A finite element approach to pricing Barrier options

by
Mark Richards

Submitted in partial fulfillment of the requirements for the degree

Magister Scientiae

in the Department of Mathematics and Applied Mathematics
in the Faculty of Natural and Agricultural Sciences

University of Pretoria
Pretoria

December 2015

Declaration

I, the undersigned, declare that the dissertation, which I hereby submit for the degree Magister Scientiae at the University of Pretoria, is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution.

Signature:

Name: Mark Timothy Richards

Date: 14 December 2015

Contents

Abstract

In this dissertation we consider the valuation of discretely monitored barrier options under the infinite element method. The infinite element method is an extension to the standard finite element method that accepts problems with unbounded spacial domains (such as the Black-Scholes PDE), without resorting to domain truncation. The degeneracy of the Black-Scholes PDE when the underlying asset reaches zero, requires that the method be formulated within the context of weighted Sobolev spaces. We will demonstrate the convergence of the proposed method and provide a rigorous investigation into the underlying weighted Sobolev spaces in which the convergence is to be demonstrated.

List of Symbols

Symbol	Meaning	Page
$\Subset$	Compact containment	
$\hookrightarrow$	Continuous embedding	132
$\subset\subset$	Compact embedding	133
$[\ .\ ]^+$	$max(0,\ .\ )$	
$[\ .\ ]^-$	$-max(0,\ .\ )$	
$(\ .\ ,\ .\ )_X$	Inner product with respect to the space X	
$\|\ .\ \|_X$	Norm with respect to the space X	
$D^{(i)}[\ .\ ]$	Distributional Derivative of order i	125
$supp(\ .\ )$	The support of a function	123
ℓ	Lebesgue measure	
$*$	Convolution operator	134
J_ϵ	Molliffier function	134
Ω	Domain in $\mathbb{R}$, specific details are given in each chapter	
$\delta\Omega$	Boundary of the domain Ω	
$\mathcal{T}_h$	Triangulation of a domain	62
$\mathring{I}$	$I \setminus \delta I$	
$\hat{I}$	Reference element $[0,1]$	67
R_h	Ritz projection operator	100
$L^p(\Omega)$	Lebesgue space of order p	124
$C^n(\Omega)$	Space of continuous functions with continuous derivatives of order n	123
$C^\infty(\Omega)$	Space of infinitely differentiable functions	123
$C_0^\infty(\Omega)$	Space of infinitely differentiable functions with compact support	123
$P_n(\Omega)$	Space of polynomials of degree n	63

P_h^n	Space of piecewise polynomials of degree n	63
Π_h	Piecewise linear interpolation operator	73
$L^p_{loc}(\Omega)$	Space of locally integrable functions	124
$W(\Omega)$	The set of weight functions	7
$L^p(\Omega, \omega)$	Weight Lebesgue space of order p	8
$W^{m,p}(\Omega)$	The m, p-th order Sobolev space	9
$B_p(\Omega)$	An important class of weight functions	9
t or τ	Point in time	
V	Option value	147
$x(t)$	Value of the underlying at time t	141
K	Option strike price	141
T	Option expiration date	141
r	Risk free interest rate	147
σ	Volatility of the underlying	147
L	Lower barrier	144
U	Upper Barrier	144

Introduction

An option is a financial contract that affords its holder the opportunity to buy or sell a given asset for a predetermined price at some time in the future. Options to buy an asset are termed call options, whilst options to sell are termed put options. Options do not require the holder to make use of their right to buy or sell the asset and as such the holder will only do so if it is within their best interests. Options therefore only provide a payoff under certain conditions and thus their value is generally substantially lower than that of the asset on which they are written. For this reason options are popular with investors, as they allow one to take a strong market position for a relatively small capital outlay and limited down side risk, as well as allowing the hedging (mitigation) of specific risks faced by the investor.

As global financial markets have evolved and expanded, so have the needs of investors, who have moved to seek ever more complex options to take precise market views and hedge their positions. To meet these needs, a wide variety of so called exotic (non-standard) options are now available in over the counter (OTC) markets worldwide. One of the first exotic options to appear in the market was the barrier option and it remains one of the most actively traded today.

