. Since by Lemma 5.1.3, $d(z) \leq \ell_{(\hat{x},\hat{t})}(z)$, we then have that

$$\psi(\hat{x},\hat{t}) = \bar{\ell}_{(\hat{x},\hat{t})}(\hat{x},\hat{t}), \text{ and } \psi(x,t) \leq \bar{\ell}_{(\hat{x},\hat{t})}(x,t) \text{ in a neighborhood of } (\hat{x},\hat{t}).$$

Therefore, ψ also touches the function $\bar{\ell}_{(\hat{x},\hat{t})}$ from below at $(\hat{x},\hat{t})$, and so

$$\frac{\partial}{\partial t}\psi = \frac{\partial}{\partial t}\bar{\ell}_{(\hat{x},\hat{t})}(\hat{x},\hat{t}), \text{ and } \Delta_x\psi \leq \Delta_x\bar{\ell}_{(\hat{x},\hat{t})}(\hat{x},\hat{t}).$$

It then follows from (5.5) (and recalling the definitions of $\tilde{\mu}$ and $\tilde{\lambda}_i$) that

$$\psi_t - \tilde{\mu}(x,t)\sum_{i,j} a_{ij}(x,t)\psi_{x_i x_j} - \sum_i \tilde{\lambda}_i(x,t)\psi_{x_i} + \gamma(x,t)\psi \geq 0$$

Since this is true for all points $(\hat{x},\hat{t})$ and all smooth functions ψ touching $\bar{d}$ from below at $(\hat{x},\hat{t})$, we have shown that $\bar{d}(x,t)$ solves

$$\bar{d}_t - \tilde{\mu}(x,t)\sum_{i,j} a_{ij}(x,t)\bar{d}_{x_i x_j} - \sum_i \tilde{\lambda}_i(x,t)\bar{d}_{x_i} + \gamma(x,t)\bar{d} \geq 0$$

in the viscosity sense.

The theorem is now proved by letting $\beta_i = \tilde{\lambda}_i$ and $\alpha_{ij} = \tilde{\mu} a_{ij}$. Note that $\{\alpha_{ij}\}$ is uniformly positive definite on compact sets as $\{a_{ij}\}$ is uniformly positive definite on compact sets and $\tilde{\mu}$ is uniformly bounded away from 0 on compact sets.

□

In order to complete the proof of (5.3), we need the following version of the strong maximum principle for supersolutions in the viscosity sense:

Theorem 5.1.5. *Suppose that $f \in C\big(\bar{\Omega}; [0,\infty)\big)$ satisfies*

$$\frac{\partial f}{\partial t} - \sum_{i,j} \alpha_{ij}(x,t) \frac{\partial^2 f}{\partial x_i \partial x_j} - \sum_i \beta_i(x,t) \frac{\partial f}{\partial x_i} + \gamma(x,t) f \geq 0 \tag{5.6}$$

in the viscosity sense. Suppose further that in Ω, $\gamma \geq 0$, each of the α_{ij}, β_i, and γ are locally bounded, and $\{\alpha_{ij}\}$ is locally uniformly positive definite. Then

$$f(x_0, t_0) = 0 \text{ for some } (x_0, t_0) \in \Omega \implies f(x,t) = 0 \text{ for all } (x,t) \in \bar{\Omega} \text{ with } t \leq t_0.$$

Note that by applying this theorem to $f = \bar{d} \geq 0$ and using Theorem 5.1.4, we get (5.3) as an immediate consequence.

Theorem 5.1.5 follows from an extension of the proof of the classical strong maximum principle given by Nirenberg in [11] to the case of viscosity solutions given by F. Da Lio in [6]. Da Lio's proof handles more general PDE but has the drawback that the PDE must have continuous coefficients. We do not make any continuity assumptions on the coefficients of our PDE (5.6), and so our PDE does not directly fall under the assumptions of Da Lio's result. Nevertheless, as our PDE (5.6) is of such a simple form, Da Lio's method of extending Nirenberg's proof can still be applied to prove Theorem 5.1.5. We present this proof in the next chapter.

Chapter 6

Proof of the Strong Maximum Principle for Viscosity Solutions

In this section we prove the following strong maximum principle for viscosity solutions:

Theorem 6.0.6. *Suppose that $f \in C\big(\bar{\Omega}; [0,\infty)\big)$ satisfies*

$$F(f) := \frac{\partial f}{\partial t} - \sum_{i,j} \alpha_{ij}(x,t)\frac{\partial^2 f}{\partial x_i \partial x_j} - \sum_i \beta_i(x,t)\frac{\partial f}{\partial x_i} + \gamma(x,t) f \geq 0$$

in the viscosity sense, where, in Ω, $\gamma \geq 0$, each of the α_{ij}, β_i, and γ are locally bounded, and $\{\alpha_{ij}\}$ is locally uniformly positive definite. Then

$$f(x_0,t_0) = 0 \text{ for some } (x_0,t_0) \in \Omega \implies f(x,t) = 0 \text{ for all } (x,t) \in \bar{\Omega} \text{ with } t \leq t_0.$$

Our proof reproduces a specific case of a more general result of F. Da Lio [6] , whose proof follows the classic argument of L. Nirenberg [11]. We first prove the following claim:

Claim 6.0.7. *If $f(x_0,t_0) = 0$ for some $(x_0,t_0) \in \mathrm{int}\Omega$ then $f(x,t_0) = 0$ for all $x \in X$. That is, $f(x,t)$ is identically 0 on the horizontal strip $X \times \{t = t_0\}$.*

Proof 6.0.8. Suppose this is not the case. Then letting $P_0 = (x_0,t_0)$, we have a point $P_1 = (x_1,t_1)$ such that $t_1 = t_0$ and $f(P_1) > 0$. By the continuity of f, we can find a pair of points $P_2 = (x_2, t_2), P_3 = (x_3, t_3)$ in $X \times \{t = t_0\}$ such that $f(P_2) = 0$, $f(P_3) > 0$, and $\text{dist}(P_2, P_3) < \text{dist}(P_3, \partial\Omega)$ (If P_0 and P_1 are sufficiently close we can just take those as P_2 and P_3. If not, we can find P_2 and P_3 on the line connecting P_0 and P_1). Our picture looks like Figure 6-1.

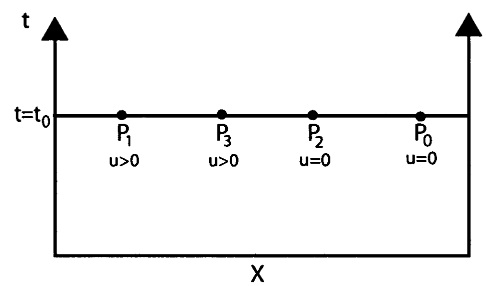

Figure 6-1:

Since $f(P_3) > 0$, by the continuity of f there exists a ball of radius $\varepsilon > 0$ centered at P_3 such that $f(x,t) > 0$ uniformly in this ball (i.e. $f(x,t) \geq C > 0$ for some C in this ball). Note that since $f(P_2) = 0$, it must be that $\varepsilon < \text{dist}(P_2, P_3) < \text{dist}(P_3, \partial\Omega)$ and so this ball is contained in Ω. Next, construct a family of ellipsoids $\mathcal{E}_\lambda$ which are centered at P_3, whose vertical axis has length $\frac{\varepsilon}{2}$, and whose horizontal axes each have length λ. Our picture now looks like Figure 6-2 (Note that in our figures we will drop elements as they are no longer needed):

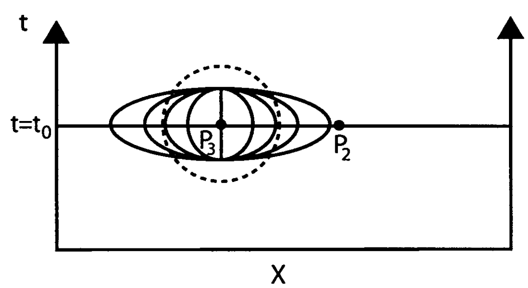

Figure 6-2:

Starting with $\lambda = 0$ and letting λ increase, by the continuity of f we eventually find an ellipsoid $\mathcal{E}_\lambda$ such that

$$f(x,t) > 0 \text{ for all } (x,t) \in \text{int}\mathcal{E}_\lambda \text{ and } f(x^*, t^*) = 0 \text{ for some point } (x^*, t^*) \in \partial\mathcal{E}_\lambda$$

Note that this will happen before the ellipsoid extends past the boundary of Ω as $f(P_2) = 0$ and P_2 is closer to P_3 than the boundary of Ω. Let $P^* = (x^*, t^*)$, and note that $x^* \neq x_3$ (i.e. P^* is not at the top or bottom of the ellipsoid). Our picture now looks like Figure 6-3.

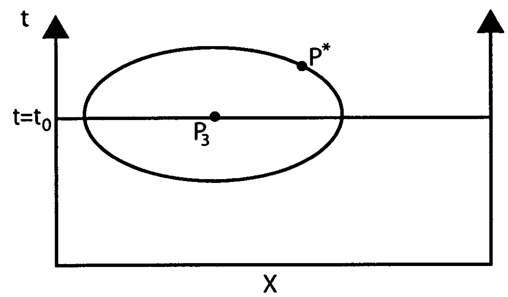

Figure 6-3:

Next, let B_1 be a closed ball of radius R centered at a point $(\tilde{x}, \tilde{t})$ such that B_1 lies inside $\mathcal{E}_\lambda$ and touches $\mathcal{E}_\lambda$ at P^*. Furthermore we take R small enough that $\tilde{x} \neq x^*$ which we can do since P^* is not at the top or bottom of the ellipsoid. Our picture now looks like Figure 6-4.

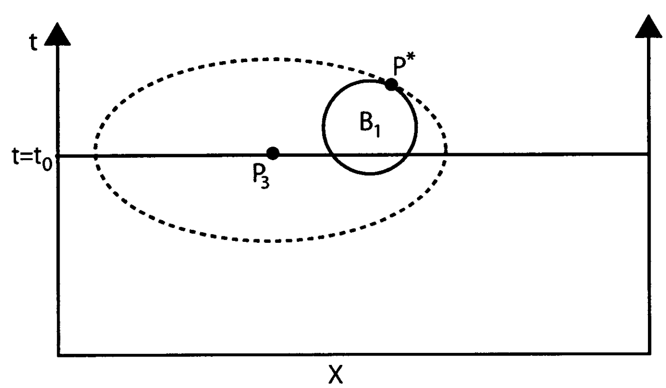

Figure 6-4:

Let $g(x,t) = -e^{-\eta R^2} + e^{-\eta|(x,t)-(\tilde{x},\tilde{t})|^2}$, where η is a constant we will tweak later, and note that

$$\begin{aligned} g(x,t) &> 0 \text{ for } (x,t) \in \operatorname{int} B_1 \\ g(x,t) &= 0 \text{ for } (x,t) \in \partial B_1 \\ g(x,t) &< 0 \text{ for } (x,t) \notin B_1 \end{aligned}$$

Next, note that

$$
\begin{aligned}
F(g) &= \frac{\partial g}{\partial t} - \sum_{i,j} \alpha_{ij}(x,t)\frac{\partial^2 g}{\partial x_i \partial x_j} - \sum_i \beta_i(x,t)\frac{\partial g}{\partial x_i} + \gamma(x,t)g \\
&= -2\eta(t-\tilde{t})e^{-\eta|(x,t)-(\tilde{x},\tilde{t})|^2} - 4\eta^2 \sum_{i,j} \alpha_{ij}(x,t)(x_i-\tilde{x}_i)(x_j-\tilde{x}_j)e^{-\eta|(x,t)-(\tilde{x},\tilde{t})|^2} \\
&\quad +2\eta \sum_i \alpha_{ii}(x,t)e^{-\eta|(x,t)-(\tilde{x},\tilde{t})|^2} + 2\eta \sum_i \beta_i(x,t)(x_i-\tilde{x}_i)e^{-\eta|(x,t)-(\tilde{x},\tilde{t})|^2} \\
&\quad +\gamma(x,t)[-e^{-\eta R^2} + e^{-\eta|(x,t)-(\tilde{x},\tilde{t})|^2}]
\end{aligned}
$$

By our assumptions on α_{ij}, β, and γ we can find a ball B_2 centered at P^* such that after tweaking η, $F(g) < 0$ uniformly on B_2: Note that since $\gamma(x,t) \geq 0$, we have that

$$
\begin{aligned}
F(g) \;\leq\; & 2\eta e^{-\eta|(x,t)-(\tilde{x},\tilde{t})|^2}[-(t-\tilde{t}) - 2\eta \sum_{i,j} \alpha_{ij}(x,t)(x_i-\tilde{x}_i)(x_j-\tilde{x}_j) \\
& + \sum_i \alpha_{ii}(x,t) + \sum_i \beta_i(x,t)(x_i-\tilde{x}_i) + \frac{1}{2\eta}\gamma(x,t)]
\end{aligned}
$$

Choose a B_2 centered at P^* sufficiently small that $B_2 \subset \Omega$ and $(\tilde{x},\tilde{y}) \notin B_2$. Then by our uniform positive definiteness assumption on $\{\alpha_{ij}\}$,

$$
\sum_{i,j} \alpha_{ij}(x,t)(x_i-\tilde{x}_i)(x_j-\tilde{x}_j) \geq K_1|x-\tilde{x}|^2 \geq K_2 > 0
$$

in B_2 for some constants K_1, K_2. By our boundedness assumption on the coefficients, all of the other terms above are bounded on B_2, and so after choosing a sufficiently large η, $F(g) < 0$ uniformly on B_2.

Let $U = B_1 \cap B_2$. Our picture now looks like Figure 6-5.

We claim that on ∂U, $f(x,t) \geq \varepsilon g(x,t)$ for some $\varepsilon > 0$. On one part of ∂U (the upper arc in our picture), this is trivially true as $g = 0$ along this part of the boundary, and we know that $f \geq 0$ everywhere. On the other part of the boundary

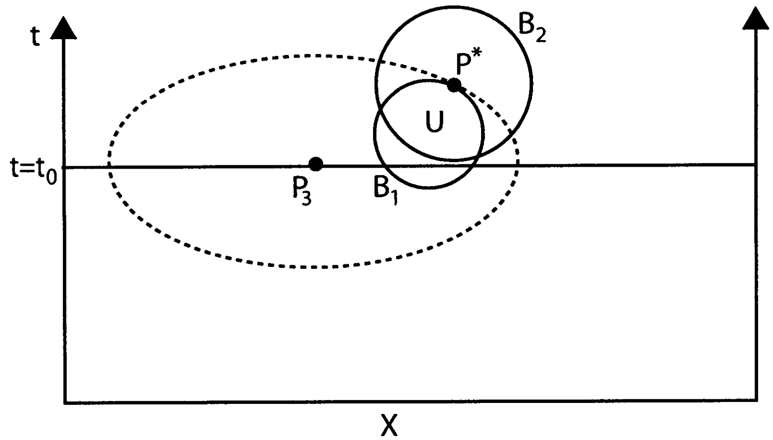

Figure 6-5:

(the lower part in our picture, which we take to include its boundary so that it is closed and hence compact) we have that $f > 0$ and hence $f > 0$ uniformly on this part of the boundary. Since $g \leq 1$ we can choose an ε sufficiently small that $f \geq \varepsilon g$ on this part of, and hence on all of ∂U.

We next claim that in fact $f \geq \varepsilon g$ in all of U: Suppose not. Let $(\hat{x}, \hat{t})$ be the point in U such that

$$f(\hat{x}, \hat{t}) - \varepsilon g(\hat{x}, \hat{t}) = \inf_{(x,t) \in U} \{f(x,t) - \varepsilon g(x,t)\} < 0$$

Then, as g is smooth, and $F(f) \geq 0$ in the viscosity sense, it follows from the definition of viscosity supersolution that

$$F(g)(\hat{x}, \hat{t}) = \frac{1}{\varepsilon} F(\varepsilon g)(\hat{x}, \hat{t}) \geq 0$$

Since $U \subset B_2$ and $F(g) < 0$ in B_2, this yields a contradiction.

As $g < 0$ outside of B_1, it follows that $f \geq \varepsilon g$ on all of B_2. Since we have that

$$\begin{aligned} f(x,t) &\geq \varepsilon g(x,t) \text{ in } B_2 \\ f(x^*, t^*) &= 0 = \varepsilon g(x^*, t^*) \end{aligned}$$

we see that εg touches f from below at (x^*, t^*) and since f is a viscosity supersolution, this implies that

$$F(g)(x^*, t^*) = \frac{1}{\varepsilon} F(\varepsilon g)(x^*, t^*) \geq 0$$

But this contradicts our earlier calculation that $F(g) < 0$ in B_2.

We have so far showed that if $f(x_0, t_0) = 0$ for some $(x_0, t_0) \in \text{int}\Omega$ then $f \equiv 0$ on the horizontal strip $X \times \{t = t_0\}$. We next want to show that in fact $f(x,t) \equiv 0$ for $t < t_0$ as well. To this end, we will need the following lemma:

Lemma 6.0.9. *Suppose $P^* \in int\Omega$, and let R be a rectangular box with P^* in the center of its top face. That is, R takes the form*

$$\{x_1^* - a_1, x_1^* + a_1\} \times \{x_2^* - a_2, x_2^* + a_2\} \times \cdots \times \{x_n^* - a_n, x_n^* + a_n\} \times \{t^* - a, t^*\}$$

for some positive constants $a, a_1, \ldots, a_n$. Then if $f(P^*) = 0$, there exists a point P in the interior of R such that $f(P) = 0$.

So far this seems far from what we want as this lemma only guarantees us that one point in every such rectangle will have $f = 0$. But in fact, once we know this, we can argue by geometry that in fact every point in the interior of R has $f = 0$ and then in turn argue that $f = 0$ on all of $X \times [0, t_0]$. We save these arguments for after the proof of this lemma. The proof of this lemma resembles the proof of the previous claim.

Proof 6.0.10. Suppose not. Then $f > 0$ everywhere in the interior of R. Let

$$h(x,t) = -(t - t^*) - K|x - x^*|^2$$

where $K > 0$ is a constant to be tweaked later. Note that the set $\{(x,t) | h(x,t) = 0\}$ is a paraboloid, and below it $h > 0$ and above it $h < 0$. Our picture looks like Figure 6-6.

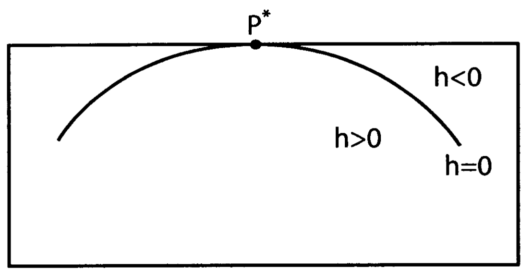

Figure 6-6:

We compute that

$$\begin{aligned} F(h) &= \frac{\partial h}{\partial t} - \sum_{i,j} \alpha_{ij}(x,t)\frac{\partial^2 h}{\partial x_i \partial x_j} - \sum_i \beta_i(x,t)\frac{\partial h}{\partial x_i} + \gamma(x,t)h \\ &= -1 - [-2K\sum_{i,j}\alpha_{ij}(x,t)\delta_{ij} - 2K\sum_i \beta_i(x,t)(x_i - x_i^*)] \\ &\quad +\gamma(x,t)[-(t-t^*) - K|x-x^*|^2] \end{aligned}$$

We can find a ball B centered at P^* such that after tweaking K, $F(h) < 0$ uniformly on B: As $\gamma(x,t) \geq 0$, we have that

$$\begin{aligned} F(h) &= -1 + 2K[\sum_i \alpha_{ii}(x,t) + \sum_i \beta_i(x,t)(x_i - x_i^*)] \\ &\quad +\gamma(x,t)(t^* - t) \end{aligned}$$

First take a ball $B_1 \subset \Omega$ centered at P^*. Then by our boundedness assumption on γ, $\gamma(x,t) \leq \Gamma$ on B_1 for some constant $\Gamma > 0$. Next take $B \subset B_1$ to be a smaller ball centered at P^* and small enough that $\gamma(x,t)(t^* - t) \leq \frac{1}{2}$. Then on B, $\sum_i \alpha_{ii}(x,t) + \sum_i \beta_i(x,t)(x_i - x_i^*) \leq C$ for some constant $C > 0$. And so

$$F(h) \leq -\frac{1}{2} + 2KC$$

and so after taking $K > 0$ to be sufficiently small, we have that $F(h) < 0$ uniformly on B.

Let $U = B \cap \{(x,t)|h(x,t) > 0\}$. Our picture now looks like Figure 6-7.

We claim that $f \geq \varepsilon h$ on ∂U for a sufficiently small $\varepsilon > 0$. On the upper part of

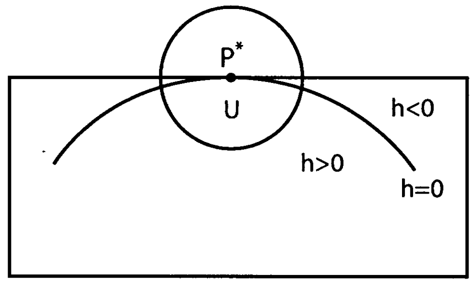

Figure 6-7:

∂U we have that $h = 0$ so this is trivially true. On the lower part (which we consider to be closed and hence compact) we have by assumption that $f > 0$ and hence $f > 0$ uniformly. Since h is bounded from above on R, we can find the desired $\varepsilon > 0$.

We next claim that in fact $f \geq \varepsilon h$ in all of U. This follows from the same argument made in the proof of the previous claim from the fact that $F(f) \geq 0$ in the viscosity sense, h is smooth, and $F(h) < 0$ in U

Finally, since $h < 0$ on $B \backslash U$ we have that $f \geq \varepsilon h$ on all of U. Since $f(x^*, t^*) = 0 = \varepsilon h(x^*, t^*)$ we see that εh touches f from below at (x^*, t^*) and so since f is a viscosity supersolution this implies that $F(h)(x^*, t^*) \geq 0$. But this contradicts the fact that $F(h) < 0$ in B.

We now argue that in fact, under the conditions of the previous lemma, we have that $f = 0$ for *all* of the points in the interior of R. Suppose this is not the case. Then there exists some point P_1 in the interior of R such that $u(P_1) > 0$. Traveling along the line from P_1 to P^*, there is a first point P_2 such that $f(P_2) = 0$ (it may be that $P_2 = P^*$). We can now draw a rectangle R' such that P_2 lies in the center of the top of R' and P_1 lies on the bottom of R' (see Figure 6-8).

By our previous lemma, there is some point Q in the interior of R such that $f(Q) = 0$. And so by our earlier claim we have that every point on the horizontal strip containing Q also has $f = 0$. In particular there is one such point on the line between P_1 and P_2. This yields a contradiction, however, as every point on the line between P_1 and P_2 satisfies $f > 0$ (see Figure 6-9).

It is now clear that if $f = 0$ at some point $P_0 = (x_0, t_0) \in \text{int}\Omega$, then $f \equiv 0$ on $X \times [0, t_0]$ since if this weren't the case, there would be a point S in $X \times [0, t_0]$ such

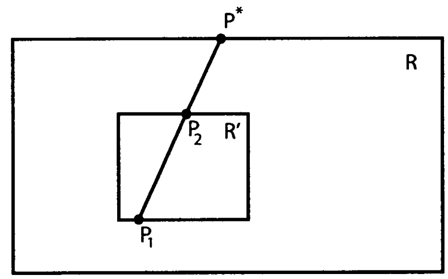

Figure 6-8:

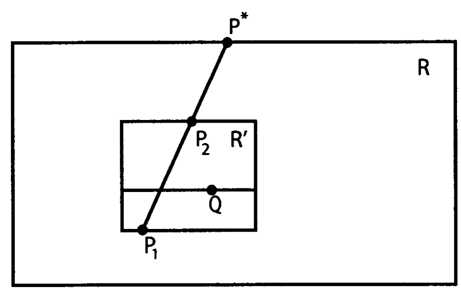

Figure 6-9:

that $f(S) > 0$. As X is connected, there exists a continuous path increasing in time from S to P_0. Traveling along this path, we hit a first point P^* where $f = 0$. Then if we draw a rectangular box with P^* in the center of its top, we have that $f = 0$ in the interior of this box. But this contradicts the fact that P^* is the first point where $f = 0$ on the line from S to P_0 (see Figure 6-10).

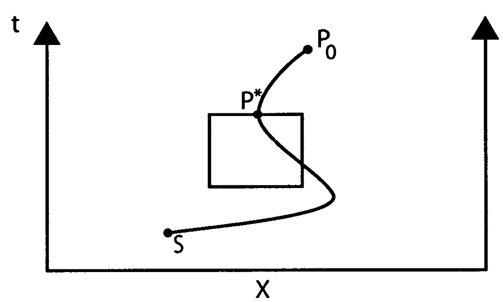

Figure 6-10:

Chapter 7

Second Result: Introduction

In this second half of the thesis we give an approximation scheme for solutions to reflected SDE of the Stratonovich type. We will show that the distribution of the solution to such a reflected SDE is the weak limit of the distribution of the solutions of the reflected SDEs one gets by replacing the driving Brownian motion by its N-dyadic linear interpolation. In the last chapter, we explore the implications of this and in particular show how we can infer geometric properties of the solutions to a Stratonovich reflected SDE from those of the solutions to the approximating reflected SDE. The results of this half of the thesis are from joint work with Professor Daniel W. Stroock.

After proving this result, we discovered the 2001 paper [9] by A. Kohatsu-Higa which contains a more general result. Specifically, Kohatsu-Higa shows that a more complex class of reflected SDEs is stable with respect to the driving process. In particular, taking the N-dyadic linear interpolation of Brownian motion as the driving process and passing to the limit, one gets our result. Moreover, one gets convergence pathwise almost surely.

Our proof is however substantially different from the one in [9]. As we only prove weak convergence for our approximations, we are able to give a proof which doesn't require as much background and relies mainly on familiar methods of proof (e.g. Kolmogorov's criterion and the martingale problem). We therefore believe our proof has merit as well.

In his 1999 paper [12], R. Petterson also gives a proof of our result, but his proof is limited to the case when the domain of reflection is convex. In our paper we consider more general domains.

In this chapter we first lay out the background for our result. We explain the Skorohod problem and give the definition of a reflected SDE. We then give a statement of our result. Finally, in the last section we explain the structure of this half of the thesis.

7.1 Skorohod Problem

We begin by introducing the (deterministic) Skorohod problem: Let $\mathcal{O} \subset \mathbb{R}^d$ be a domain and to each $x \in \partial\mathcal{O}$ assign a nonempty collection $\nu(x) \subset \mathbb{S}^{d-1}$ of of vectors in the $(d-1)$-dimensional unit sphere. Given a continuous function $w_t : [0,\infty) \to \mathbb{R}^d$ with $w_0 \in \mathcal{O}$, we say that a *solution to the Skorohod problem for* $(\mathcal{O}, \nu(x))$ is a continuous function $x_t : [0,\infty) \to \bar{\mathcal{O}}$ and a continuous function of locally bounded variation $L_t : [0,\infty) \to \mathbb{R}^d$ such that

$$\begin{aligned} x_t =& w_t + L_t, \\ |L|_t =& \int_0^t 1_{\{x_s \in \partial\mathcal{O}\}} d|L|_s, \\ \text{and } L_t =& \int_0^t \nu(x_s) d|L|_s, \end{aligned}$$

where $|L|_t$ denotes the total variation of L_t by time t, and the third line is a shorthand way of expressing

$$\frac{dL_t}{d|L|_t} \in \nu(x_t), \quad \forall t.$$

We will often speak of w_t as being the "input" and x_t as being the "output" of the Skorohod problem.

We present the following intuitive picture of the Skorohod problem: Think of the function w_t as being a set of instructions we feed to a robot telling it how to move, i.e. at time t, the robot should be at position w_t. Left to carry out its instructions, the robot happily traces the path described by the function w_t.

Suppose now that the boundary $\partial\mathcal{O}$ acts as a wall through which the robot cannot move and when the robot contacts the wall at a point $x \in \partial\mathcal{O}$ it is pushed back in one of the directions $\nu \in \nu(x)$. The robot, however, is oblivious to this and thinks it is still on its original course and continues to carry out its movement instructions. And so its final path, which we denote x_t, is a perturbation of its programmed path w_t (See Figure 7-1).

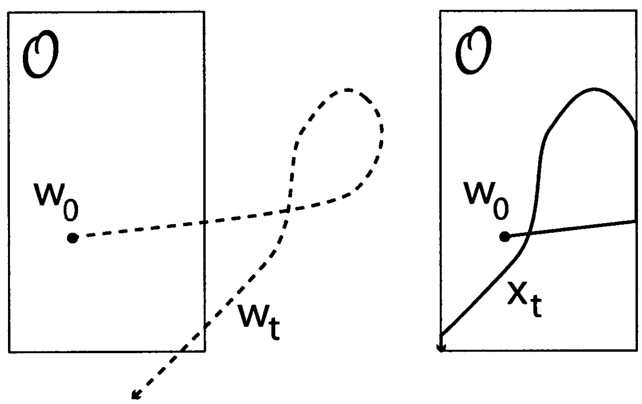

Figure 7-1:

If we let $L_t = x_t - w_t$, then L_t records the total perturbation up to time t of the robot from its intended path. It makes sense that L_t should only change when $x_t \in \partial\mathcal{O}$ and then L_t should only change in one of the directions $\nu \in \nu(x_t)$.

In this thesis we will consider a certain class of domain which was first considered by P. L. Lions and A. S. Sznitman in [10].

Definition 7.1.1. A set $\mathcal{O}$ is *admissible* if the following hold:

1. (a) There exists a $C_0 \geq 0$ such that for all $x' \in \mathcal{O}$, $x \in \partial\mathcal{O}$, and $\nu \in \nu(x)$,

$$(x' - x) \cdot \nu + C_0|x - x'|^2 \geq 0.$$

   (b) Furthermore, if for any point $x \in \partial\mathcal{O}$ there exists a non-zero vector $v \in \mathbb{R}^d$ and constant $C \geq 0$ such that

$$(x' - x) \cdot v + C|x - x'|^2 \geq 0, \quad \forall x' \in \mathcal{O},$$

then $\frac{v}{|v|} \in \nu(x)$.

2. There exists a function $\phi \in C^2(\mathbb{R}^d; \mathbb{R})$ such that

$$\nabla \phi(x) \cdot \nu \geq \alpha > 0$$

for all $x \in \partial \mathcal{O}$, $\nu \in \nu(x)$.

3. We have the following "circle covering" condition: There exist $n \geq 1$, $\lambda > 0$, $R > 0$, $a_1, \dots, a_n \in \mathbb{R}^d$ with $|a_i| = 1 \; \forall i$, and $x_1, \dots, x_n \in \partial O$ such that

$$\partial O \subset \bigcup_{i=1}^{n} B(x_i, R)$$

and

$$x \in \partial O \cap B(x_i, 2R) \Longrightarrow \nu(x) \cdot a_i \geq \lambda > 0.$$

When $\partial \mathcal{O}$ is smooth, Part 1 of Defn 7.1.1 says that for each $x \in \partial \mathcal{O}$, $\nu(x)$ consists of a single element, the inward pointing unit normal to $\partial \mathcal{O}$ at x. For more general $\mathcal{O}$, Part 1 of Defn 7.1.1 says that $\nu(x)$ consists of all of the (there may be more than one at, say, a corner) inward pointing "normal vectors" to $\partial \mathcal{O}$ at x. Part 1a) gives a uniform exterior sphere condition, and Part 1b) ensures that no "normal vectors" are excluded from $\nu(x)$.

Part 2 and Part 3 of Definition 7.1.1 are regularity requirements on $\partial \mathcal{O}$ which ensure that the "normal vectors" don't fluctuate too wildly. In particular when $\mathcal{O}$ is bounded, Part 2 implies Part 3:

Theorem 7.1.2. *If $\mathcal{O}$ is bounded and admissible, then Part 2 implies Part 3 in Definition 7.1.1.*

Proof 7.1.3. As $\partial \mathcal{O}$ is compact it follows that $\nabla \phi$ is uniformly continuous on $\partial \mathcal{O}$. Take $\varepsilon = \frac{\alpha}{2}$. Then there exists a $\delta > 0$ such that

$$x, y \in \partial \mathcal{O}, |x - y| < \delta \Longrightarrow |\nabla \phi(x) - \nabla \phi(y)| < \varepsilon.$$

So for $y \in B(x,\delta)$, we have that for $\nu \in \nu(y)$,

$$\nabla\phi(x)\cdot\nu = \nabla\phi(y)\cdot\nu + (\nabla\phi(x) - \nabla\phi(y))\cdot\nu \geq \alpha - \varepsilon = \frac{\alpha}{2}$$

As $\partial\mathcal{O}$ is compact we can cover it with a finite collection of balls, $B(x_1, \frac{\delta}{2}), \ldots, B(x_n, \frac{\delta}{2})$, with $x_k \in \partial\mathcal{O}$. Take $a_k = \frac{\nabla\phi(x_k)}{|\nabla\phi(x_k)|}$ and let $\lambda = \frac{\alpha}{2\max\limits_{1\leq k\leq n}|\nabla\phi(x_k)|} > 0$. Then for $y \in B(x_k,\delta)$ and $\nu \in \nu(y)$, we have that

$$a_k\cdot\nu = \frac{\nabla\phi(x_k)\cdot\nu}{|\nabla\phi(x_k)|} \geq \frac{\alpha}{|\nabla\phi(x_k)|} \geq \lambda > 0$$

And so letting $R = \delta$ we are done.

In their paper [10], Lions and Sznitman show that there exists a unique solution to the Skorohod problem for any admissible domain $\mathcal{O}$.

7.2 Reflected Brownian Motion

We now can define reflected Brownian motion, i.e. Brownian motion which "reflects" off of the boundary to stay in some domain $\mathcal{O}$. We do this pathwise. Simply take any d-dimensional Brownian motion W_t defined on a probability space $(\Omega, \mathcal{F}, P)$ and define $X_t(\omega)$ to be the output to the Skorohod problem with $W_t(\omega)$ as the input. We call this process X_t *reflected Brownian motion* and it is clear that it behaves as we would hope.

7.3 Reflected SDE

Defining reflected SDE in general is not as simple a task as we cannot simply apply the deterministic Skorohod problem to each path. We give the full setup:

Let $\mathcal{O} \subset \mathbb{R}^d$ be a domain and to each $x \in \partial\mathcal{O}$ assign a non-empty collection $\nu(x)$ of unit vectors $\nu \in \mathbb{R}^d$. Suppose also that we have an r-dimensional Brownian motion W_t. For $x_0 \in \mathcal{O}$, $b \in C(\bar{\mathcal{O}};\mathbb{R}^d)$, and $\sigma \in C(\bar{\mathcal{O}};\mathrm{Mat}(\mathbb{R},d,r))$ [1], we seek a

[1] $\mathrm{Mat}(\mathbb{R},d,r)$ denotes the space of real valued $d\times r$ matrices.

continuous process $X_t : [0,\infty) \to \bar{\mathcal{O}}$ and a continuous process of locally bounded variation $L_t : [0,\infty) \to \mathbb{R}^d$ such that

$$\begin{aligned} X_t =& x_0 + \int_0^t \sigma(X_s)dW_s + \int_0^t b(X_s)ds + L_t, \\ |L|_t =& \int_0^t 1_{\{X_s \in \partial\mathcal{O}\}} d|L|_s, \text{ and } L_t = \int_0^t \nu(X_s)d|L|_s. \end{aligned} \tag{7.1}$$

Note that unlike the case of reflected Brownian motion, the coefficients σ and b depend on the final path X_t and so we cannot simply solve the unreflected SDE and then apply the deterministic Skorohod problem.

In [10], Lions and Sznitman show, using a Picard iteration scheme, that when $\mathcal{O}$ is admissible and b and σ are Lipschitz there exists a unique solution to the reflected SDE above.

In this thesis we consider reflected SDE of Stratonovich type, i.e. the first line of (7.1) is replaced by

$$X_t = x_0 + \int_0^t \sigma(X_s) \circ dW_s + \int_0^t b(X_s)ds + L_t$$

for a bounded admissible domain $\mathcal{O}$ and where $\sigma \in C^2(\bar{\mathcal{O}}; \mathrm{Mat}(\mathbb{R}, d, r))$ and b is Lipschitz continuous.