Barrier options extend vanilla (standard) put and call options by placing a restriction on the values that the underlying may assume over the life of the option. These restrictions take the form of certain levels (barriers) that, upon being reached by the underlying trigger various pay-off features. These features may be classified as being of either *knock-in* or *knock-out* type. A knock-in condition asserts that the option will expire worthless unless the barrier is triggered at some point over the life of the option, whilst a knock-out condition causes the option to expire worthless upon the triggering of the barrier. Barrier options may further be classified as being of either up or down type. If the barrier condition is triggered by the value of the underlying exceeding some value $U > 0$, the option is termed an up-style option, while a down-style barrier option is one in which the barrier condition is triggered by the value of the underlying falling below a certain level $L > 0$.

In classical theory, barrier options are assumed to be monitored in continuous time. In other words, the knock-in and knock-out conditions are applied if the underlying triggers the barrier at any point during the option's life. This is however an unrealistic assumption, as in practice it is impossible to monitor the value of the underlying at each instant in time. In fact, the majority of traded barrier options specify only a fixed number of times $0 \leq t_1 < t_2 < ... < t_N \leq T$ at which the underlying is evaluated against the barrier and allow the underlying to assume any value during the interim time periods. These options are termed discretely monitored barrier options.

While discretely monitored barrier options are far more commonly traded than their continuously monitored counterparts, they present significantly more difficulties in terms of valuation and there is, as yet, no consensus on the most effective valuation method for these options. Merton [30] demonstrated that under the well known Black-Scholes [7] framework one may derive analytic solutions for the value of continuously monitored barrier options. Fusai, Gianluca and Abrahams [22] have shown that this idea may be extended to discretely monitored options, however the analytic formula presented proves impractical as it must itself be evaluated numerically, a task that proves very difficult as the number of monitoring points increases beyond 5. Broadie, Glasserman and Kou [12] present an alternative in the form of a continuity correction that allows one to approximate the value of a discretely monitored barrier option with that of a continuously monitored option with a shifted barrier.

Numerically, options are most commonly valued via binomial (see Cox, Ross and Rubinstein [15]) or trinomial (see Boyle [9]) lattice methods. In the case of barrier options however, such methods are prone to large errors and very slow convergence if the barrier is no optimally positioned with respect to a horizontal layer of nodes. For classical barrier options, problem is treated by Boyle and Lau [10] and Ritchken [32], whilst Ahn, Figlewski and Gao [3] value discretely monitored barrier options under the trinomial lattice framework by making use of an adaptive mesh model and increasing the number of nodes in the vicinity of the barrier.

Recalling that under the Black-Scholes framework an option value may be expressed as the solution to the Black-Scholes partial differential equation (PDE), classical numerical techniques such as the finite element method may also be applied within the field of option pricing. That said, while the finite element method is commonly used to solve PDE's arising in many fields (solid mechanics, fluid flow, thermodynamics and structural analysis to name a few), its use remains fairly rare in finance. A simple exposition of the method in terms of fairly vanilla derivatives is however presented within the work of Seydel [34].

Achdou and Pironneau [1] consider the valuation of a European put option in the context of the Black-Scholes equation with local volatility and term structured interest rates, by making use of the finite element method. They follow a similar approach to that presented in this dissertation to develop the weak formulation of the valuation problem, but rather consider the problem on a truncated domain and hence conduct their analysis in different spaces to that used here. They furthermore present a very brief exposition of the treatment for a standard double barrier option.

The finite element method is also employed to value options under a stochastic volatility model by Apel, Winkler and Wystup [38]; Asian, basket and look-back options by Zvan, Forsyth and Vetzal [42] [20]; Asian and Parisian options by Zhu and Stokes [41] and barrier options, power options and a variety of basket options by Topper [37]. A possible reason for the hesitancy to apply the finite element method within finance is that the Black Scholes PDE differs fundamentally from many of the equations to which the finite element method is commonly applied.

The first of these differences is that the Black-Scholes PDE is defined over an unbounded spacial domain, as the value of the underlying cannot be bounded above. Since the finite element method relies on being able to partition the computational domain into a finite number of sub-domains, each of finite measure, by its very nature it cannot accept a problem defined on an unbounded domain. To address this short coming, it is standard practice within literature to truncate the spacial domain and only consider the Black-Scholes PDE up to some maximum value of the underlying x_{max}. The second major difference between the Black-Scholes PDE and many others, is the fact that it degenerates (the coefficients of both the convection and diffusion terms vanish) as the value of the underlying approaches zero. This characteristic implies that the standard Sobolev spaces (see Section A.2 in the appendix) in which the finite element analysis is usually performed are not applicable and must be replaced with far more complicated weighted versions (see Section 2.1). In order to avoid the complexities that arise from the use of these spaces, many authors resort to applying a logarithmic transformation to reduce the Black-Scholes PDE to the well known heat equation (See [1]). It is however easy to demonstrate that this transformation results in an uneven distribution of nodes on the spacial domain, with fewer nodes being present as the value of the underlying increases. Similarly to any interpolation method, this reduces the accuracy of the solution as the value of the underlying increases, a fact particularly noticeable within the vicinity of the strike price of the option.

Sanfelici [33] proceeds to address the difficulties posed by the Black-Scholes PDE without resorting to domain truncation or logarithmic transformations and presents an adapted version of the finite element method (termed the infinite element method) that is designed to accept problems on unbounded domains. This method is applied within the context of discretely monitored barrier options and the convergence analysis is conducted within the context of weighted Sobolev spaces.

In this dissertation we will follow the work of Sanfelici [33] and consider the valuation of discretely monitored barrier options within the context of the infinite element method. We do however aim to improve upon the work of Sanfelici by providing a much more rigorous treatment of this method and the associated convergence theory. We will place a particular emphasis on the weighted Sobolev theory that is required to demonstrate convergence and in this regard provide a near complete introduction to this topic, as well as providing rigorous demonstrations of weighted analogs to numerous classical results that play a critical part within convergence theory . We will furthermore discuss the selection of suitable weighted spaces for the problem at hand, rigorously demonstrate the existence and uniqueness of the solution to the weak formulation of the valuation problem and provide an introductory discussion of the infinite element method and the selection of the basis functions that generate the associated spaces.

We begin our investigation in Chapter 1, by introducing weighted Sobolev spaces and presenting a number of weighted analogs to classical results that will prove critical during our investigation. We will then proceed, in Chapter 2, to investigate the nature of weighted spaces that arise due to the degeneracy of the Black-Scholes PDE, derive the weak formulation of the valuation problem in terms of these spaces and finally discuss existence and uniqueness topics in the context of weighted spaces. Chapter 3 begins with a gentle introduction to the classical Galerkin finite element method and then proceeds to introduce the infinite element method and derive the semi-discrete version of the valuation problem. We then move, in Chapter 4, to apply the previously developed weighted Sobolev theory and rigorously demonstrate the convergence of the method. The dissertation is then concluded by a brief examination of the numerical application of the developed scheme, after which we will present concluding remarks, as well as a number of appendices that will serve to introduce various topics with which the reader may be unfamiliar.

Chapter 1

Mathematical Preliminaries

In this chapter we begin by introducing weighted Lebesgue and Sobolev spaces and examining under what conditions these spaces are complete. We then proceed to present weighted analogs to the well known Sobolev embedding theorems and then demonstrate the convergence of mollifications (see section A.3 in the appendix) in various weighted norms. Finally, we conclude by examining a special class of smooth functions, termed cut-off functions, that will be crucial in proving a number of density results in later chapters. An introduction to classical Sobolev theory is presented in an appendix at the end of this dissertation and it is recommended that readers unfamiliar with this topic peruse the appendix before proceeding with this chapter.

1.1 Definitions and Completeness

In this section we will define weighted analogs of the standard Lebesgue and Sobolev spaces and present a number of properties of these spaces that will prove useful in later sections. We begin by defining a class of functions that will be suitable to act as weights for these spaces.

Definition 1.1. (See Kufner and Opic [28])
Let $W(\Omega)$ denote the set of all Lebesgue measurable, positive and finite valued functions on Ω. If $\omega \in W(\Omega)$ then we call ω a weight function.

Making use of these weight functions we may now define the weighted analog of the Lebesgue spaces L^p. We will then proceed to demonstrate that these spaces are in fact Banach spaces, a fact that will prove very useful later in the section. Before proceeding with this definition, we make the important observation that context of Lebesgue spaces (and similarly weighted Lebesgue, Sobolev and weighted Sobolev spaces), functions are viewed to be equivalent if the norm of their difference is 0, or equivalently if they differ at most on a set of measure 0. It therefore follows that although we treat (and even refer to) the elements of theses spaces as functions, they are in fact equivalence classes of functions.