We now explain our result. Suppose we are given the standard d-dimensional Wiener space $(\Omega, \mathcal{F}_t, \mathbb{W})$ where $\Omega = C([0,\infty); \mathbb{R}^d)$, $\mathbb{W}$ is Wiener measure, and $\mathcal{F}_t$ is the usual filtration generated by paths under Wiener measure. We will use the usual convention of writing $W_t(\omega)$ for $\omega(t)$ and will often suppress the dependence of W_t on ω.

We introduce the following notation which we will use throughout this thesis: We will use the notation $\lfloor u \rfloor_N$ to denote the greatest N-dyadic rational (i.e. number of the form $m2^{-N}$, $m \in \mathbb{Z}$) which is less than or equal to u. Similarly we will use the notation $\lceil u \rceil_N$ to denote the smallest N-dyadic rational which is greater than or equal to u. When the choice of N is clear from context, we will suppress the letter N. Define $\Delta W^N_{\lfloor t \rfloor} := (W_{\lceil t \rceil_N} - W_{\lfloor t \rfloor_N})$ and define W^N_t to be the piecewise linear interpolation of

W_t over the N-dyadic rationals, i.e.

$$W_t^N(\omega) = W_{\lfloor t \rfloor}(\omega) + 2^N(t - \lfloor t \rfloor)(\Delta W_{\lfloor t \rfloor}^N(\omega))$$

Then for each ω let $X_t^N(\omega)$ and $L_t^N(\omega)$ denote the solution to the reflected ODE

$$dX_t^N = \sigma(X_t^N)dW_t^N(\omega) + b(X_t^N)dt + dL_t^N, \quad X_0^N = x_0 \tag{7.2}$$

We show then that, thought of as a stochastic process, the solution X_t^N converges weakly in distribution (i.e. as a measure on the space $C([0,\infty);\mathbb{R}^d)$) to the process X_t which solves the Stratonovich version of (7.1).

In their famous paper [18], Wong and Zakai show that, in the case without reflection, X_t^N converges to X_t in the uniform topology almost surely. Our result can therefore be seen as a weaker version of this result for the case of reflected SDE.

For the proofs in this thesis we will often consider the Hormander form way of expressing Stratonovich SDE. We write the Stratonovich SDE

$$dX_t = \sigma(X_t) \circ dW_t + b(X_t)dt$$

in Hormander form as

$$dX_t = \sum_{i=1}^{r} V_i(X_t)d(W_i)_t + V_0(X_t)dt \tag{7.3}$$

where the vector fields V_i are given by the columns of σ and $V_0 = b$.

An advantage of Hormander form over Stratonovich form is notational simplicity. In particular, the operator associated with (7.3) can be expressed as

$$[\mathcal{L}f](x) := \frac{1}{2}\sum_{i=1}^{r}[D_{V_i}^2 f](x) + [D_{V_0} f](x) \tag{7.4}$$

where $D_V = \sum_{j=1}^{d} V_j(x)\frac{\partial}{\partial x_j}$ is the differential operator associated with the vector field V. We could of course express $\mathcal{L}$ in terms of σ and b but the expression wouldn't be as simple.

7.4 Structure of the Presentation

The structure of this second half of the thesis is as follows. First the proof of our result is presented in 4 steps:

In Step 1 (Chapter 8), we argue that we have existence and uniqueness for the approximating ODE (7.2). We do this by a Picard iteration scheme following Lions and Sznitman's proof of Theorem 3.1 in [10].

In Step 2 (Chapter 9), we show that the measures P^N on $\left(C([0,\infty);\mathbb{R}^d)\right)^2$ induced by the pair (X_t^N, L_t^N) are tight by showing that X_t^N and L_t^N satisfy Kolmogorov's criterion.

In Step 3 (Chapter 10), we show that any limit point P of the P^N satisfies both a martingale problem and a submartingale problem for operators similar to (7.4).

Finally, in Step 4 (Chapter 11), we argue that the processes X_t and L_t under P are in fact a weak solution to the Stratonovich reflected SDE. Therefore, by the uniqueness result of Lions and Sznitman (Theorem 3.1 of [10]) we have that in fact the sequence of measures P^N converges weakly to P. This is our main result.

In the last chapter, Chapter 12, we present some applications/observations which follow from our result. In particular, we look at geometric properties of the solutions to Stratonovich reflected SDE in certain domains. Such properties have been used to prove the "hot spots conjecture" in various domains ([2] gives a good overview of the conjecture and this technique).

One final note about our convention for constants: Throughout this paper there will be many constants. For notational simplicity these are usually all lumped under the constant C. That is, C may change from line to line or indeed two instances of C on the same line may not be the same! When it is important to emphasize the dependence of the constant C on a parameter, say T, we will use the notation $C(T)$.

Chapter 8

Step 1: Existence of solutions to the approximating reflected ODE

Suppose that $\mathcal{O}$ is bounded and admissible, $x_0 \in \mathcal{O}$, and σ and b are Lipschitz continuous and let $(\Omega, \mathcal{F}_t, \mathbb{W})$ be the standard d-dimensional Wiener space, where we will use the usual convention of writing $W_t(\omega)$ for $\omega(t)$. We will show that for each fixed N and ω, there exists a solution, which we denote by $X_t^N(\omega)$ and $L_t^N(\omega)$ to the following reflected ODE:

$$\begin{aligned} X_t^N =& x_0 + \int_0^t \sigma(X_s^N) dW_s^N(\omega) + \int_0^t b(X_s^N) ds + L_t^N, \\ |L^N|_t =& \int_0^t 1_{\{X_s^N \in \partial\mathcal{O}\}} d|L^N|_s, \text{ and } L_t^N = \int_0^t \nu(X_s^N) d|L^N|_s, \end{aligned} \tag{8.1}$$

where X_t^N is a continuous function taking values in $\bar{\mathcal{O}}$ and L_t^N is a continuous function of locally bounded variation.

We solve (8.1) by a Picard iteration scheme following the proof of Theorem 3.1 in [10]: Fix N and ω and look at the mapping $X_t \longrightarrow Y_t$ where Y_t is given by

$$dY_t = \sigma(X_t) dW_t^N + b(X_t) dt$$

and the mapping $Y_t \longrightarrow Z_t$, where Z_t is the solution (guaranteed to exist and be unique by Theorem 2.2 of [10]) to the deterministic Skorohod problem

$$Z_t = Y_t + L_t$$

satisfying the usual conditions. Let $F : C([0,\infty);\mathbb{R}^d) \to C([0,\infty);\mathbb{R}^d)$ be the composition of these two maps, i.e. $Z_t = F(X_t)$. We will show that F has a unique fixed point. We first prove the following theorem:

Theorem 8.0.1. *Given two paths X_t and X'_t, let $Z_t = F(X_t)$ and $Z'_t = F(X'_t)$. Then*

$$|Z_T - Z'_T|^2 \leq C\int_0^T |Z_t - Z'_t|^2 dt + C\int_0^T |X_t - X'_t|^2 dt \tag{8.2}$$

Proof 8.0.2. Let ϕ be the function associated with $\mathcal{O}$ (see part 2 of Definition 7.1.1). For a constant γ, we have that

$$\begin{aligned}
&e^{-\gamma[\phi(Z_t)+\phi(Z'_t)]}d(e^{\gamma[\phi(Z_t)+\phi(Z'_t)]}|Z_t - Z'_t|^2) \\
=&2(Z_t - Z'_t)\cdot[(\sigma(X_t)dW_t^N + b(X_t)dt + dL_t) - (\sigma(X'_t)dW_t^N + b(X'_t)dt + dL'_t)] \\
&+ |Z_t - Z'_t|^2\gamma[\nabla\phi(Z_t)\cdot(\sigma(X_t)dW_t^N + b(X_t)dt + dL_t) \\
&+ \nabla\phi(Z'_t)\cdot(\sigma(X'_t)dW_t^N + b(X'_t)dt + dL'_t)] \\
=&[(2(Z_t - Z'_t) + \gamma|Z_t - Z'_t|^2\nabla\phi(Z_t))\cdot\nu(Z_t)]d|L|_t \\
&+ [(2(Z'_t - Z_t) + \gamma|Z_t - Z'_t|^2\nabla\phi(Z'_t))\cdot\nu(Z'_t)]d|L'|_t \\
&+ [2(Z_t - Z'_t)\cdot(\sigma(X_t) - \sigma(X'_t)) + \gamma|Z_t - Z'_t|^2(\nabla\phi(Z_t)\sigma(X_t) + \nabla\phi(Z'_t)\sigma(X'_t))]dW_t^N \\
&+ [2(Z_t - Z'_t)\cdot(b(X_t) - b(X'_t)) + \gamma|Z_t - Z'_t|^2(\nabla\phi(Z_t)\cdot b(X_t) + \nabla\phi(Z'_t)\cdot b(X'_t))]dt
\end{aligned}$$

Taking $\gamma = \frac{-2C_0}{\alpha}$, we have that (c.f. part 1a) of Definition 7.1.1) the first two terms are less than or equal to 0. As σ, b, and $\nabla\phi$ are Lipschitz continuous and bounded on $\bar{\mathcal{O}}$ ($\mathcal{O}$ is bounded), since $\frac{dW^N}{dt}$ is bounded on $\bar{\mathcal{O}}$ (N is fixed), and since $Z_0 = Z'_0 = x_0$, we have that

$$|Z_T - Z'_T|^2 \leq C\int_0^T |Z_t - Z'_t|^2 dt + C\int_0^T |Z_t - Z'_t||X_t - X'_t| dt + C\int_0^T |X_t - X'_t|^2 dt$$

Since $|Z_t - Z'_t||X_t - X'_t| \leq C|Z_t - Z'_t|^2 + C|X_t - X'_t|^2$, (8.2) is proved.

Once we have Theorem 8.0.1, it follows from a standard Picard iteration argument that F has a unique fixed point: First note that by Gronwall's inequality we have that

$$||Z - Z'||^2_{[0,T]} \leq C(T)||X - X'||^2_{[0,T]}$$

Plugging this into (8.2), we get that

$$||Z - Z'||^2_{[0,T]} \leq C(T) \int_0^T ||X - X'||^2_{[0,t]} dt$$

We now begin our Picard iteration. Let $(X_0)_t \equiv x_0 \in \mathcal{O}$, $X_{m+1} = F(X_m)$. Then

$$||X_{m+1} - X_m||^2_{[0,T]} \leq C(T) \int_0^T ||X_m - X_{m-1}||^2_{[0,t]} dt \tag{8.3}$$

Since $||X_1 - X_0|| \leq 2 \sup_{x \in \bar{\mathcal{O}}} |x| < \infty$, we have by iteration that

$$||X_{m+1} - X_m||^2_{[0,T]} \leq C(T) \frac{T^m}{m!}$$

It follows that

$$\sum_{m=0}^{\infty} ||X_{m+1} - X_m||^2_{[0,T]} < \infty$$

and so X_m converges uniformly in $[0, T]$ to a path X. It is easy to see by (8.3) that this path is unique. Since we have the convergence for all $T > 0$, there is a unique path X to which the X_m converge in $C([0, \infty); \mathbb{R}^d)$ equipped with the topology of uniform convergence on compacts. This path X solves (8.1) and we denote it by $X_t^N(\omega)$. Once we have $X_t^N(\omega)$, we get the corresponding $L_t^N(\omega)$ by applying F to $X_t^N(\omega)$.

Chapter 9

Step 2: Tightness of the Approximating Measures

In the previous chapter we showed for each N and ω the existence and uniqueness of paths $X_t^N(\omega)$ and $L_t^N(\omega)$ which solve the approximating reflected ODE (8.1). The corresponding pair of processes (X_t^N, L_t^N) induces a measure, which we denote by P^N, on the space $\left(C([0,\infty);\mathbb{R}^d)\right)^2$ which we will refer to as (X, L)-*pathspace*.

In this chapter, we show that the measures P^N on (X, L)-pathspace are tight. In the last section we also present some estimates which we will need for the next chapter. In this chapter we make the assumptions that $\mathcal{O}$ is bounded and admissible, and σ and b are Lipschitz continuous.

9.1 Tightness of the P^N

We begin with the following lemma:

Lemma 9.1.1. *For $s < t$ lying within the same N-dyadic interval, we have that*

$$|X_t^N - X_s^N| \leq C|W_t^N - W_s^N| + C(t-s) \tag{9.1}$$

Proof 9.1.2. Note that, for each ω, X_t^N is the solution to the deterministic Skorohod problem with input Y_t^N, where Y_t^N solves the SDE

$$dY_t^N = \sigma(X_t^N)dW_t^N + b(X_t^N)dt, \quad Y_0^N = x_0$$

By Theorem 2.2 in [10], since the input Y_t^N has bounded variation, so will the output X_t^N and

$$d|X^N|_t \leq d|Y^N|_t$$

As σ and b are bounded in $\mathcal{O}$, we have that

$$d|Y^N|_t \leq Cd|W^N|_t + Cdt$$

and so putting these together we have that

$$|X_t^N - X_s^N| \leq |X^N|_t - |X^N|_s \leq C(|W^N|_t - |W^N|_s) + C(t-s) = C|W_t^N - W_s^N| + C(t-s)$$

where the last equality holds because s and t are in the same N-dyadic interval.

We now state the main theorem of this section

Theorem 9.1.3. *(Kolmogorov Criterion) For each integer $m \geq 1$ and for all $0 \leq s < t \leq T$,*

$$E[|W_t^N - W_s^N|^{2m}] \leq C(m,T)(t-s)^{2^{m-1}} \tag{9.2}$$

$$E[|X_t^N - X_s^N|^{2m}] \leq C(m,T)(t-s)^{2^{m-1}} \tag{9.3}$$

$$E[|L_t^N - L_s^N|^{2m}] \leq C(m,T)(t-s)^{2^{m-1}} \tag{9.4}$$

In particular, the constant $C(m,T)$ does not depend on N.

The tightness of the P^N follows in a standard way from this theorem so the remainder of this section will be devoted to proving this theorem. We start by proving

(9.2).

Suppose first that $s < t$ lie within the same N-dyadic interval. Then, using the notation

$$\Delta W^N_{\lfloor u \rfloor} := W_{\lfloor u \rfloor + 2^{-N}} - W_{\lfloor u \rfloor}$$

we have that

$$\begin{aligned} E[|W^N_t - W^N_s|^{2m}] &= (2^N)^{2m}(t-s)^{2m} E[|\Delta W^N_{\lfloor s \rfloor}|^{2m}] \\ &= C(2^N)^{2m}(t-s)^{2m}(2^{-N})^{2m-1} \\ &= C(t-s)^{2m-1}\left(\frac{t-s}{2^N}\right)^{2m-1} \\ &\leq C(t-s)^{2m-1} \end{aligned}$$

as hoped.

Suppose next that $s < t$ are in adjacent N-dyadic intervals. Then

$$\begin{aligned} E[|W^N_t - W^N_s|^{2m}] &\leq CE[|W^N_t - W^N_{\lfloor t \rfloor}|^{2m}] + CE[|W^N_{\lfloor t \rfloor} - W^N_s|^{2m}] \\ &\leq C(t - \lfloor t \rfloor)^{2m-1} + C(\lfloor t \rfloor - s)^{2m-1} \\ &\leq C(t-s)^{2m-1} \end{aligned}$$

as hoped.

Finally, consider the case when $s < t$ are not in adjacent N-dyadic intervals. Then

$$\begin{aligned} E[|W^N_t - W^N_s|^{2m}] \leq & CE[|W^N_t - W^N_{\lfloor t \rfloor}|^{2m}] + CE[|W^N_{\lfloor t \rfloor} - W^N_{\lceil s \rceil}|^{2m}] \\ & + CE[|W^N_{\lceil s \rceil} - W^N_s|^{2m}] \\ = & CE[|W^N_t - W^N_{\lfloor t \rfloor}|^{2m}] + CE[|W_{\lfloor t \rfloor} - W_{\lceil s \rceil}|^{2m}] \\ & + CE[|W^N_{\lceil s \rceil} - W^N_s|^{2m}] \\ \leq & C(t - \lfloor t \rfloor)^{2m-1} + C(\lfloor t \rfloor - \lceil s \rceil)^{2m-1} + C(\lceil s \rceil - s)^{2m-1} \\ \leq & C(t-s)^{2m-1} \end{aligned}$$

as hoped. And so we have proved (9.2).

We will next prove (9.3), but first it will help to have the following Lemmas:

Lemma 9.1.4. *For $s < t$ lying within the same N-dyadic interval and $m \geq 0$, we have that*

$$E[(\int_s^t |W_u^N - W_{\lfloor u \rfloor}^N| \, d|W^N|_u)^{2^m}] \leq C(t-s)^{2^m} \tag{9.5}$$

and

$$E[(\int_s^t (u - \lfloor u \rfloor) \, d|W^N|_u)^{2^m}] \leq C(t-s)^{2^m} \tag{9.6}$$

Proof 9.1.5. Repeated applications of the Cauchy-Schwartz inequality show that

$$(\int_s^t f d\mu)^{2^m} \leq (t-s)^{2^m - 1} \int_s^t f^{2^m} d\mu$$

So from this we see that

$$\begin{aligned}
&E[(\int_s^t |W_u^N - W_{\lfloor u \rfloor}^N| d|W^N|_u)^{2^m}] \\
\leq &E[(t-s)^{2^m-1} \int_s^t |W_u^N - W_{\lfloor u \rfloor}^N|^{2^m} \left|\frac{dW_u^N}{du}\right|^{2^m} du] \\
= &(t-s)^{2^m-1} E[\sum \int_{\lfloor u \rfloor}^{\lfloor u \rfloor + 2^{-N}} |W_u^N - W_{\lfloor u \rfloor}^N|^{2^m} \left|\frac{dW_u^N}{du}\right|^{2^m} du] \\
= &(t-s)^{2^m-1} E[\sum (2^N)^{2^{m+1}} |\Delta W_{\lfloor u \rfloor}^N|^{2^{m+1}} \int_{\lfloor u \rfloor}^{\lfloor u \rfloor + 2^{-N}} (u - \lfloor u \rfloor)^{2^m} du] \\
\leq &(t-s)^{2^m-1} E[\sum (2^N)^{2^{m+1}} |\Delta W_{\lfloor u \rfloor}^N|^{2^{m+1}} (2^{-N})^{2^m+1}] \\
= &(t-s)^{2^m-1} (2^N)^{2^m-1} \sum E[|\Delta W_{\lfloor u \rfloor}^N|^{2^{m+1}}] \\
= &C(t-s)^{2^m-1} (2^N)^{2^m-1} \sum (2^{-N})^{2^m} \\
\leq &C(t-s)^{2^m-1} (2^{-N})(2^N(t-s) + 2) \\
\leq &C(t-s)^{2^m}
\end{aligned}$$

where we have used the fact that the sum has at most $(2^N(t-s)+2)$ terms and the fact that $t - s > 2^{-N}$. Thus we have proved (9.5).