Definition 1.2. (See Kufner and Opic [28])
Let $1 \leq p < \infty$ and ω be a weight function, we then define the weighted Lebesgue space $L^p(\Omega, \omega)$ as the collection of all functions that satisfy

$$\|u\|_{L^p(\Omega,\omega)} := \left(\int_\Omega \omega(x) |u(x)|^p \mathrm{d}x \right)^{1/p} < \infty \ .$$

Theorem 1.3. *The weighted Lebesgue space $L^p(\Omega, \omega)$ is a Banach space.*

Proof. Let $A \subset \Omega$ be Lebesgue measurable, and define

$$\nu(A) = \int_A \omega(x) \mathrm{d}x = \int_A \omega \mathrm{d}\ell \ ,$$

where ℓ denotes the standard Lebesgue measure.
Now, since ω is a positive-valued function, it is well known from measure theory that ν is a measure. Furthermore, it is clear that ν is absolutely continuous with respect to the Lebesgue measure ℓ and hence ω is a Radon-Nikodym derivative and thus, (see de Barra [16]) for any $f \in L^p(\Omega, \omega)$ we have that

$$\int_A |f(x)|^p \omega(x) \mathrm{d}x = \int_A |f|^p \mathrm{d}\nu \ .$$

It therefore follows that the space $L^p(\Omega, \omega)$ is equivalent to the space $L^p(\Omega, \mathrm{d}\nu)$ and is hence a Banach space. □

Definition 1.4. (See Kufner and Opic [28])
Let $m \in \mathbb{N}$, $1 \leq p < \infty$ and ω be a set of m weight functions $\{\omega_i\}$. The weighted Sobolev norm is then defined as

$$\|u\|_{W^{m,p}(\Omega,\omega)} := \left(\sum_{i=0}^{m} \|D^{(i)}u\|^p_{L^p(\Omega,\omega_i)} \right)^{1/p} \quad ,$$

where $D^{(i)}$ denotes the distributional derivative of order i (see the appendix for a brief introduction to distributional derivatives).
The weighted Sobolev space $W^{m,p}(\Omega,\omega)$ may then be defined as the collection of all functions for which the above norm is finite.
Notationally, we will write

$$|u|_{W^{m,p}(\Omega,\omega)} = \|D^{(m)}u\|^p_{L^p(\Omega,\omega_m)} \quad .$$

In the theory of weak solutions, which will be central to our analysis in later sections, it will prove useful to know under what conditions these spaces will be complete. To this end, we present the following results due to Kufner and Opic [28].

Definition 1.5. (Kufner and Opic [28])
Given $1 \leq p < \infty$, we denote by $B_p(\Omega)$ the collection of all weight functions $\omega(x)$ such that: $\omega(x)^{-1/(p-1)} \in L^1_{loc}(\Omega)$.

Lemma 1.6. *Let $\omega \in B_p(\Omega)$, $\phi \in C_0^\infty(\Omega)$ and $m \in \mathbb{N}$. Then, for $u \in L^p(\Omega,\omega)$ and $i = 0, 1, ..., m$, define*

$$L_i(u) = \int_\Omega u(x)\phi^{(i)}(x)\mathrm{d}x \ .$$

It follows that L_i is a bounded linear functional on $L^p(\Omega,\omega)$.

Proof. Clearly $L_i : L^p(\Omega,\omega) \to \mathbb{R}$ is linear, hence it remains to show that it is bounded. To this end, notice that for any $u \in L^p(\Omega,\omega)$

$$|L_i(u)| = |\int_\Omega u(x)\phi^{(i)}(x)\mathrm{d}x|$$

$$\leq \int_\Omega |u(x)\phi^{(i)}(x)|\mathrm{d}x$$
$$= \int_\Omega |u(x)|(\omega(x))^{1/p}|\phi(x)|(\omega(x))^{-1/p}\mathrm{d}x \ .$$

Hölder's inequality then implies that for q such that $\frac{1}{p}+\frac{1}{q}=1$

$$|L_i(u)| \leq \|u\|_{L^p(\Omega,\omega)} \left(\int_\Omega |\phi^{(i)}(x)|^q(\omega(x))^{-q/p}\mathrm{d}x\right)^{1/q} .$$