The proof of (9.6) follows similarly (and in fact the inequality (9.6) is far from sharp).

We are now ready to prove (9.3). We first handle the case when $s < t$ are in the same N-dyadic interval. In this case we have by (9.1) that

$$\begin{aligned} E[|X_t^N - X_s^N|^{2^m}] &\leq CE[|W_t^N - W_s^N|^{2^m}] + CE[(t-s)^{2^m}] \\ &\leq C(t-s)^{2^{m-1}} \end{aligned}$$

where we have used (9.2) and the fact we are in the interval $[0,T]$ for the last inequality.

Next, we consider the case when $s < t$ are in adjacent intervals. In this case,

$$\begin{aligned} E[|X_t^N - X_s^N|^{2^m}] &\leq CE[|X_t^N - X_{\lfloor t\rfloor}^N|^{2^m}]] + CE[|X_{\lfloor t\rfloor}^N - X_s^N|^{2^m}] \\ &\leq C(t-\lfloor t\rfloor)^{2^{m-1}} + C(\lfloor t\rfloor - s)^{2^{m-1}} \\ &\leq C(t-s)^{2^{m-1}} \end{aligned}$$

It remains to prove (9.3) in the case when $t - s > 2^{-N}$. For this we will need the following inequality:

Theorem 9.1.6. *Let ϕ be the function associated with $\mathcal{O}$ (see part 2 of Definition 7.1.1). Then there is a constant γ such that*

$$|X_t^N - X_s^N|^2 \leq \int_s^t Z_u^N dW_u^N + \int_s^t V_u^N du \tag{9.7}$$

where

$$Z_u^N = e^{-\gamma\phi(X_t^N)+\gamma\phi(X_u^N)}(2(X_u^N - X_s^N) + |X_u^N - X_s^N|^2\gamma\nabla\phi(X_u^N))\sigma(X_u^N)$$

and

$$V_u^N = e^{-\gamma\phi(X_t^N)+\gamma\phi(X_u^N)}(2(X_u^N - X_s^N) + |X_u^N - X_s^N|^2\gamma\nabla\phi(X_u^N))b(X_u^N)$$

Proof 9.1.7. We compute,

$$\begin{aligned}
&d(e^{\gamma\phi(X_t^N)}|X_t^N - X_s^N|^2) \\
&\quad = e^{\gamma\phi(X_t^N)}(2(X_t^N - X_s^N) \cdot (\sigma(X_t^N)dW_t^N + b(X_t^N)dt + dL_t^N)) \\
&\quad\quad + e^{\gamma\phi(X_t^N)}|X_t^N - X_s^N|^2(\gamma\nabla\phi(X_t^N) \cdot (\sigma(X_t^N)dW_t^N + b(X_t^N)dt + dL_t^N)) \\
&\quad = e^{\gamma\phi(X_t^N)}(2(X_t^N - X_s^N) \cdot \nu(X_t^N) + \gamma|X_t^N - X_s^N|^2\nabla\phi(X_t^N) \cdot \nu(X_t^N))d|L^N|_t \\
&\quad\quad + e^{\gamma\phi(X_t^N)}(2(X_t^N - X_s^N) + \gamma|X_t^N - X_s^N|^2\nabla\phi(X_t^N))\sigma(X_t^N)dW_t^N \\
&\quad\quad + e^{\gamma\phi(X_t^N)}(2(X_t^N - X_s^N) + \gamma|X_t^N - X_s^N|^2\nabla\phi(X_t^N))b(X_t^N)dt
\end{aligned}$$

Taking $\gamma = \frac{-2C_0}{\alpha}$, we have (c.f. part 1a) of Definition 7.1.1) that

$$\begin{aligned}
&d(e^{\gamma\phi(X_t^N)}|X_t^N - X_s^N|^2) \\
&\quad \leq e^{\gamma\phi(X_t^N)}(2(X_t^N - X_s^N) + \gamma|X_t^N - X_s^N|^2\nabla\phi(X_t^N))\sigma(X_t^N)dW_t^N \\
&\quad\quad + e^{\gamma\phi(X_t^N)}(2(X_t^N - X_s^N) + \gamma|X_t^N - X_s^N|^2\nabla\phi(X_t^N))b(X_t^N)dt
\end{aligned}$$

from which (9.7) follows.

With this theorem in hand we are ready to proceed. Although our proof is *not* by induction, we will need to prove the case $m = 1$ first. We have by (9.7) that

$$\begin{aligned}
E[|X_t^N - X_s^N|^2] &\leq E[\int_s^t Z_u^N dW_u^N + \int_s^t V_u^N du] \\
&= E[\int_s^t (Z_u^N - Z_{\lfloor u \rfloor}^N)dW_u^N + \int_s^t Z_{\lfloor u \rfloor}^N dW_u^N + \int_s^t V_u^N du] \\
&= E[\int_s^t (Z_u^N - Z_{\lfloor u \rfloor}^N)dW_u^N] + 0 + E[\int_s^t V_u^N du]
\end{aligned}$$

We estimate the term on the right:

$$\begin{aligned}
E[\int_s^t V_u^N du] &\leq E[\int_s^t |V_u^N| du] \\
&\leq C(t-s)
\end{aligned}$$

since $|V_u^N|$ is bounded independent of N and u. We next estimate the term on the left:

$$\begin{aligned}E[\int_s^t (Z_u^N - Z_{\lfloor u\rfloor}^N)dW_u^N] \leq & E[\int_s^t |Z_u^N - Z_{\lfloor u\rfloor}^N|d|W^N|_u] \\ \leq & CE[\int_s^t |X_u^N - X_{\lfloor u\rfloor}^N|d|W^N|_u] \\ \leq & CE[\int_s^t |W_u^N - W_{\lfloor u\rfloor}^N|d|W^N|_u] \\ & + CE[\int_s^t (u - \lfloor u\rfloor)d|W^N|_u] \\ \leq & C(t-s)\end{aligned}$$

where in the second line we have used the fact that Z_u^N is a Lipschitz function of X_u^N and in the last line we have used (9.5) and (9.6). We have therefore shown that

$$E[|X_t^N - X_s^N|^2] \leq C(t-s) \tag{9.8}$$

which proves (9.3) in the case when $m = 1$. We can now prove (9.3) for larger m, but first we will need the following lemma:

Lemma 9.1.8. *The random variable*

$$\int_s^t Z_{\lfloor u\rfloor}^N dW_u^N$$

is a normal random variable with mean 0 *and variance which is bounded above by* $C(t-s)$.

Proof 9.1.9. As $Z_{\lfloor u\rfloor}^N$ is adapted, only the bound on the variance requires comment. Note that by the Ito isometry,

$$\begin{aligned}E[(\int_s^t Z_{\lfloor u\rfloor}^N dW_u^N)^2] &= \int_s^t E[|Z_{\lfloor u\rfloor}^N|^2]du] \\ &\leq \int_s^t E[|X_{\lfloor u\rfloor}^N - X_s^N|^2]du] \\ &\leq C\int_s^t (\lfloor u\rfloor - s)du \\ &\leq C(t-s)^2\end{aligned}$$

where for the second to last inequality we have used (9.8).

We can finally prove (9.3) for $m \geq 2$. Note that by taking the appropriate power of (9.7), we have that

$$\begin{aligned} E[|X_t^N - X_s^N|^{2^{m+1}}] \leq & E[(\int_s^t Z_u^N dW_u^N + \int_s^t V_u^N du)^{2^m}] \\ \leq & CE[(\int_s^t (Z_u^N - Z_{\lfloor u \rfloor}^N) dW_u^N)^{2^m}] \\ & + CE[(\int_s^t Z_{\lfloor u \rfloor}^N dW_u^N)^{2^m}] + CE[(\int_s^t V_u^N du)^{2^m}] \end{aligned}$$

So it remains to bound these three terms. From Lemma 9.1.8 it follows that the second term is bounded from above by $C(t-s)^{2^m}$. Since $|V_u^N|$ is bounded uniformly in N and u, it follows that the third term is bounded from above by $C(t-s)^{2^m}$. Finally, for the first term we have that

$$\begin{aligned} CE[(\int_s^t (Z_u^N - Z_{\lfloor u \rfloor}^N) dW_u^N)^{2^m}] \leq & CE[(\int_s^t |Z_u^N - Z_{\lfloor u \rfloor}^N| d|W^N|_u)^{2^m}] \\ \leq & CE[(\int_s^t |X_u^N - X_{\lfloor u \rfloor}^N| d|W^N|_u)^{2^m}] \\ \leq & CE[(\int_s^t |W_u^N - W_{\lfloor u \rfloor}^N| d|W^N|_u)^{2^m}] \\ & + CE[(\int_s^t (u - \lfloor u \rfloor) d|W^N|_u)^{2^m}] \\ \leq & C(t-s)^{2^m} \end{aligned}$$

where in the last inequality we have used (9.5) and (9.6). Thus we have shown that

$$E[|X_t^N - X_s^N|^{2^{m+1}}] \leq C(m,T)(t-s)^{2^m}$$

and proved (9.3).

We next prove (9.4). Note that as

$$dX_t^N = \sigma(X_t^N) dW_t^N + b(X_t^N) dt + dL_t^N$$

we have that

$$E[|L_t^N - L_s^N|^{2^m}] \leq CE[|X_t^N - X_s^N|^{2^m}] + CE[(\int_s^t \sigma(X_u^N)dW_u^N)^{2^m}] + CE[(\int_s^t b(X_u^N)du)^{2^m}]$$

We have shown that the first term is bounded from above by $C(t-s)^{2^{m-1}}$. As b is bounded, the third term is bounded from above by $C(t-s)^{2^{m-1}}$ as well (remember $0 \leq s < t \leq T$). For the second term we have that

$$\begin{aligned} CE[(\int_s^t \sigma(X_u^N)dW_u^N)^{2^m}] \leq & CE[(\int_s^t (\sigma(X_u^N) - \sigma(X_{\lfloor u \rfloor}^N))dW_u^N)^{2^m}] \\ & + CE[(\int_s^t \sigma(X_{\lfloor u \rfloor}^N)dW_u^N)^{2^m}] \\ \leq & CE[(\int_s^t (\sigma(X_u^N) - \sigma(X_{\lfloor u \rfloor}^N))dW_u^N)^{2^m}] + C(t-s)^{2^{m-1}} \end{aligned}$$

where the last inequality we have used the fact that $\int_s^t \sigma(X_{\lfloor u \rfloor}^N)dW_u^N$ is Gaussian and σ is bounded. Finally,

$$\begin{aligned} CE[(\int_s^t (\sigma(X_u^N) - \sigma(X_{\lfloor u \rfloor}^N))dW_u^N)^{2^m}] \leq & CE[(\int_s^t |X_u^N - X_{\lfloor u \rfloor}^N| d|W^N|_u)^{2^m}] \\ \leq & CE[(\int_s^t |W_u^N - W_{\lfloor u \rfloor}^N| d|W^N|_u)^{2^m}] \\ & + CE[(\int_s^t (u - \lfloor u \rfloor) d|W^N|_u)^{2^m}] \\ \leq & C(t-s)^{2^{m-1}} \end{aligned}$$

where in the last inequality follows from (9.5) and (9.6). Putting these all together we have shown that

$$E[|L_t^N - L_s^N|^{2^m}] \leq C(t-s)^{2^{m-1}}$$

and so we have proved (9.4).

9.2 Additional Estimates

In this section we show that for all $0 \le s < t$,

$$E[\int_s^t |L_u^N - L_{\lfloor u \rfloor}^N| d|W_i^N|_u] \to 0 \text{ as } N \to \infty \tag{9.9}$$

Recall our assumptions for this chapter that $\mathcal{O}$ is bounded and admissible, and σ and b are Lipschitz continuous.

We make the following definitions:

Definition 9.2.1.

$$\begin{aligned}
\mathcal{X}_\eta^N(T) &:= \sup_{0 \le s < t \le T} \left\{ \frac{|X_t^N - X_s^N|}{|t-s|^\eta} \right\} \\
\mathcal{W}_\beta^N(T) &:= \sup_{0 \le s < t \le T} \left\{ \frac{|(W_i^N)_t - (W_i^N)_s|}{|t-s|^\beta} \right\} \\
\mathcal{L}_\gamma^N(T) &:= \sup_{0 \le s < t \le T} \left\{ \frac{|L_t^N - L_s^N|}{|t-s|^\gamma} \right\}
\end{aligned}$$

Note that we have suppressed the dependence of $\mathcal{W}_\beta^N$ on i. This is to simplify notation, and in the end we will only need the expectation of $\mathcal{W}_\beta^N$ which is independent of i.

We will need the following Lemma:

Lemma 9.2.2. *For each $\eta > 0$, and for all $0 \le s < t$,*

$$|L^N|_t - |L^N|_s \le (t-s) C R^{-\frac{1}{\eta}} (\mathcal{X}_\eta^N(t))^{\frac{1}{\eta}} \left|\left|L^N\right|\right|_{[s,t]} \tag{9.10}$$

where $\left|\left|L^N\right|\right|_{[s,t]} := \sup_{s \le u \le t} \left|L_u^N\right|$.

Our proof follows the proof of Lemma 1.2 in [10].

Proof 9.2.3. As $\mathcal{O}$ is admissible it satisfies Part 3 of Definition 7.1.1. Let $O_1, \ldots, O_n$ denote the open sets $B(x_1, 2R) \cap O, \ldots, B(x_n, 2R) \cap O$ and let O_0 be an open set such that $\bar{O}_0 \subset O$ and $\bar{O} \subset \bigcup_{i=1}^n B(x_i, R) \cup O_0$. It follows then that for each u, X_u^N is in one of the O_k. In particular, X_s^N is in one or more of the O_k and we "assign it" to one of them arbitrarily. We next construct a sequence of times T_k corresponding to each time the process X_u^N leaves the O_k it is currently assigned to, letting $T_0 = s$. Upon leaving one of the O_k, if the process is in any of the $B(x_j, R)$ it is assigned to that O_j (if there is more than one O_j, choose the one with the smallest index), otherwise it is assigned to O_0. We examine how the total variation of L^N changes between these re-assignment times:

In the case X_u^N is assigned to O_0, it never touches the boundary and $|L^N|_{T_{m+1}} - |L^N|_{T_m} = 0$. In the case X_u^N is assigned to one of the O_k we have that

$$\begin{aligned}(L^N_{T_{m+1}} - L^N_{T_m}) \cdot a_{i_m} &= \int_{T_m}^{T_{m+1}} \nu(X_u^N) \cdot a_{i_m} d|L^N|_u \\ &\geq \lambda(|L^N|_{T_{m+1}} - |L^N|_{T_m})\end{aligned}$$

And so in any case we have that

$$|L^N|_{T_{m+1}} - |L^N|_{T_m} \leq C|L^N_{T_{m+1}} - L^N_{T_m}| \leq C||L^N||_{[s,t]}$$

By our re-assignment times construction we have that $|X^N_{T_{m+1}} - X^N_{T_m}| \geq R$. It therefore follows that

$$\frac{R}{(T_{m+1} - T_m)^\eta} \leq \frac{|X^N_{T_{m+1}} - X^N_{T_m}|}{(T_{m+1} - T_m)^\eta} \leq \mathcal{X}_\eta^N(t)$$

therefore

$$\sup\{m : T_m < t\} \leq \frac{(t-s)}{(T_{m+1} - T_m)} \leq (t-s)\left(\frac{\mathcal{X}_\eta^N(t)}{R}\right)^{\frac{1}{\eta}}$$

from which it follows that

$$\begin{aligned}|L^N|_t - |L^N|_s &\leq \sum_{\{m:T_m<t\}} |L^N|_{T_{m+1}} - |L^N|_{T_m} \\ &\leq \sum_{\{m:T_m<t\}} C||L^N||_{[s,t]} \leq (t-s)CR^{-\frac{1}{\eta}}(\mathcal{X}_\eta^N(t))^{\frac{1}{\eta}}||L^N||_{[s,t]}\end{aligned}$$

We will also need the following lemma.

Lemma 9.2.4. *For each $\beta > 0$, and for all $0 \le s < t$,*

$$\int_s^t |L_u^N - L_{\lfloor u\rfloor}^N| d|W_i^N|_u \le (|L^N|_t - |L^N|_{\lfloor s\rfloor}) 2^{-N\beta} \mathcal{W}_\beta^N(t) \tag{9.11}$$

Proof 9.2.5.

$$\begin{aligned}
\int_s^t |L_u^N - L_{\lfloor u\rfloor}^N| d|W_i^N|_u &= \int_s^{\lceil s\rceil} |L_u^N - L_{\lfloor s\rfloor}^N| d|W_i^N|_u + \sum \int_{\lfloor u\rfloor}^{\lceil u\rceil} |L_u^N - L_{\lfloor u\rfloor}^N| d|W_i^N|_u \\
&\quad + \int_{\lfloor t\rfloor}^t |L_u^N - L_{\lfloor t\rfloor}^N| d|W_i^N|_u \\
&\le \int_s^{\lceil s\rceil} |L^N|_u - |L^N|_{\lfloor s\rfloor} d|W_i^N|_u + \sum \int_{\lfloor u\rfloor}^{\lceil u\rceil} |L^N|_u - |L^N|_{\lfloor u\rfloor} d|W_i^N|_u \\
&\quad + \int_{\lfloor t\rfloor}^t |L^N|_u - |L^N|_{\lfloor t\rfloor} d|W_i^N|_u \\
&\le \int_s^{\lceil s\rceil} |L^N|_{\lceil s\rceil} - |L^N|_{\lfloor s\rfloor} d|W_i^N|_u + \sum \int_{\lfloor u\rfloor}^{\lceil u\rceil} |L^N|_{\lceil u\rceil} - |L^N|_{\lfloor u\rfloor} d|W_i^N|_u \\
&\quad + \int_{\lfloor t\rfloor}^t |L^N|_t - |L^N|_{\lfloor t\rfloor} d|W_i^N|_u
\end{aligned}$$

Now we have that

$$\begin{aligned}
\int_{\lfloor u\rfloor}^{\lceil u\rceil} |L^N|_{\lceil u\rceil} - |L^N|_{\lfloor u\rfloor} d|W_i^N|_u &= (|L^N|_{\lceil u\rceil} - |L^N|_{\lfloor u\rfloor}) |(W_i^N)_{\lceil u\rceil} - (W_i^N)_{\lfloor u\rfloor}| \\
&\le (|L^N|_{\lceil u\rceil} - |L^N|_{\lfloor u\rfloor}) 2^{-N\beta} \mathcal{W}_\beta^N(t)
\end{aligned}$$

We have a similar estimate for the other two terms and so combining everything together we have that

$$\int_s^t |L_u^N - L_{\lfloor u\rfloor}^N| d|W_i^N|_u \le (|L^N|_t - |L^N|_{\lfloor s\rfloor}) 2^{-N\beta} \mathcal{W}_\beta^N(t)$$

as desired.

We now prove (9.9). It is simple to check that

$$||L^N||_{[s,t]} \le t^\gamma \mathcal{L}_\gamma^N(t)$$

Putting this together with (9.10) and (9.11) we have that

$$E[\int_s^t |L_u^N - L_{\lfloor u\rfloor}^N| d|W_i^N|_u] \leq 2^{-N\beta}(t - \lfloor s\rfloor)CR^{-\frac{1}{\eta}}t^{\gamma}E[(\mathcal{X}_\eta^N(t))^{\frac{1}{\eta}}\mathcal{L}_\gamma^N(t)\mathcal{W}_\beta^N(t)]$$

and so it suffices to show that $E[(\mathcal{X}_\eta^N(t))^{\frac{1}{\eta}}\mathcal{L}_\gamma^N(t)\mathcal{W}_\beta^N(t)] \leq K < \infty$, independent of N. For this we will use the following theorem (c.f. Theorem 3.4.16 in [13]):

Theorem 9.2.6. *Let Z_t be a process such that*

$$E[|Z_t - Z_s|^r] \leq C|t - s|^{1+a}, 0 \leq s \leq t \leq T$$

Then for each $b \in (0, \frac{a}{r})$, there exists a constant K such that

$$P(\sup_{0\leq s<t\leq T}\left\{\frac{|Z_t - Z_s|}{|t - s|^b}\right\} \geq R) \leq \frac{KC}{R^r}$$

Take $\eta = \beta = \gamma = \frac{1}{4}$. This theorem gives us (c.f. the estimates (9.2),(9.3), and (9.4)) that $\mathcal{X}_\eta^N(t)$, $\mathcal{W}_\beta^N(t)$, and $\mathcal{L}_\gamma^N(t)$ have moments of all orders which are bounded independently of N. And so it follows that $E[(\mathcal{X}_\eta^N(t))^{\frac{1}{\eta}}\mathcal{L}_\gamma^N(t)\mathcal{W}_\beta^N(t)] \leq K < \infty$ as desired.

Chapter 10

Step 3: The Martingale and Submartingale Problem

In the previous section, we showed that the measures P^N are tight on (X, L)-pathspace. In this section we argue that any measure P on (X, L)-path space which is the weak limit of the P^N solves a martingale and submartingale problem.

In the previous chapters, we have only needed the assumptions that $\mathcal{O}$ is bounded and admissible, and σ and b are Lipschitz continuous. In this chapter, we will need to make the additional assumption that $\sigma \in C^2(\bar{\mathcal{O}}; \mathrm{Mat}(\mathbb{R}, d, r))$.

We have shown that there exists a unique solution (X_t^N, L_t^N) (pathwise) to the reflected ODE

$$dX_t^N = \sigma(X_t^N)dW_t^N + b(X_t^N)dt + dL_t^N, \quad X_0^N = x_0 \tag{10.1}$$

For the proofs in this section, it will simplify notation to express (10.1) in Hormander form:

$$dX_t^N = \sum_{i=1}^{r} V_i(X_t^N)d(W_i^N)_t + V_0(X_t^N)dt + dL_t^N, \quad X_0^N = x_0 \tag{10.2}$$

where the vector fields V_i are just the columns of σ for $i = 1, \ldots, r$ and $V_0 = b$. In this section we will prove the following theorem:

Theorem 10.0.7. *Let P be the weak limit of (a subsequence of) our measures P^N*

on (X,L) pathspace. Then $\forall f \in C^2(\mathbb{R}^d)$,

$$f(X_t - L_t) - f(x_0) - \int_0^t \frac{1}{2}\sum_{i=1}^r [D^2_{V_i} f(\cdot - L_s)](X_s) + [D_{V_0} f(\cdot - L_s)](X_s) ds \qquad (10.3)$$

is a P-martingale with respect to the filtration $\mathcal{F}_t$ generated by the paths under the measure P.

Also, $\forall f \in C^2(\mathbb{R}^d)$ such that $\frac{\partial f}{\partial \nu}(x) \geq 0$, $\forall x \in \partial \mathcal{O}$, $\forall \nu \in \nu(x)$,

$$f(X_t) - f(x_0) - \int_0^t \frac{1}{2}\sum_{i=1}^r [D^2_{V_i} f](X_s) + [D_{V_0} f](X_s) ds \qquad (10.4)$$

is a P-sub-martingale with respect to the filtration $\mathcal{F}_t$ generated by the paths under the measure P.

For this we will need the following lemmas.

Lemma 10.0.8.

$$E^{P^N}[\int_s^t |L_u - L_{\lfloor u \rfloor_N}| d|W_i^N|_u] \to 0 \qquad (10.5)$$

Proof 10.0.9. See previous section.

Lemma 10.0.10.

$$E[\int_s^t O(|W_u^N - W_{\lfloor u \rfloor}^N|^2) d(W_i^N)_u] \to 0 \text{ as } n \to \infty \qquad (10.6)$$

where $O(|W_u^N - W_{\lfloor u \rfloor}^N|^2)$ denotes any quantity whose absolute value is dominated by $C|W_u^N - W_{\lfloor u \rfloor}^N|$ for some constant $C > 0$.

Proof 10.0.11. We have that

$$
\begin{aligned}
E[\int_s^t O(|W_u^N - W_{\lfloor u\rfloor}^N|^2) d(W_i^N)_u] \leq & E[\int_s^t C|W_u^N - W_{\lfloor u\rfloor}^N|^2 d|W_i^N|_u] \\
= & C\sum E[\int_{\lfloor u\rfloor}^{\lceil u\rceil} |W_u^N - W_{\lfloor u\rfloor}^N|^2 d|W_i^N|_u] \\
= & C\sum E[\int_{\lfloor u\rfloor}^{\lceil u\rceil} (2^N)^3 (u - \lfloor u\rfloor)^2 |\Delta W_{\lfloor u\rfloor}^N|^2 |\Delta (W_i^N)_{\lfloor u\rfloor}|] \\
\leq & C\sum E[|\Delta W_{\lfloor u\rfloor}^N|^2 |\Delta (W_i^N)_{\lfloor u\rfloor}|] \\
\leq & C\sum (2^{-N})^{\frac{3}{2}} \to 0
\end{aligned}
$$

as there are at most $2^N(t-s)+2$ terms in the sum.

Lemma 10.0.12. *Let $F : (\mathbb{R}^d)^4 \to \mathbb{R}$ be a bounded, Lipschitz continuous function. Let $s < t$ both be M-dyadic rationals for some integer M. Let $A \subset \mathcal{F}_s$. Then*

$$
\begin{aligned}
E^{P^N}[&\int_s^t F(X_u, X_{\lfloor u\rfloor}, X_u - L_u, X_{\lfloor u\rfloor} - L_{\lfloor u\rfloor})((W_j^N)_u - (W_j^N)_{\lfloor u\rfloor}) d(W_i^N)_u, A] \\
&- E^{P^N}[\delta_{ij}\frac{1}{2}\int_s^t F(X_u, X_u, X_u - L_u, X_u - L_u) du, A] \to 0
\end{aligned}
$$

Proof 10.0.13. First note that

$$
\begin{aligned}
&E^{P^N}[\int_s^t F(X_u, X_{\lfloor u\rfloor}, X_u - L_u, X_{\lfloor u\rfloor} - L_{\lfloor u\rfloor})((W_j^N)_u - (W_j^N)_{\lfloor u\rfloor}) d(W_i^N)_u, A] \\
&- E^{P^N}[\int_s^t F(X_{\lfloor u\rfloor}, X_{\lfloor u\rfloor}, X_{\lfloor u\rfloor} - L_{\lfloor u\rfloor}, X_{\lfloor u\rfloor} - L_{\lfloor u\rfloor})((W_j^N)_u - (W_j^N)_{\lfloor u\rfloor}) d(W_i^N)_u, A] \\
\leq & E^{P^N}[\int_s^t C(|W_u^N - W_{\lfloor u\rfloor}^N| + (u - \lfloor u\rfloor))|(W_j^N)_u - (W_j^N)_{\lfloor u\rfloor}| d|W_i^N|_u] \to 0
\end{aligned}
$$

where the last line follows from an argument like the one used to prove Lemma 10.0.10. Next note that

$$
\begin{aligned}
&E^{P^N}[\delta_{ij}\frac{1}{2}\int_s^t F(X_u, X_u, X_u - L_u, X_u - L_u)du, A] \\
&- E^{P^N}[\delta_{ij}\frac{1}{2}\int_s^t F(X_{\lfloor u\rfloor}, X_{\lfloor u\rfloor}, X_{\lfloor u\rfloor} - L_{\lfloor u\rfloor}, X_{\lfloor u\rfloor} - L_{\lfloor u\rfloor})du, A] \\
\leq &\delta_{ij}\frac{1}{2}E^{P^N}[\int_s^t C(|W_u^N - W_{\lfloor u\rfloor}^N| + (u - \lfloor u\rfloor))du] \to 0
\end{aligned}
$$

where the last line follows from an argument like the one used to prove Lemma 10.0.10. Therefore, it suffices to show that

$$
\begin{aligned}
E^{P^N}[&\int_s^t F(X_{\lfloor u\rfloor}, X_{\lfloor u\rfloor}, X_{\lfloor u\rfloor} - L_{\lfloor u\rfloor}, X_{\lfloor u\rfloor} - L_{\lfloor u\rfloor})((W_j^N)_u - (W_j^N)_{\lfloor u\rfloor})d(W_i^N)_u, A] \\
&- E^{P^N}[\delta_{ij}\frac{1}{2}\int_s^t F(X_{\lfloor u\rfloor}, X_{\lfloor u\rfloor}, X_{\lfloor u\rfloor} - L_{\lfloor u\rfloor}, X_{\lfloor u\rfloor} - L_{\lfloor u\rfloor})du, A] \to 0 \quad (10.7)
\end{aligned}
$$

In fact, once N is greater than or equal to M, we have that s and t are N-dyadic rationals and the difference above is 0 for each N. To show this, it suffices by linearity to show that

$$
\begin{aligned}
E^{P^N}[&\int_{m2^{-N}}^{(m+1)2^{-N}} F(X_{\lfloor u\rfloor}, X_{\lfloor u\rfloor}, X_{\lfloor u\rfloor}-L_{\lfloor u\rfloor}, X_{\lfloor u\rfloor}-L_{\lfloor u\rfloor})((W_j^N)_u-(W_j^N)_{\lfloor u\rfloor})d(W_i^N)_u, A] \\
&= E^{P^N}[\delta_{ij}\frac{1}{2}\int_{m2^{-N}}^{(m+1)2^{-N}} F(X_{\lfloor u\rfloor}, X_{\lfloor u\rfloor}, X_{\lfloor u\rfloor} - L_{\lfloor u\rfloor}, X_{\lfloor u\rfloor} - L_{\lfloor u\rfloor})du, A] \quad (10.8)
\end{aligned}
$$

but this follows immediately from the fact that

$$
\begin{aligned}
E\Big[&\int_{\lfloor u\rfloor}^{\lceil u\rceil}((W_j^N)_u - (W_j^N)_{\lfloor u\rfloor})d(W_i^N)_u\Big|\mathcal{F}_{\lfloor u\rfloor}\Big] \\
&= E\Big[\int_{\lfloor u\rfloor}^{\lceil u\rceil}(2^N)^2(u - \lfloor u\rfloor)|\Delta(W_j^n)_{\lfloor u\rfloor}||\Delta(W_i^n)_{\lfloor u\rfloor}|du\Big|\mathcal{F}_{\lfloor u\rfloor}\Big] = \frac{1}{2}\delta_{ij}2^{-N}
\end{aligned}
$$

And so we have proved Lemma 10.0.12.

We are now ready to give our proof of Theorem 10.0.7. First we prove (10.3).

Note that from (10.1), that for $f \in C_c^\infty$,

$$\begin{aligned}
&f(X_t^N - L_t^N) - f(x_0) \\
&\quad = \sum_{i=1}^{r} \int_0^t V_i(X_s^N) \cdot \nabla f(X_s^N - L_s^N) d(W_i^N)_s + \underbrace{\int_0^t V_0(X_s^N) \cdot \nabla f(X_s^N - L_s^N) ds}_{\text{drift}} \\
&\quad = \sum_{i=1}^{r} \int_0^t [V_i(X_s^N) \cdot \nabla f(X_s^N - L_s^N) - V_i(X_{\lfloor s \rfloor}^N) \cdot \nabla f(X_{\lfloor s \rfloor}^N - L_{\lfloor s \rfloor}^N)] d(W_i^N)_s \\
&\qquad + \underbrace{\sum_{i=1}^{r} \int_0^t V_i(X_s^N) \cdot \nabla f(X_{\lfloor s \rfloor}^N - L_{\lfloor s \rfloor}^N) d(W_i^N)_s}_{\text{MG}} + \text{drift}
\end{aligned}$$

We recognize that the second to last term is a martingale for each N, and so we will henceforth refer to that term as "MG". Similarly, we will simply refer to the last term in the first line as "drift" as the last term arises from the drift term $b(X_t^N)dt$

and won't change during our computations. Continuing, we have that

$$
\begin{aligned}
&f(X_t^N - L_t^N) - f(x_0) \\
&= \sum_{i=1}^r \int_0^t [V_i(X_s^N) \cdot \nabla f(X_s^N - L_s^N) - V_i(X_{\lfloor s \rfloor}^N) \cdot \nabla f(X_{\lfloor s \rfloor}^N - L_{\lfloor s \rfloor}^N)] d(W_i^N)_s \\
&\quad + \text{ MG } + \text{ drift} \\
&= \sum_{i=1}^r \int_0^t V_i(X_s^N) \cdot [\nabla f(X_s^N - L_s^N) - \nabla f(X_{\lfloor s \rfloor}^N - L_{\lfloor s \rfloor}^N)] \\
&\quad + [V_i(X_s^N) - V_i(X_{\lfloor s \rfloor}^N)] \cdot \nabla f(X_{\lfloor s \rfloor}^N - L_{\lfloor s \rfloor}^N) d(W_i^N)_s + \text{ MG } + \text{ drift} \\
&= \sum_{i=1}^r \int_0^t \sum_{j=1}^r \Big[V_i(X_s^N)^T D^2 f(X_{\lfloor s \rfloor}^N - L_{\lfloor s \rfloor}^N) V_j(X_s^N)((W_j^N)_s - (W_j^N)_{\lfloor s \rfloor}) \\
&\quad + V_j(X_s^N)^T \nabla V_i(X_{\lfloor s \rfloor}^N) \nabla f(X_{\lfloor s \rfloor}^N - L_{\lfloor s \rfloor}^N)((W_j^N)_s - (W_j^N)_{\lfloor s \rfloor}) \Big] \\
&\quad + V_i(X_s^N)^T D^2 f(X_{\lfloor s \rfloor}^N - L_{\lfloor s \rfloor}^N) V_0(X_s^N)(s - \lfloor s \rfloor) \\
&\quad + V_0(X_s^N)^T \nabla V_i(X_{\lfloor s \rfloor}^N) \nabla f(X_{\lfloor s \rfloor}^N - L_{\lfloor s \rfloor}^N)(s - \lfloor s \rfloor) \\
&\quad + O(|W_s^N - W_{\lfloor s \rfloor}^N|^2) + O((s - \lfloor s \rfloor)^2) \\
&\quad + (L_s^N - L_{\lfloor s \rfloor}^N)^T \nabla V_i(X_{\lfloor s \rfloor}^N) \nabla f(X_{\lfloor s \rfloor}^N - L_{\lfloor s \rfloor}^N) d(W_i^N)_s + \text{ MG } + \text{ drift} \\
&= \int_0^t \mathcal{L}^N f(s) ds + \text{ MG}
\end{aligned}
\tag{10.9}
$$

where we have denoted by $\mathcal{L}^N f(s)$ the integrand of everything (including the drift term) on the right hand side excluding the martingale term. Similarly, let $\mathcal{L} f(s) = \frac{1}{2} \sum_{i=1}^r [D_{V_i}^2 f(\cdot - L_s)](X_s) + [D_{V_0} f(\cdot - L_s)](X_s)$. Our goal is therefore to show that under P,

$$f(X_t - L_t) - f(x_0) - \int_0^t \mathcal{L} f(s) ds$$

is a martingale for each $f \in C^2$. It suffices to show that $\forall A \subset \mathcal{F}_s$,

$$E^P[f(X_t - L_t) - f(X_s - L_s) - \int_s^t \mathcal{L} f(u) du, A] = 0 \tag{10.10}$$

Since the set of f satisfying (10.10) is closed under C^2, it suffices to prove (10.10) for $f \in C_c^\infty$. Since X and L are continuous, it suffices to prove (10.10) for $s < t$ where s

and t are M-dyadic rationals for some M. Finally, it suffices to show that

$$E^P[(f(X_t - L_t) - f(X_s - L_s) - \int_s^t \mathcal{L}f(u)du)g(X_{t_1}, L_{t_1}, \ldots, X_{t_n}, L_{t_n})] = 0$$

where $0 \leq t_1 \leq \ldots \leq t_n \leq s$ and each $g \in C_b((\mathbb{R}^d)^{2n}; \mathbb{R})$ as $1_A, A \in \mathcal{F}_s$ is the limit of such functions. As the term inside the expectation is a bounded and continuous function on (X, L)-pathspace, we have that

$$\begin{aligned} &E^{P^N}[(f(X_t - L_t) - f(X_s - L_s) - \int_s^t \mathcal{L}f(u)du)g(X_{t_1}, \ldots, L_{t_n})] \\ &\qquad \to E^P[(f(X_t - L_t) - f(X_s - L_s) - \int_s^t \mathcal{L}f(u)du)g(X_{t_1}, \ldots, L_{t_n})] \end{aligned}$$