Noting that $q=\frac{p}{p-1}$ and setting $A = supp(\phi) \Subset \Omega$, it follows that

$$|L_i(u)| \leq \|u\|_{\Omega,p,\omega} \left(\int_A |\phi^{(i)}(x)|^q(\omega(x))^{-1/(p-1)}\mathrm{d}x\right)^{(p-1)/p}$$
$$\leq \|u\|_{L^p(\Omega,\omega)} \sup_{x\in A}(|\phi^{(i)}(x)|) \left(\int_A (\omega(x))^{-1/(p-1)}\mathrm{d}x\right)^{(p-1)/p}$$
$$= C\|u\|_{L^p(\Omega,\omega)} \quad ,$$

for

$$C = \sup_{x\in A}(|\phi^{(i)}(x)|) \left(\int_A (\omega(x))^{-1/(p-1)}\mathrm{d}x\right)^{(p-1)/p} \qquad .$$

It therefore follows that as required, L_i is bounded. □

Theorem 1.7. *Let $m \in \mathbb{N}$ and $\omega = \{\omega_0, \omega_1, ..., \omega_m\}$ be a set of weight functions such that for each $i = 0, 1, ..., m$, $\omega_i \in B_p(\Omega)$. It then follows that for any $1 < p < \infty$, $W^{m,p}(\Omega, \omega)$ is complete.*

Proof. Let (u_n) be a Cauchy sequence in $W^{m,p}(\Omega,\omega)$. It is then clear that for each $i = 0, 1, ..., m$, $(D^{(i)}[u_n])$ is Cauchy in $L^p(\Omega,\omega_i)$. Now, since $L^p(\Omega,\omega_i)$ is complete it follows that for each $i = 0, 1, ..., m$ there exists a function $u_{(i)} \in L^p(\Omega,\omega_i)$ such that

$$D^{(i)}[u_n] \to u_{(i)} \in L^p(\Omega,\omega_i) \quad . \tag{1.1}$$

As in the previous result, for any fixed $\phi \in C_0^\infty$ and $i = 0, 1, ..., m$ we set

$$L_i(u) = \int_\Omega u(x)\phi^{(i)}(x)\mathrm{d}x \quad .$$

It follows from Lemma 1.6 that for each $i = 0, 1, ..., m$, L_i is a bounded linear functional on $L^p(\Omega, \omega_i)$ and hence

$$L_i(u_n) \to L_i(u_{(0)})$$

and

$$L_0(D^{(i)}[u_n]) \to L_0(u_{(i)}) \quad .$$

Furthermore, by the definition of the weak derivative, it follows that

$$\begin{aligned} L_i(u_n) &= \int_\Omega u_n(x)\phi^{(i)}(x)\mathrm{d}x \\ &= -\int_\Omega D^{(1)}[u_n(x)]\phi^{(i-1)}(x)\mathrm{d}x \\ &= (-1)^2\int_\Omega D^{(2)}[u_n(x)]\phi^{(i-2)}(x)\mathrm{d}x \\ &. \\ &. \\ &. \\ &= (-1)^i\int_\Omega D^{(i)}[u_n(x)]\phi(x)\mathrm{d}x \quad . \end{aligned}$$

We therefore have that ,

$$L_i(u_0) = \lim_{n\to\infty} L_i(u_n) = (-1)^i \lim_{n\to\infty} L_0(D^{(i)}[u_n] = (-1)^i L_0(u_{(i)})$$

and hence

$$L^i(u_{(0)}) = \int_\Omega u_{(0)}\phi^{(i)}(x)\mathrm{d}x = (-1)^i L_0(u_{(i)}) = (-1)^i\int_\Omega u_{(i)}\phi(x)\mathrm{d}x \quad . \qquad (1.2)$$

Since (1.2) holds for all $\phi \in C_0^\infty(\Omega)$, it follows from the definition of the weak derivative that for each $i = 1, 2, ..., m$,

$$u_{(i)} = D^{(i)}[u_{(0)}] \text{ almost everywhere on } \Omega \quad . \tag{1.3}$$