For each N, we have by (10.9) that

$$E^{P^N}[(f(X_t - L_t) - f(X_s - L_s) - \int_s^t \mathcal{L}^N f(u)du)g(X_{t_1}, \ldots, L_{t_n})] = 0.$$

And so it remains to show that

$$E^{P^N}[(\int_s^t [\mathcal{L}f(u) - \mathcal{L}^N f(u)]du)g(X_{t_1}, \ldots, L_{t_n})] \to 0$$

Comparing the terms of $\mathcal{L}f(u)$ and $\mathcal{L}^N f(u)$, it is clear that we will be done after showing the following seven results:

$$\begin{aligned} &E^{P^N}[(\int_s^t \frac{1}{2}\sum_{i=1}^r [D^2_{V_i} f(\cdot - L_u)](X_u)du \\ &\qquad - \sum_{i=1}^r \int_s^t \sum_{j=1}^r V_i(X_u)^T D^2 f(X_{\lfloor u \rfloor} - L_{\lfloor u \rfloor})V_j(X_u)((W_j^N)_u - (W_j^N)_{\lfloor u \rfloor}) \\ &+ V_j(X_u)^T \nabla V_i(X_{\lfloor u \rfloor})\nabla f(X_{\lfloor u \rfloor} - L_{\lfloor u \rfloor})((W_j^N)_u - (W_j^N)_{\lfloor u \rfloor})d(W_i^N)_u)g(X_{t_1}, \ldots, L_{t_n})] \to 0 \end{aligned} \tag{10.11}$$

$$E^{P^N}[(\sum_{i=1}^{r}\int_s^t V_i(X_u)^T D^2 f(X_{\lfloor u\rfloor} - L_{\lfloor u\rfloor})V_0(X_u)(u-\lfloor u\rfloor)d(W_i^N)_u)g(X_{t_1},\ldots,L_{t_n})] \to 0 \tag{10.12}$$

$$E^{P^N}[(\sum_{i=1}^{r}\int_s^t V_0(X_u)^T \nabla V_i(X_{\lfloor u\rfloor})\nabla f(X_{\lfloor u\rfloor} - L_{\lfloor u\rfloor})(u-\lfloor u\rfloor)d(W_i^N)_u)g(X_{t_1},\ldots,L_{t_n})] \to 0 \tag{10.13}$$

$$E^{P^N}[(\int_s^t [D_{V_0}f(\cdot - L_u)](X_u) - \int_s^t V_0(X_u)\cdot\nabla f(X_u - L_u)du)g(X_{t_1},\ldots,L_{t_n})] \to 0 \tag{10.14}$$

$$E^{P^N}[(\sum_{i=1}^{r}\int_s^t (L_u - L_{\lfloor u\rfloor})^T \nabla V_i(X_{\lfloor u\rfloor})\nabla f(X_{\lfloor u\rfloor} - L_{\lfloor u\rfloor})d(W_i^N)_u)g(X_{t_1},\ldots,L_{t_n})] \to 0 \tag{10.15}$$

$$E^{P^N}[(\sum_{i=1}^{r}\int_s^t O(|W_u^N - W_{\lfloor u\rfloor}^N|^2)d(W_i^N)_u)g(X_{t_1},\ldots,L_{t_n})] \to 0 \tag{10.16}$$

$$E^{P^N}[(\sum_{i=1}^{r}\int_s^t O((u-\lfloor u\rfloor)^2)d(W_i^N)_u)g(X_{t_1},\ldots,L_{t_n})] \to 0 \tag{10.17}$$

But we have all these from our lemmas and the fact that all of the functions appearing in (10.11)-(10.17) are bounded. Result (10.14) is immediate as the left hand side is equal to 0 independent of N. Result (10.11) follows from Lemma 10.0.12. Result (10.15) follows from Lemma 10.0.8. Result (10.16) follows from Lemma 10.0.10 and results (10.12), (10.13), and (10.17) follow from variants on the proof of Lemma 10.0.10. And so, (10.3) is proved.

We now turn our attention to (10.4). We proceed as in the proof of (10.3): For each $f \in C_c^\infty(\mathbb{R}^d)$ such that $\frac{\partial f}{\partial \nu}(x) \geq 0$, $\forall x \in \partial \mathcal{O}$, $\forall \nu \in \nu(x)$,

$$
\begin{aligned}
&f(X_t^N) - f(x_0) \\
&= \sum_{i=1}^{r} \int_0^t V_i(X_s^N) \cdot \nabla f(X_s^N) d(W_i^N)_s + \underbrace{\int_0^t V_0(X_s^N) \cdot \nabla f(X_s^N) ds}_{\text{drift}} \\
&\quad + \int_0^t \nabla f(X_s^N) \cdot \nu(X_s^N) d|L^N|_s \\
&= \sum_{i=1}^{r} \int_0^t [V_i(X_s^N) \cdot \nabla f(X_s^N) - V_i(X_{\lfloor s \rfloor}^N) \cdot \nabla f(X_{\lfloor s \rfloor}^N)] d(W_i^N)_s \\
&\quad + \underbrace{\underbrace{\sum_{i=1}^{r} \int_0^t V_i(X_s^N) \cdot \nabla f(X_{\lfloor s \rfloor}^N) d(W_i^N)_s}_{\text{MG}} + \underbrace{\int_0^t \nabla f(X_s^N) \cdot \nu(X_s^N) d|L^N|_s}_{\geq 0}}_{\text{sub MG}} + \text{drift}
\end{aligned}
$$

The rest of the argument is more or less the same as the argument used to prove (10.3) only with some equalities replaced by inequalities to reflect the fact we are dealing with sub martingales and with f evaluated at X instead of $X - L$. In particular we again make use of Lemmas 10.0.8, 10.0.10, and 10.0.12.

Chapter 11

Step 4: Final Argument

In this chapter we finish our argument by showing that the processes X_t and L_t under P are a weak solution to the appropriate Stratonovich reflected SDE. By the uniqueness result of Lions and Sznitman (Theorem 3.1 of [10]) it will follow that our sequence of measures P^N converges weakly to P.

Let P be any limiting measure of the sequence of measures P^N on (X, L)-pathspace. We have shown that for all $f \in C^2(\mathbb{R}^d)$ we have that under P,

$$f(X_t - L_t) - f(x_0) - \int_0^t \mathcal{L}f(s)ds \text{ is a } P \text{ martingale} \tag{11.1}$$

and that for all $f \in C^2(\mathbb{R}^d)$ such that $\frac{\partial f}{\partial \nu} \geq 0$, $\forall x \in \partial \mathcal{O}$, $\forall \nu \in \nu(x)$, we have that

$$f(X_t) - f(x_0) - \int_0^t \tilde{\mathcal{L}}f(s)ds \text{ is a } P \text{ sub-martingale} \tag{11.2}$$

where

$$\mathcal{L}f(s) = \frac{1}{2}\sum_{i=1}^{r}[D^2_{V_i}f(\cdot - L_s)](X_s) + [D_{V_0}f(\cdot - L_s)](X_s)$$

and

$$\tilde{\mathcal{L}}f(s) = \frac{1}{2}\sum_{i=1}^{r}[D^2_{V_i}f(\cdot)](X_s) + [D_{V_0}f(\cdot)](X_s)$$

Taking the coordinate funcitons $f(x) = x_i$ in (11.1), we see that

$$M_t := X_t - x_0 - \int_0^t \frac{1}{2}\sum_{i=1}^r [D_{V_i}V_i](X_s) + V_0(X_s)ds - L_t \text{ is a martingale}$$

Taking the functions $f(x) = x_i x_j$ in (11.1), we see that

$$d < M_i, M_j > (t) = (\sigma(X_t)\sigma(X_t)^T)_{ij}dt$$

It therefore follows by standard techniques (c.f. Section 4.5 in [15]) that, extending our probablity space if neccesary, there exists an r-dimensional Brownian motion B_t such that

$$X_t - x_0 - \int_0^t \frac{1}{2}\sum_{i=1}^r [D_{V_i}V_i](X_s) + V_0(X_s)ds - L_t = \int_0^t \sigma(X_s)dB_s. \tag{11.3}$$

Specifically, we can extend our probability space $((C([0,\infty);\mathbb{R}^d))^2, P)$ to $((C([0,\infty);\mathbb{R}^d))^2 \times C([0,\infty);\mathbb{R}^d), \tilde{P})$, where the marginal distribution of $\tilde{P}$ in the appended space is Wiener measure.

Remark 11.0.14. Note that (11.3) can be written in "Stratonovich form" as

$$dX_t = \sigma(X_t) \circ dB_t + b(X_t)dt + L_t, \quad X_0 = x_0$$

Applying Ito's formula to (11.3) we have that for $f \in C_b^2(\mathbb{R}^d)$ that

$$f(X_t) - f(x_0) - \int_0^t \tilde{\mathcal{L}}f(s)ds - \xi_f(t) \text{ is a } \tilde{P}\text{-martingale}, \tag{11.4}$$

where

$$\xi_f(t) := \int_0^t \nabla f(X_s) \cdot dL_s.$$

It is clear that

$$L_t = (\xi_{x_1}(t), \ldots, \xi_{x_n}(t))$$

and

$$d\xi_f(t) = \sum_i \frac{\partial f}{\partial x_i}(X_t) d\xi_{x_i}(t). \tag{11.5}$$

We will need the following lemma:

Lemma 11.0.15. *L_t is a bounded variation process.*

Proof 11.0.16. Fix $T > 0$. As L_t is continuous, we have that

$$|L|_T = \lim_M g_M(L_t, T)$$

where

$$g_M(L_t, T) := \sum_{m=1}^{2^M} |L_{Tm2^{-M}} - L_{T(m-1)2^{-M}}|$$

is a non-decreasing sequence of positive functions on $C([0,\infty), \mathbb{R}^d)$. We know that

$$g_M(L_t^N, T) \leq |L^N|_T \leq T^\gamma \mathcal{L}_\gamma^N$$

and so it follows that

$$E[g_M(L_t^N, T)] \leq CT^\gamma$$

For each constant $K > 0$ we have that $g_M \wedge K$ is a bounded continuous function on $C([0,\infty), \mathbb{R}^n)$ and so

$$E^{P^N}[g_M(L_t, T) \wedge K] \to E^P[g_M(L_t, T) \wedge K]$$

whence it follows that

$$E^P[g_M(L_t, T) \wedge K] \leq CT^\gamma$$

Then letting $K \to \infty$ we have, by the monotone convergence theorem, that

$$E^P[g_M(L_t, T)] \leq CT^\gamma$$

Finally, letting $M \to \infty$, we apply the monotone convergence theorem again to see that

$$E^P[|L|_T] \leq CT^\gamma < \infty$$

from which the lemma follows.

Combining the above lemma with (11.5), it follows that $\xi_f(t)$ is a bounded variation process as well. Comparing (11.2) with (11.4), we have by the uniqueness of the Doob-Meyer martingale decomposition that for all $f \in C_b^2(\mathbb{R}^d)$ such that $\frac{\partial f}{\partial \nu} \geq 0$, $\forall x \in \partial \mathcal{O}$, $\forall \nu \in \nu(x)$,

$$\xi_f(t) \text{ is a non-decreasing function, } \tilde{P}\text{-a.s.}$$

Before proving the next theorem, we will introduce the following two lemmas from [14]. We remark that the proof of these lemmas carries over to our setting with the function ϕ in their setup replaced with the function ϕ from our setup.

Lemma 11.0.17. *(Lemma 2.3 in [14]) For all $f \in C_b^2(\mathbb{R}^d)$ we have that off of a set of $\tilde{P}$-measure* 0,

$$\int_s^t 1_{\{X_u \in \mathcal{O}\}} d\xi_\phi(u) = 0, \quad \forall s < t.$$

Lemma 11.0.18. *(Lemma 2.5 in [14]) Let U be a neighborhood of a point $x \in \partial\mathcal{O}$ and suppose that $f \in C_b^2(\mathbb{R}^d)$ is such that $\frac{\partial f}{\partial \nu} \geq 0$, $\forall x \in U \cap \partial\mathcal{O}$, $\forall \nu \in \nu(x)$. Then, off of a set of $\tilde{P}$-measure* 0,

$$\int_s^t 1_{\{X_u \in U\}} d\xi_\phi(u) \geq 0, \quad \forall s < t.$$

We present a key theorem:

Theorem 11.0.19.

$$L_t = \int_0^t \nu(X_s) d|L|_s, \quad \tilde{P}\text{-a.s.}$$

That is, if we let $\mu(t) := \frac{dL_t}{d|L|_t}$, then $\mu(t) \in \nu(X_t)$ for all t.

The proof of this theorem follows the proof of Theorem 2.4 in [14]. First we will need the following lemmas.

Lemma 11.0.20. *The set valued mapping $x \in \partial\mathcal{O} \longmapsto \nu(x)$ is continuous in the sense that if $x_n \in \partial\mathcal{O}$ is a sequence converging to $x \in \partial\mathcal{O}$ and $\nu_n \in \nu(x_n)$ is a sequence converging to ν then $\nu \in \nu(x)$.*

Proof 11.0.21. As $\mathcal{O}$ is admissible, we have by Part 1a) of Definition 7.1.1 that for each n,

$$(x' - x_n) \cdot \nu_n + C_0|x' - x_n|^2 \geq 0, \quad \forall x' \in \mathcal{O}$$

Taking the limit as n tends to infinity we get that

$$(x' - x) \cdot \nu + C_0|x' - x|^2 \geq 0, \quad \forall x' \in \mathcal{O}$$

It then follows by Part 1b) of Definition 7.1.1 that $\nu \in \nu(x)$.

Lemma 11.0.22. *For $f \in C_b^2(\mathbb{R}^d)$, let $a : \partial\mathcal{O} \to \mathbb{R}$ be the function defined by*

$$a(x) := \inf_{\nu \in \nu(x)} \left(\frac{\nabla f(x) \cdot \nu}{\nabla \phi(x) \cdot \nu} \right)$$

and let $b : \partial\mathcal{O} \to \mathbb{R}$ be the function defined by

$$b(x) := \sup_{\nu \in \nu(x)} \left(\frac{\nabla f(x) \cdot \nu}{\nabla \phi(x) \cdot \nu} \right)$$

(note that we have suppressed the dependence of a and b on f). Then $a(x)$ is lower semi-continuous and $b(x)$ is upper semi-continuous on $\partial\mathcal{O}$.

Proof 11.0.23. We will just show that $a(x)$ is lower semi-continuous as the corresponding proof for $b(x)$ is analogous. Let $x_n \in \partial\mathcal{O}$ be a sequence converging to $x \in \partial\mathcal{O}$. For each n there exists a ν_n such that

$$a(x_n) > \frac{\nabla f(x_n) \cdot \nu_n}{\nabla \phi(x_n) \cdot \nu_n} - \frac{1}{n}$$

It therefore follows that

$$\liminf_{n \to \infty} a(x_n) \geq \liminf_{n \to \infty} \frac{\nabla f(x_n) \cdot \nu_n}{\nabla \phi(x_n) \cdot \nu_n} \geq a(x)$$

where the last inequality follows from Lemma 11.0.20.

Lemma 11.0.24. *For each $f \in C^2(\mathbb{R}^d)$, $d\xi_f(t)$ is absolutely continuous with respect to $d\xi_\phi(t)$ and the Radon-Nikodym derivative, $\frac{d\xi_f}{d\xi_\phi}(t)$, satisfies*

$$\inf_{\nu \in \nu(X_t)} \left(\frac{\nabla f(X_t) \cdot \nu}{\nabla \phi(X_t) \cdot \nu} \right) \leq \frac{d\xi_f}{d\xi_\phi}(t) \leq \sup_{\nu \in \nu(X_t)} \left(\frac{\nabla f(X_t) \cdot \nu}{\nabla \phi(X_t) \cdot \nu} \right). \tag{11.6}$$

Proof 11.0.25. Fix $f \in C^2(\mathbb{R}^d)$ and note that $f \longmapsto \xi_f$ is linear. Letting $\bar{f} = f + \lambda\phi$ then, as $\mathcal{O}$ is bounded, for sufficiently large λ, $\frac{\partial \bar{f}}{\partial \nu} \geq 0, \forall x \in \partial\mathcal{O}$, $\forall \nu \in \nu(x)$. It therefore follows that

$$0 \leq d\xi_{\bar{f}} = d\xi_f + \lambda d\xi_\phi$$

Similarly, letting $\bar{f} = -f + \lambda\phi$ we find that

$$0 \geq \lambda d\xi_\phi - d\xi_f$$

Combining these we have that

$$-\lambda d\xi_\phi \leq d\xi_f \leq \lambda d\xi_\phi$$

and so $d\xi_f(t)$ is absolutely continuous with respect to $d\xi_\phi(t)$.

We next prove the left inequality of (11.6). By Lemma 11.0.22, for each $x \in \partial\mathcal{O}$, and $\varepsilon > 0$, there exists a sufficently small neighborhood U of x such that $\forall y \in U \cap \partial\mathcal{O}$, $\forall \nu \in \nu(y)$,

$$a(x) - \varepsilon \leq \frac{\nu \cdot \nabla f(y)}{\nu \cdot \nabla \phi(y)} \tag{11.7}$$

From this it follows that $\forall y \in U \cap \partial\mathcal{O}$, $\forall \nu \in \nu(y)$,

$$(a(x) - \varepsilon)\frac{\partial \phi}{\partial \nu}(y) \leq \frac{\partial f}{\partial \nu}(y)$$

and this in turn implies that, by Lemma 11.0.18, on a set Γ_1 such that $\tilde{P}(\Gamma_1) = 1$, we have that for all $s < t$,

$$(a(x) - \varepsilon) \int_s^t 1_{\{X_u \in U\}} d\xi_\phi(u) \leq \int_s^t 1_{\{X_u \in U\}} d\xi_f(u),$$

and so

$$(a(x) - \varepsilon) 1_{\{X_u \in U\}} \leq 1_{\{X_u \in U\}} \frac{d\xi_f}{d\xi_\phi}(u), \quad \tilde{P} \times d\xi_\phi\text{-a.s}$$

Now for each k, let $\varepsilon_k = \frac{1}{k}$. For each k, by the compactness of $\partial\mathcal{O}$, there exist a finite number of points $x_{j,k}$ with corresponding neighborhoods $U_{j,k}$ chosen to satisfy (11.7) which cover $\partial\mathcal{O}$. We can and do take these neighborhoods such that $\mathrm{diam}(U_{j,k}) < \frac{1}{k}$. As there are a countable number of sets $U_{j,k}$ we can find a single set Γ_2 with $d\xi_\phi(\Gamma_2) = 1$ such that

$$(a(x_{j,k}) - \frac{1}{k}) 1_{\{X_u \in U_{j,k}\}} \leq 1_{\{X_u \in U_{j,k}\}} \frac{d\xi_f}{d\xi_\phi}(u), \quad \forall j, k \text{ on } \Gamma_1 \times \Gamma_2$$

Furthermore, by Lemma 11.0.17 we may take Γ_2 such that $X_u \in \partial\mathcal{O}, \quad \forall u \in \Gamma_2$. Now fix $u^* \in \Gamma_2$. For each k, there is some $j(k)$ such that $X_{u^*} \in U_{j(k),k} \cap \partial\mathcal{O}$ and so it follows that

$$(a(x_{j(k),k}) - \frac{1}{k}) \leq \frac{d\xi_f}{d\xi_\phi}(u^*)$$

letting $k \to \infty$, by the lower semi-continuity of $a(x)$ and the fact that $x_{j(k),k} \to X_{u^*}$, we have that

$$a(X_{u^*}) \leq \frac{d\xi_f}{d\xi_\phi}(u^*)$$

As u^* was an arbitrary element of the $d\xi_\phi$-full set Γ_2, we have proved the left inequality of (11.6). The proof of the right inequality is analogous.

We will now present a proof of Theorem 11.0.19.

Proof 11.0.26. Following Stroock and Varadhan in [14], we will use the notation $d\xi_0(t)$ for $d|L|_t$. First note that by (11.5), for all $f \in C_b^2$ we have that

$$d\xi_f \ll d\xi_{x_1} + \ldots + d\xi_{x_n} \ll d\xi_0$$

and since, by Lemma 11.0.24, $d\xi_{x_i} \ll d\xi_\phi$, we have that

$$d\xi_0 \ll d\xi_\phi$$

Therefore, $d\xi_0$-a.s. we have that

$$\nabla f(X_s) \cdot \mu(s) = \frac{d\xi_f}{d\xi_\phi}(s)[\nabla \phi(X_s) \cdot \mu(s)]. \tag{11.8}$$

Indeed, using (11.5), we can reduce both sides to $\frac{d\xi_f}{d\xi_0}(s)$.

We first consider the case that $\nabla \phi(X_s) \cdot \mu(s) = \beta \geq 0$. Combining (11.6) and (11.8), we have that

$$\nabla f(X_s) \cdot \mu(s) \geq \beta \inf_{\nu \in \nu(X_s)} \left(\frac{\nabla f(X_s) \cdot \nu}{\nabla \phi(X_s) \cdot \nu} \right)$$

here, s and X_s are fixed, so since the above only depends on f through its gradient, we have in fact that

$$v \cdot \mu(s) \geq \beta \inf_{\nu \in \nu(X_s)} \left(\frac{v \cdot \nu}{\nabla \phi(X_s) \cdot \nu} \right)$$

for each vector $v \in \mathbb{R}^d$. Taking v to be $(x' - X_s)$, we then have that

$$\begin{aligned}
(x' - X_s) \cdot \mu(s) &\geq \beta \inf_{\nu \in \nu(X_s)} \left(\frac{(x' - X_s) \cdot \nu}{\nabla \phi(X_s) \cdot \nu} \right) && (11.9)\\
&\geq \beta \inf_{\nu \in \nu(X_s)} \left(\frac{-C_0 |x' - X_s|^2}{\nabla \phi(X_s) \cdot \nu} \right) && (11.10)\\
&\geq \frac{-C_0 \beta}{\alpha} |x' - X_s|^2 && (11.11)
\end{aligned}$$

and so

$$\nabla (x' - X_s) \cdot \mu(s) + \frac{C_0 \beta}{\alpha} |x' - X_s|^2 \geq 0$$

so, as $\mathcal{O}$ is admissible, it follows from Part 1b) of Definition 7.1.1 that $\frac{\beta}{\alpha}\mu(s)$ and hence $\mu(s)$ is a positive multiple of some $\nu \in \nu(X_s)$. Since $\mu(s)$ has norm 1 it follows that $\mu(s) \in \nu(X_s)$.

In the case where $\nabla \phi(X_s) \cdot \mu(s) = -\beta \leq 0$, applying the above to $-\mu(s)$ we see that

$$-\mu(s) \in \nu(X_s)$$

So we have that for each s, that either

$$\mu(s) \in \nu(X_s) \text{ or } -\mu(s) \in \nu(X_s)$$

We claim it is the former. Since $\frac{\partial\phi}{\partial\nu}(x) \geq 0, \forall x \in \mathcal{O}, \forall \nu \in \nu(x)$, we have that $d\xi_\phi(s)$ is a positive measure $\tilde{P}$-a.s. Noting that

$$\begin{aligned} d\xi_\phi &= \sum \frac{\partial\phi}{\partial x_i} d\xi_{x_i} \\ &= [\nabla\phi(X_s) \cdot \mu(s)] d\xi_0 \end{aligned}$$

we have (c.f. Part 2 of Definition 7.1.1) that $\mu(s) \in \nu(X_s)$, $d\xi_0$-a.s. And so Theorem 11.0.19 is proved.

It is clear that as P is the weak limit of (a subsequence of) the P^N and since $X_t \in \bar{\mathcal{O}}, \forall t$ under P^N, it follows that $X_t \in \bar{\mathcal{O}}, \forall t$ under P and hence under $\tilde{P}$.

So in summary, we have shown that under $\tilde{P}$, we have a weak solution to the reflected SDE:

$$X_t = x_0 + \int_0^t \sigma(X_s) \circ dW_s + \int_0^t b(X_s) ds + L_t, \tag{11.12}$$

$$|L|_t = \int_0^t 1_{\{X_s \in \partial\mathcal{O}\}} d|L|_s, \text{ and } L_t = \int_0^t \nu(X_s) d|L|_s \tag{11.13}$$

where X_t is a continuous process lying in $\bar{\mathcal{O}}$ and L_t is a continuous bounded variation process.

In Theorem 3.1 of [10] states that if H is the Frechet space of continuous adapted processes X whose semi-norms

$$||X||_t := E[\sup_{0 \leq s \leq t} |X_s^4|]^{\frac{1}{4}}$$

are finite, then there is a unique $X \in H$ which satisfies (11.12). As $\mathcal{O}$ is bounded, under $\tilde{P}$, X is an element of H which satisfies (11.12). Therefore we have that for each limit point P of the measures P^N, the distribution of X under P is the same unique element of H which satisfies (11.12). It therefore follows that the distributions of X under P^N converge weakly to the unique weak solution X of (11.12).

Chapter 12

Observations and Applications

In this chapter we record some observations and applications of our result. We have shown that for a bounded and admissible set $\mathcal{O}$, under suitable regularity conditions, if (X^N, L^N) are solutions to the reflected SDE

$$dX_t^N = \sigma(X_t^N)dW_t^N + b(X_t^N)dt + dL_t^N$$

and (X, L) is the solution to the Stratonovich reflected SDE

$$dX_t = \sigma(X_t) \circ dW_t + b(X_t)dt + dL_t$$

then X_t^N converges to X_t in distribution.

It is interesting to note that while the L_t^N and L_t are bounded in variation uniformly almost surely (see the proof of Lemma 11.0.15), L_t^N does *not* converge to L_t in variation almost surely. We can see this even in the simple case where $\mathcal{O}$ is the half-line:

Example 12.0.27. Let $d = 1$, $\mathcal{O} = \mathbb{R}^+$, $b \equiv 0$, and $\sigma \equiv 1$, i.e. (X^N, L^N) are solutions to the reflected SDE

$$dX_t^N = dW_t^N + dL_t^N$$

and (X, L) is the solution to the reflected SDE

$$dX_t = dW_t + dL_t$$

(that is, X_t is reflected Brownian motion in the half-line). In this case, for a.e. ω, $L_t^N(\omega)$ does not converge to $L_t(\omega)$ in variation.

Proof 12.0.28. We fix ω and suppress the dependence of L_t^N and L_t on ω. It is well known that

$$L_t^N = \sup_{0 \le s \le t} [-W_s^N] \text{ and } L_t = \sup_{0 \le s \le t} [-W_s]$$

It is clear that L_t^N is piecewise linear, and so dL_t^N is absolutely continuous with respect to Lebesgue measure dt for each N. If L_t^N did converge to L_t in variation then dL_t would be absolutely continuous with respect to dt as well.

But dL_t is almost surely singular to dt: L_t is the local time of X_t at 0. Another representation of reflected Brownian motion on the half-line is $\tilde{X}(t) = |W_t|$. The local time of $\tilde{X}_t$ at 0 is clearly the same as the local time of W_t at 0 and this is well known to be almost surely singular with respect to Lebesgue measure. As X_t and $\tilde{X}_t$ are representations of the same process we are done.

The main application of our result that we consider is the following: Suppose that for each N, the paths X_t^N satisfy a certain geometric property almost surely and the set S of paths which satisfy this geometric property is closed in $C([0, \infty); \mathbb{R}^d)$. It then follows that the paths of X_t also satisfy this geometric property almost surely since

$$P(S) \ge \limsup_{N \to \infty} P^N(S) = 1$$

We illustrate this with some examples.

Example 12.0.29. In $\mathbb{R}^2$, let $\mathcal{O}$ be the rectangle $[-1, 1] \times [0, 2]$ and consider the Stratonovich reflected SDE

$$dX_t = \sigma(X_t) \circ dW_t + dL_t, \quad X_0 = x_0,$$

where $\sigma(x) = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}$. Then

$$\text{If } |x_0| > 1, \quad |X_t| \leq |x_0| \text{ for } t > 0 \text{ a.s.} \tag{12.1}$$

and

$$\text{If } |x_0| < 1, \quad |X_t| \geq |x_0| \text{ for } t > 0 \text{ a.s.} \tag{12.2}$$

Proof 12.0.30. Note that the vector field $\sigma(x)$ points counter-clockwise around the origin. For a fixed ω and N, $X_t^N(\omega)$ will simply be a path which moves along arcs of constant radius until it hits the boundary of the rectangle where it is nudged back inside. It is clear that $X_t^N(\omega)$ satisfies (12.1) and (12.2) since when $|x_0| > 1$, the path will be nudged to a lower radius and when $|x_0| < 1$, the path will be nudged to a higher radius (See Figure 12-1).

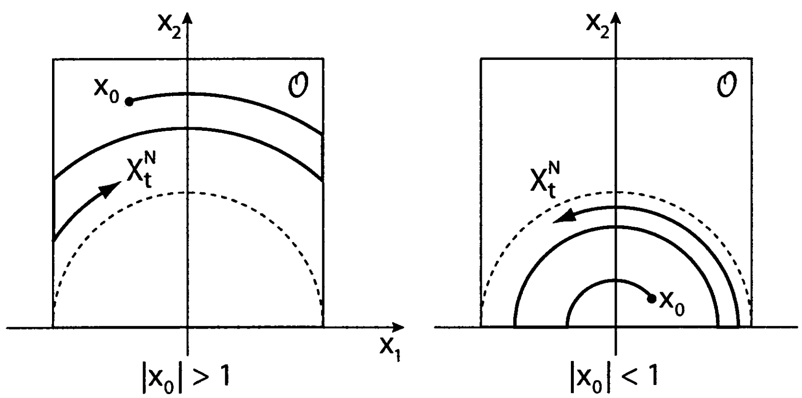

Figure 12-1:

It is clear that the set of paths satisfying each of (12.1) and (12.2) is closed in $C([0,\infty);\mathbb{R}^d)$. And so it follows that X_t satisfies (12.1) and (12.2).

We next consider coupled reflected Brownian motion. We will first need the following theorem.

Theorem 12.0.31. *Suppose $\mathcal{O}$ is bounded and admissible. Then $\mathcal{O} \times \mathcal{O}$ is bounded and admissible as well, where for each $(x, y) \in \partial(\mathcal{O} \times \mathcal{O})$ we take the set of normal* *vectors $\nu(x,y)$ to be the set of vectors*

$$\left\{ \begin{pmatrix} a_1\nu_x \\ a_2\nu_y \end{pmatrix} : \nu_x \in \nu(x), \nu_y \in \nu(y), a_1^2 + a_2^2 = 1 \right\}, \ \textit{when } (x,y) \in \partial\mathcal{O} \times \partial\mathcal{O},$$

$$\left\{ \begin{pmatrix} \nu_x \\ 0 \end{pmatrix} : \nu_x \in \nu(x) \right\}, \ \textit{when } (x,y) \in \partial\mathcal{O} \times \mathcal{O},$$

and

$$\left\{ \begin{pmatrix} 0 \\ \nu_y \end{pmatrix} : \nu_y \in \nu(y) \right\}, \ \textit{when } (x,y) \in \mathcal{O} \times \partial\mathcal{O}.$$

Proof 12.0.32. In the proof below we will only consider boundary points $(x,y) \in \partial\mathcal{O} \times \partial\mathcal{O}$. The other two cases are simpler and the proofs for them are analogous.

We first show that Part 1 of Definition 7.1.1 holds. Note that for each $x', y' \in \mathcal{O}$ and $\nu \in \nu(x,y)$,

$$\left[\begin{pmatrix} x' \\ y' \end{pmatrix} - \begin{pmatrix} x \\ y \end{pmatrix} \right] \cdot \nu + C_0 \left| \begin{pmatrix} x' \\ y' \end{pmatrix} - \begin{pmatrix} x \\ y \end{pmatrix} \right|^2$$
$$= \alpha_1(x'-x) \cdot \nu_x + \alpha_2(y'-y) \cdot \nu_y + C_0|x-x'|^2 + C_0|y-y'|^2 \geq 0,$$

and so Part 1a) holds. Next, suppose that for some $v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ and $C > 0$ we have that

$$\left[\begin{pmatrix} x' \\ y' \end{pmatrix} - \begin{pmatrix} x \\ y \end{pmatrix} \right] \cdot v + C_0 \left| \begin{pmatrix} x' \\ y' \end{pmatrix} - \begin{pmatrix} x \\ y \end{pmatrix} \right|^2$$
$$= \alpha_1(x'-x) \cdot v_1 + \alpha_2(y'-y) \cdot v_2 + C|x-x'|^2 + C|y-y'|^2 \geq 0.$$

Then, taking $y' \to y$ we see that

$$\alpha_1(x'-x) \cdot v_1 + C|x-x'|^2 \geq 0.$$

and so, since $\mathcal{O}$ is admissible, $v_1 = k_1\nu_x$ for some $\nu_x \in \nu(x)$, $k_1 > 0$. A similar argument gives that $v_2 = k_2\nu_y$ for some $\nu_y \in \nu(y)$, $k_2 > 0$. It is then easy to show that $\frac{v}{|v|} \in \nu(x,y)$ and so Part 1b) holds.

We next show that Part 2 holds. As $\mathcal{O}$ is bounded, ϕ is bounded in $\mathcal{O}$ and so after adding a constant to ϕ if necessary, we may assume that $\phi \geq 1$ in $\mathcal{O}$.

For $\mathcal{O} \times \mathcal{O}$, we define $\Phi(x,y) := \phi(x)\phi(y)$. Then $\nabla\Phi(x,y) = \begin{pmatrix} \phi(y)\nabla\phi(x) \\ \phi(x)\nabla\phi(y) \end{pmatrix}$ and for all $(x,y) \in \partial\mathcal{O} \times \partial\mathcal{O}, \quad \nu \in \nu(x,y)$, we have that

$$\begin{aligned} \nabla\Phi(x,y) \cdot \nu =& \alpha_1\phi(y)\nabla\phi(x) \cdot \nu_x + \alpha_2\phi(x)\nabla\phi(y) \cdot \nu_y \\ \geq& \alpha_1\phi(y)\alpha + \alpha_2\phi(x)\alpha \\ \geq& \alpha(\alpha_1 + \alpha_2) \geq \alpha \end{aligned}$$

and so Part 2 holds with the function $\Phi(x,y)$. Finally, as $\mathcal{O} \times \mathcal{O}$ is bounded, we get that Part 3 holds for free via Theorem 7.1.2.

We now discuss coupled reflected Brownian motion. A d-dimensional coupled reflected Brownian motion is a $2d$-dimensional process X_t in a product domain $\mathcal{O} \times \mathcal{O}$ which satisfies the reflected SDE

$$dX_t = \sigma(X_t)dW_t + dL_t,$$

where

$$\sigma(x) \equiv \begin{pmatrix} I \\ I \end{pmatrix}.$$

Note that as σ is constant, the Stratonovich and Ito versions of the above SDE coincide. We will express this reflected SDE in a more convenient form as the pair of reflected SDEs

$$\begin{aligned} dX_t =& dW_t + dL_t, X_0 = x_0 \\ dY_t =& dW_t + dM_t, Y_0 = y_0 \end{aligned}$$

where, abusing notation, X_t and Y_t denote the first and second d coordinates of the process. We think of X_t and Y_t as being two d-dimensional processes which are driven by the same Brownian motion W_t and which are constrained to lie in the same domain $\mathcal{O}$. The two processes move in sync except for where either process bumps into the boundary and is nudged.

We now consider the geometric properties of coupled reflected Brownian motion in two domains. Such properties were used to prove the "hot spots conjecture" for these domains (See [2] and [1] for more details).

Example 12.0.33. We first consider the case where our domain $\mathcal{O} \subset \mathbb{R}^2$ is the obtuse triangle lying with its longest face on the horizontal axis and denote its left and right acute angles by α and β. Suppose $x_0 \neq y_0$, and let Θ_t be the angle of the vector from X_t to Y_t. Then, almost surely,

$$\text{If } -\beta \leq \Theta_0 \leq \alpha, \text{ then } \forall t, \text{ either } -\beta \leq \Theta_t \leq \alpha \text{ or } X_t = Y_t. \tag{12.3}$$

Proof 12.0.34. Consider the approximating stochastic ODE

$$\begin{aligned} dX_t^N =& dW_t^N + dL_t^N, X_0^N = x_0 \\ dY_t^N =& dW_t^N + dM_t^N, Y_0^N = y_0 \end{aligned}$$

Since, by Theorem 12.0.31, $\mathcal{O} \times \mathcal{O}$ is bounded and admissible, we have that (X_t^N, Y_t^N) converges to (X_t, Y_t) in distribution. As the set of paths satisfying (12.3) is closed, it suffices to show that (12.3) holds for X_t^N and Y_t^N for each N.