We therefore have that

$$D^{(i)}[u_{(0)}] \in L^p(\Omega, \omega_i)$$

and hence

$$u_{(0)} \in W^{m,p}(\Omega, \omega) \quad .$$

Finally,

$$\begin{aligned} \|u_n - u_{(0)}\|^p_{W^{m,p}(\Omega,\omega)} \leq &\|u_n - u_0\|_{L^p(\Omega,\omega_0)} + \|D^{(1)}[u_n - u]\|_{L^p(\Omega,\omega_1)} \\ &+ \ . \ . \ . \ + \|D^{(m)}[u_n - u]\|_{L^p(\Omega,\omega_m)} \quad , \end{aligned}$$

making use of (1.3),

$$\begin{aligned} = &\|u_n - u_0\|_{L^p(\Omega,\omega_0)} + \|D^{(1)}[u_n] - u_{(1)}\|_{L^p(\Omega,\omega_1)} \\ &+ \ . \ . \ . \ + \ \|D^{(m)}[u_n] - u_{(m)}\|_{L^p(\Omega,\omega_m)} \quad . \end{aligned}$$

Taking the limit as $n \to \infty$, it follows immediately from (1.1) that

$$\lim_{n\to\infty} \|u_n - u\|_{\Omega,m,p,\omega} = 0$$

and hence $W^{m,p}(\Omega, \omega)$ is complete. □

1.2 Embedding Theorems

It is well known that under certain conditions there exist embedding properties between the standard Sobolev spaces $W^{m,p}(\Omega)$ and the space $C(\bar{\Omega})$ (see Section A.2 of the appendix). Such embedding properties are critical in the development of certain interpolation estimates and hence, in this section, we will present weighted analogs of a number of well known embedding theorems.

Embedding theorems for weighted Sobolev spaces are not extensively covered within the literature and in most cases such results are only presented for certain classes of weight functions. For our purposes it will suffice to only consider the "power-type" weight functions (See Kufner [27]).

Definition 1.8. Given a bounded domain $\Omega \subset \mathbb{R}$, a real number $\alpha \geq 0$ and a set $\Gamma_0 \subset \partial\Omega$, for $x \in \Omega$ we define the weight function

$$\rho_{\Gamma_0}(x)^\alpha = \inf_{y \in \Gamma_0} |x-y|^\alpha \quad .$$

We note that the boundedness of Ω ensures that ρ_{Γ_0} is well defined. Furthermore, for notational convenience, we will write $\rho(x)^\alpha$ in cases where confusion as to the definition of Γ_0 cannot arise.

Notationally, we will write $W^{m,p}(\Omega, \rho^\alpha_{\Gamma_0})$ for weighted spaces that make use of these power-type weights and in the important case where $\Gamma_0 = \{0\}$ and $\alpha = 2$, we will write $W^{m,p}(\Omega, X)$.

Embedding theorems between different orders of power-weighted Sobolev spaces are presented by Kufner [27]; while for some α, embedding results are given in the book of Grisvard [23]. These text do not however consider the complete embedding relations required here and as such, we rather present (without proof) results from a paper due to Timerbaev [35], entitled "Embedding theorems for weighted Sobolev spaces". At the time of writing, the author of this dissertation could not find an English translation of this paper and for this reason we present the original theorems in Russian and then proceed to provide appropriate translations.

As in the case of the standard Sobolev embedding theorems, Timerbaev begins by placing some restrictions of the domain on which the spaces are defined, by noting "ограниченной звездной области Ω ". This translates to the requirement that the domain Ω is bounded and star-shaped.

Definition 1.9. A domain Ω is said to be star-shaped if there exists a point $x_0 \in \Omega$ such that for each $x \in \Omega$ the line segment between x and x_0

lies within Ω.

We note that in $\mathbb{R}$, a star-shaped domain is equivalent to a convex domain and hence this condition holds trivially as, for the purposes of this dissertation we will view Ω to be a bounded interval.

Timerbaev then writes " произвольном подмножестве Γ_0 границы Γ", or Γ_0 being an arbitrary subset of the boundary Γ (we rather use the notation $\partial\Omega$ to denote the boundary of Ω).

Timerbaev now introduces the space $C^m_\beta(\bar{\Omega})$, by writing "Обозначим также через $C^m_\beta(\bar{\Omega})$ пространство функций $u(x)$ таких, что $\rho(x)^\beta D^{(i)}u(x) \in C(\bar{\Omega})$ для $i \leq m$." or in English, We denote by $C^m_\beta(\bar{\Omega})$ the collection of functions $u(x)$ such that

$$\rho(x)^\beta D^{(i)}u(x) \in C(\bar{\Omega}), \text{ for } i \leq m \quad . \tag{1.4}$$

Timerbaev then gives the following embedding theorems.

Теорема 1. Пусть $1 < p < +\infty$, $\gamma = l - m - n/p > 0$, $l > m$, $\alpha, \beta \geq 0$. Еслу $\beta \leq \alpha + \gamma$, То $W^{l,p}(\Omega, \rho^{l\beta})$ компакмно вложено в $C^m_\alpha(\bar{\Omega})$.

Теорема 2. Пусть $1 < p \leq q < +\infty$, $\gamma = l - m - n/p + n/q > 0$, $l > m$, $\alpha, \beta \geq 0$. Еслу $\beta \leq \alpha + \gamma$, То $W^{l,p}(\Omega, \rho^{l\beta})$ компакмно вложено в $W^{m,q}(\Omega, \rho^{m\alpha})$.

In English, we have:

Theorem 1.10. *Let $1 < p < +\infty$, $\gamma = l - m - n/p > 0$, $l > m$, $\alpha, \beta \geq 0$. If $\beta \leq \alpha + \gamma$, then $W^{l,p}(\Omega, \rho^{l\beta})$ is compactly embedded in $C^m_\alpha(\bar{\Omega})$.*

Theorem 1.11. *Let $1 < p \leq q < +\infty$, $\gamma = l - m - n/p + n/q > 0$, $l > m$, $\alpha, \beta \geq 0$. If $\beta \leq \alpha + \gamma$, then $W^{l,p}(\Omega, \rho^{l\beta})$ is compactly embedded in $W^{m,q}(\Omega, \rho^{m\alpha})$*

Choosing $\Omega = (0, a)$ for some $a > 0$, $\Gamma_0 = \{0\}$ and $\alpha = 0$, $m = 0$, $\beta = 2/l$, we obtain the following simplified result.

Theorem 1.12. *Let $1 < p < \infty$ and $0 < k$. Now set $\gamma = k - \frac{2}{k} - \frac{1}{p}$, then provided that $\gamma > 0$ the following compact embedding holds*

$$W^{k,p}(\Omega, X) \subset\subset C(\bar{\Omega}) \quad .$$

Similarly, setting $\alpha = 2/m$ and $\beta = 2/l$

Theorem 1.13. *Let $1 < p \leq q < \infty$ and $m < k$. Now set $\gamma = k - m - \frac{1}{p} + \frac{1}{q}$, then provided that $\gamma > 0$ the following compact embedding holds*

$$W^{k,p}(\Omega, X) \subset\subset W^{m,q}(\Omega, X) \quad .$$

1.3 Mollification in Weighted Norms

It is well known that the proofs of many of the key results in Sobolev theory require one to pass to an approximating sequence of smooth functions via the process of mollification (See section A.3 in the appendix). Theorem A.26 (part 3) plays a crucial role in this regard as it provides conditions under which such a sequence converges in L^p-norms. In order to employ a similar technique in the setting of weighted spaces, we require a similar theorem for weighted Lebesgue spaces. This section will be dedicated to the proof of such a result for a certain power-weighted space. We note that the proof of this result is adapted from the proof of the unweighted case, as given by Adams and Fournier [2]. We begin by supposing that Ω is a bounded, open interval of the form $(0, a)$, for some $a > 0$.