Some thought shows that this is true: Fix N and consider an N-dyadic time interval $[(m+1)2^{-N}, m2^{-N}]$. Within this interval, X_t^N and Y_t^N will attempt to travel along the vector $W_{(m+1)2^{-N}} - W_{m2^{-N}}$. Pushing against the boundary kills the component of motion perpendicular to the wall. And so at the end of the N-dyadic interval, the processes will have moved to the same location they would have gone to if they were allowed to leave the domain but were then projected back (See Figure 12-2).

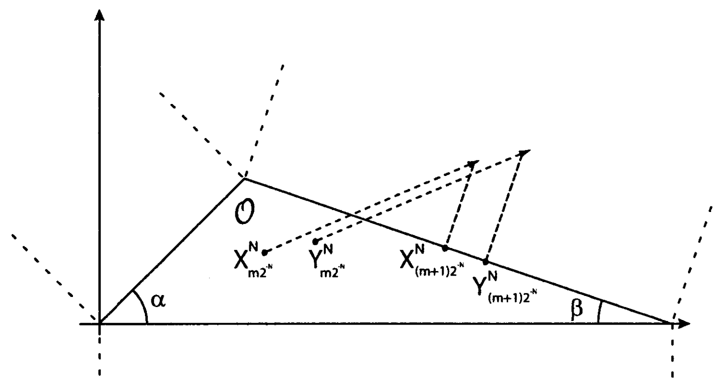

Figure 12-2:

Note that if either process would have ended up in one of the "exterior cones", it is projected to the corresponding corner.

Now given any two points x and y in $\mathbb{R}^2$, if we define by $\pi(x)$ and $\pi(y)$ their projections onto the triangle, some thought shows that whenever the angle between x and y lies in the interval $[-\beta, \alpha]$ so too will the angle between $\pi(x)$ and $\pi(y)$, unless $\pi(x) = \pi(y)$. It follows that X_t^N and Y_t^N satisfy (12.3) on each N-dyadic interval and therefore satisfy (12.3).

Example 12.0.35. We now consider coupled reflected Brownian motion in a Lip domain. A Lip domain is a Lipschitz domain in $\mathbb{R}^2$ which is bounded below by a function $f_1(x)$ and above by another function $f_2(x)$ each of which is Lipschitz continuous with constant ≤ 1. The domains are so named because they look like a pair of lips (See Figure 12-3).

Consider coupled reflected Brownian motion in a lip domain $\mathcal{O}$ where the defining functions $f_1(x)$ and $f_2(x)$ are smooth. Then $\mathcal{O}$ is a bounded admissible domain. Recall the definition of Θ_t from the previous example. We have the following geometric property almost surely for the paths X_t and Y_t:

$$\text{If } -\frac{\pi}{4} \leq \Theta_0 \leq \frac{\pi}{4}, \text{ then } \forall t, \text{ either } -\frac{\pi}{4} \leq \Theta_t \leq \frac{\pi}{4} \text{ or } X_t = Y_t. \tag{12.4}$$

Proof 12.0.36. Again we consider the approximating reflected ODE and its solutions X_t^N and Y_t^N. We will take advantage of the fact that these approximate processes have

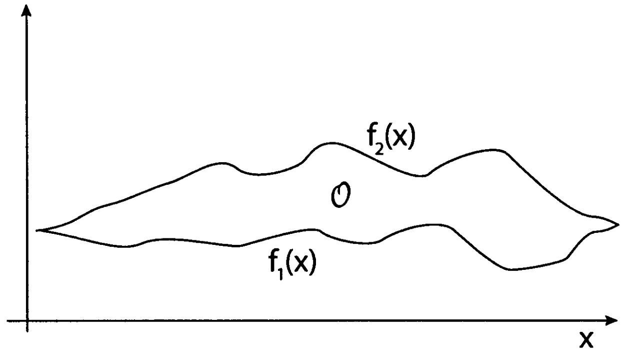

Figure 12-3:

well defined notions of velocity. It again suffices to prove (12.4) for these approximate processes. The fact that the Lipschitz constant is between -1 and 1 for f_1 and f_2 means that the "ceiling" and "floor" each slope upward or downward at 45 ° at most.

Suppose X_t^N and Y_t^N do not satisfy (12.4). Let $T = \inf t : \Theta_t \notin [-\frac{\pi}{4}, \frac{\pi}{4}]$. As Θ_t only changes when one of X_t^N and Y_t^N is at the boundary, without loss of generality assume Y_T^N is at the boundary, and will be affected by the boundary (if not rotate the whole picture by 180 °) and that $\Theta_T = \frac{\pi}{4}$ (if not reflect the whole picture across the horizontal axis). T is in some N-dyadic interval of the form $[m2^{-N}, (m+1)2^{-N})$ and so, if not for the boundary, both points would be traveling at a constant velocity $v = 2^N(W_{(m+1)2^{-N}} - W_{m2^{-N}})$. Instead, for Y_T^N, the presence of the boundary projects the velocity vector v onto the tangent space of the boundary.

It suffices to show that the presence of the boundary cannot increase Θ_t. Note first that it is impossible for the boundary to cross the line segment ℓ_t^N connecting X_t^N and Y_t^N as doing so would violate the condition on the Lipschitz constants for the boundary functions. It is possible that the ℓ_t^N is part of the boundary, and the set lies to the left of ℓ_t^N, but this case is uninteresting as the points X_t^N and Y_t^N will be affected by the boundary in the same way and this cannot change Θ_t. The interesting cases are depicted in Figure 12-4.

Some geometric thought shows that in either case depicted, any vector v which would drive Y_t^N into the boundary could only decrease Θ_t. And so we have reached a contradiction.

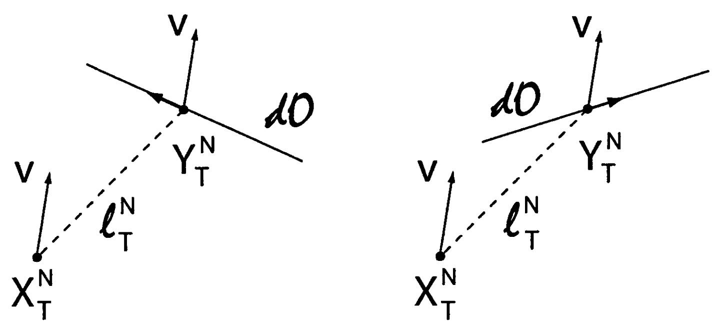

Figure 12-4:

Appendix A

Appendix for Viscosity Solutions

In this appendix we give a bare bones introduction to the theory of viscosity solutions. For a more complete introduction to viscosity solutions, I recommend either [4] or [5].

A.1 General Setting

We begin in the elliptic setting (which will in fact cover the parabolic setting once we think of t as just being the $(d+1)$st component of x, see below). A viscosity solution is a type of weak solution for a second order, non-linear PDE:

Definition A.1.1. Let $\Omega \subset \mathbb{R}^d$. Then $u(x) \in C(\Omega)$ is a *viscosity solution* for the PDE

$$F(x, u, Du, D^2u) = 0 \text{ in } \Omega \tag{A.1}$$

(where Du and D^2u represent the gradient and Hessian of u respectively) if the following two conditions hold:

1. For every $v \in C^\infty(\Omega)$ which touches u from below at x_0 (i.e. $u(x) \geq v(x), \forall x$ and $u(x_0) = v(x_0)$) we have that $F(x_0, v(x_0), Dv(x_0), D^2v(x_0)) \geq 0$.

2. For every $v \in C^\infty(\Omega)$ which touches u from above at x_0 (i.e. $u(x) \leq v(x), \forall x$

and $u(x_0) = v(x_0)$) we have that $F(x_0, v(x_0), Dv(x_0), D^2v(x_0)) \leq 0$.

A viscosity solution is clearly not a classical solution in general as we have not even assumed that u is differentiable. Nevertheless, we have made sense of (A.1) by moving the differentiation from u onto a smooth test function v. This is analogous to the more common notion of a distributional weak solution where derivatives are moved to smooth test functions using integration by parts. However, unlike distributional weak solutions, viscosity solution theory can be applied to non-linear differential equations.

Before proceeding further, we should stop and examine why viscosity solutions are our choice of weak solution. After all, the theory of viscosity solutions was developed to handle non-linear PDE and in this thesis we only consider linear PDE. Why not use distributional weak solutions? The main reason is that the notion of a viscosity solution is closely tied to maximum principles. Indeed, Nirenberg's proof [11] of the strong maximum principle for classical solutions extends without much change to cover viscosity solutions via arguments like those used in the proof of Theorem A.1.8. On the other hand, it is more complicated to work with maximum principles for distributional solutions as they are not even defined pointwise.

Returning to our discussion of viscosity solutions, we can now decompose the previous definition to define viscosity subsolutions and supersolutions. Let $USC(\Omega)$ and $LSC(\Omega)$ denote the space of upper semi-continuous and lower semi-continuous functions respectively.

Definition A.1.2. A function $u \in LSC(\Omega)$ is a *viscosity supersolution* of $F(x, u, Du, D^2u) = 0$ (alternatively "u solves $F(x, u, Du, D^2u) \geq 0$ in the viscosity sense") if condition 1. of Definition A.1.1 holds.

A function $u \in USC(\Omega)$ is a *viscosity subsolution* of $F(x, u, Du, D^2u) = 0$ (alternatively "u solves $F(x, u, Du, D^2u) \leq 0$ in the viscosity sense") if condition 2. of Definition A.1.1 holds.

Clearly, if u is both a viscosity subsolution and supersolution then it is a viscosity solution. In order for a viscosity solution to be a good notion for a weak solution we would hope that

(I) Classical solutions are viscosity solutions.

(II) Viscosity solutions are unique.

These statements do *not* hold for general F and so we restrict our focus to proper F:

Definition A.1.3. F is proper if

a) $F(x, r, p, X) \leq F(x, s, p, X)$ whenever $r \leq s$.

b) $F(x, r, p, X) \leq F(x, r, p, Y)$ whenever $Y \leq X$ (as symmetric non-negative definite matrices).

When F is proper it is easy to see that statement (I) holds.

Theorem A.1.4. *Let $u \in C^2(\Omega)$ be a classical sub/super solution to*

$$F(x, u, Du, D^2u) = 0$$

where F is proper. Then u is a viscosity sub/super solution as well.

Proof A.1.5. We handle the supersolution case (the subsolution case is analogous). Suppose $v \in C^\infty(\Omega)$ touches u from below at x_0. Then $v(x_0) = u(x_0)$, $Dv(x_0) = Du(x_0)$, and $D^2v(x_0) \leq D^2u(x_0)$. So by the properness of F,

$$F(x_0, v(x_0), Dv(x_0), D^2v(x_0)) \geq F(x_0, u(x_0), Du(x_0), D^2u(x_0)) \geq 0$$

Statement (II) for the Dirichlet problem follows from the following comparison principle (c.f. Theorem 3.3 in [5]):

Theorem A.1.6. *(Comparison principle): Let Ω be a bounded open subset of $\mathbb{R}^d$ and let F be proper and satisfy some additional regularity assumptions. Let $u \in USC(\bar{\Omega})$ be a subsolution and $v \in LSC(\bar{\Omega})$ be a supersolution of $F(x, u, Du, D^2u) = 0$ in Ω. Then*

$$u \leq v \text{ on } \partial\Omega \Longrightarrow u \leq v \text{ in } \bar{\Omega}$$

Corollary A.1.7. *(Uniqueness of viscosity solutions to the Dirichlet problem) If $u_1, u_2 \in C(\Omega)$ are solutions to*

$$\begin{cases} F(x, u, Du, D^2u) = 0 \text{ in } \Omega \\ u(x) = f(x) \text{ on } \partial\Omega \end{cases}$$

then $u_1 = u_2$.

The proof of the comparison principle is fairly involved and we won't present it here. In fact, for the purposes of this thesis, the additional regularity assumptions required for the comparison principle do not hold! Therefore we will instead use the following weaker comparison principle. The proof in this case is much easier and requires no additional regularity assumptions on F.

Theorem A.1.8. *(Comparison with smooth functions) Let Ω be a bounded open subset of $\mathbb{R}^d$ and let F be proper. If $u \in LSC(\bar{\Omega})$ is a viscosity supersolution of $F(x, u, Du, D^2u) = 0$ and $v \in C^\infty(\bar{\Omega})$ is a (classical) solution of $F(x, v, Dv, D^2v) < 0$, then*

$$u \leq v \text{ on } \partial\Omega \Longrightarrow u \leq v \text{ in } \bar{\Omega}$$

If $u \in USC(\bar{\Omega})$ is a viscosity subsolution of $F(x, u, Du, D^2u) = 0$ and $v \in C^\infty(\bar{\Omega})$ is a (classical) solution of $F(x, v, Dv, D^2v) > 0$, then

$$u \geq v \text{ on } \partial\Omega \Longrightarrow u \geq v \text{ in } \bar{\Omega}$$

Proof A.1.9. We will just prove the first statement as the proof of the second is analogous. Suppose not, then by the boundedness of Ω and the lower semi-continuity of u we know $\exists x_0 \in \Omega$ such that

$$u(x_0) - v(x_0) = \inf_{x \in \bar{\Omega}} \{u(x) - v(x)\} < 0$$

So $v(x) - v(x_0) + u(x_0)$ touches $u(x)$ from below at x_0. By virtue of u's being a viscosity super solution, we have that $F(x_0, u(x_0), Dv(x_0), D^2v(x_0)) \geq 0$. By the properness of F we then have that

$$F(x_0, v(x_0), Dv(x_0), D^2v(x_0)) \geq F(x_0, u(x_0), Dv(x_0), D^2v(x_0)) \geq 0$$

which contradicts the fact that v is a strict subsolution.

A.2 Our Setting

In this thesis we consider the linear parabolic PDE:

$$G(t, x, u, \frac{\partial u}{\partial t}, D_x u, D_x^2 u) = \frac{\partial u}{\partial t} - \sum_{i,j} \alpha_{ij}(x,t) \frac{\partial^2 u}{\partial x_i \partial x_j} - \sum_i \beta_i(x,t) \frac{\partial u}{\partial x_i} + \gamma(x,t)u = 0,$$

where $u(x,t) : \mathbb{R}^d \times [0,\infty) \longrightarrow \mathbb{R}$, $\gamma(x,t) \geq 0$, and the matrix $\{\alpha_{ij}(x,t)\}$ is non-negative definite for each (x,t). We can put this in our general (elliptic) setting by thinking of t as the $(d+1)$st coordinate and thinking of $u(x,t) : \mathbb{R}^{d+1} \longrightarrow \mathbb{R}$ as being a solution to

$$F((x,t), u, Du, D^2u) := G(t, x, u, \frac{\partial u}{\partial t}, D_x u, D_x^2 u) = 0$$

As γ is non-negative and $\{\alpha_{ij}(x,t)\}$ is non-negative definite matrix valued, it is easy to see that F is proper and so the results for viscosity solutions above hold for this PDE.

Bibliography

[1] Rami Atar and Krzysztof Burdzy. On Neumann eigenfunctions in lip domains. *J. Amer. Math. Soc.*, 17(2):243–265 (electronic), 2004.

[2] Rodrigo Bañuelos and Krzysztof Burdzy. On the "hot spots" conjecture of J. Rauch. *J. Funct. Anal.*, 164(1):1–33, 1999.

[3] K. N. Chueh, C. C. Conley, and J. A. Smoller. Positively invariant regions for systems of nonlinear diffusion equations. *Indiana Univ. Math. J.*, 26(2):373–392, 1977.

[4] Michael G. Crandall. Viscosity solutions: a primer. In *Viscosity solutions and applications (Montecatini Terme, 1995)*, volume 1660 of *Lecture Notes in Math.*, pages 1–43. Springer, Berlin, 1997.

[5] Michael G. Crandall, Hitoshi Ishii, and Pierre-Louis Lions. User's guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc. (N.S.)*, 27(1):1–67, 1992.

[6] Francesca Da Lio. Remarks on the strong maximum principle for viscosity solutions to fully nonlinear parabolic equations. *Commun. Pure Appl. Anal.*, 3(3):395–415, 2004.

[7] S. D. Eidel'man. *Parabolic Systems (Russian).* Izdat. "Nauka", Moscow, 1964.

[8] Lawrence Christopher Evans. A strong maximum principle for parabolic systems in a convex set with arbitrary boundary, 2009.

[9] Arturo Kohatsu-Higa. Stratonovich type SDE's with normal reflection driven by semimartingales. *Sankhyā Ser. A*, 63(2):194–228, 2001.

[10] P.-L. Lions and A.-S. Sznitman. Stochastic differential equations with reflecting boundary conditions. *Comm. Pure Appl. Math.*, 37(4):511–537, 1984.

[11] Louis Nirenberg. A strong maximum principle for parabolic equations. *Comm. Pure Appl. Math.*, 6:167–177, 1953.

[12] Roger Petterson. Wong-zakai approximations for reflecting stochastic differential equations. *Stochastic Analysis and Applications*, 17(4):609–617, 1999.

[13] Daniel W. Stroock. *Probability theory, an analytic view.* Cambridge University Press, Cambridge, 1993.

[14] Daniel W. Stroock and S. R. S. Varadhan. Diffusion processes with boundary conditions. *Comm. Pure Appl. Math.*, 24:147–225, 1971.

[15] Daniel W. Stroock and S. R. Srinivasa Varadhan. *Multidimensional diffusion processes*, volume 233 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences].* Springer-Verlag, Berlin, 1979.

[16] Xuefeng Wang. A remark on strong maximum principle for parabolic and elliptic systems. *Proc. Amer. Math. Soc.*, 109(2):343–348, 1990.

[17] Hans F. Weinberger. Invariant sets for weakly coupled parabolic and elliptic systems. *Rend. Mat. (6)*, 8:295–310, 1975. Collection of articles dedicated to Mauro Picone on the occasion of his ninetieth birthday.

[18] Eugene Wong and Moshe Zakai. On the convergence of ordinary integrals to stochastic integrals. *Ann. Math. Statist.*, 36:1560–1564, 1965.