Theorem 1.14. *Let $u \in L^2(\Omega, X) \cap L^1_{loc}(\Omega)$, with $supp(u) \Subset \Omega$. Then given $\epsilon > 0$, there exists $v \in C_0(\Omega)$ such that*

$$\|u - v\|_{L^2(\Omega,X)} < \epsilon.$$

Proof. We begin by noting that each function $u \in L_x^2(\Omega) \cap L_{loc}^1(\Omega)$ may be written in the form $u = u^+ - u^-$, where u^+ and u^- are both non-negative members of $L^2(\Omega, X) \cap L_{loc}^1(\Omega)$. Without loss of generality we may therefore restrict ourselves to the case in which u is non-negative. Now, since u is measurable, it is well known from measure theory that there exists a monotone increasing sequence of non-negative, measurable simple functions $\{s_n\}$ that converges pointwise to u on Ω.
Thus, clearly we have

$$0 \leq s_n \leq u \quad \forall\, n \in \mathbb{N}$$

and

$$s_n \in L^2(\Omega, X) \quad \forall\, n \in \mathbb{N} \quad .$$

Furthermore, for each $x \in \Omega$

$$x^2 \left[u(x) - s_n(x)\right]^2 \leq x^2 (u(x))^2 \in L^1(\Omega)$$

and hence Lebesgue dominated convergence theorem implies that

$$s_n \to u \in L^2(\Omega, X) \quad .$$

It therefore follows that for each $\epsilon > 0$, there exists s in $\{s_n\}$ such that

$$\|u - s\|_{L^2(\Omega,X)} < \frac{\epsilon}{2} \quad .$$

Clearly, we may assume that $s = 0$ outside of $supp(u)$ and hence due to Lusin's Theorem, there exists a $\phi \in C_0(\Omega)$ such that

$$|\phi(x)| \leq \|s\|_{\infty,\Omega} \quad \forall x \in \Omega \tag{1.5}$$

and, if ℓ denotes the Lebesgue measure,

$$\ell\{x \in \Omega : \phi(x) \neq s(x)\} < \left(\frac{\epsilon}{4\|s\|_{\infty,\Omega}\|x\|_{\infty,\Omega}}\right)^2 \quad . \tag{1.6}$$

Now

$$\begin{aligned}\|s-\phi\|_{L^2(\Omega,X)} &= \|x(s-\phi)\|_{L^2(\Omega)} \\ &\leq \|x\|_{\infty,\Omega}\ \|s-\phi\|_{L^2(\Omega)} \\ &\leq \|x\|_{\infty,\Omega}\ \|s-\phi\|_{\infty,\Omega}\ \ell\left(\{x\in\Omega : \phi(x)\neq s(x)\}\right)^{1/2} \quad ,\end{aligned}$$

making use of (1.5) and (1.6),

$$\begin{aligned}&\leq 2\|x\|_{\infty,\Omega}\ \|s\|_{\infty,\Omega}\ \frac{\epsilon}{4\|x\|_{\infty,\Omega}\|s\|_{\infty,\Omega}} \\ &< \frac{\epsilon}{2} \quad .\end{aligned}$$

As required, we therefore have that

$$\begin{aligned}\|u-\phi\|_{L^2(\Omega,X)} &\leq \|u-s\|_{L^2(\Omega,X)} + \|s-\phi\|_{L^2(\Omega,X)} \\ &< \epsilon \quad .\end{aligned}$$

□

Making use of the above result, we now have the following weighted analog of part 3 of Theorem A.26.

Theorem 1.15. *Given* $u \in L^2(\Omega,X)\cap L^1_{loc}(\Omega)$*, with* $supp(u) \Subset \Omega$*, it follows that* $\lim_{\epsilon\to 0^+}\|J_\epsilon * u - u\|_{L^2(\Omega,X)} = 0$.

Proof. For $\epsilon > 0$, we denote by J_ϵ the standard mollifier as introduced in section A.3 of the appendix. Then, making use of Hölder's inequality and recalling that $\int_{\mathbb{R}} J_\epsilon(x)\mathrm{d}x = 1$, it follows that

$$\begin{aligned}|xJ_\epsilon * u(x)| &= \left|\int_{\mathbb{R}} xJ_\epsilon(x-y)u(y)\mathrm{d}y\right| \\ &\leq \left(\int_{\mathbb{R}} x^2|u(x)|^2 J_\epsilon(x-y)\mathrm{d}y\right)^{1/2} \\ &\quad \left(\int_{\mathbb{R}} J_\epsilon(x-y)\mathrm{d}y\right)^{1/2}\end{aligned}$$

$$= \left(\int_{\mathbb{R}} x^2 |u(x)|^2 J_\epsilon(x-y) \mathrm{d}y \right)^{1/2} \quad .$$

The application of Fubini's theorem to interchange the order of integration then yields

$$\begin{aligned}
\|J_\epsilon * u\|^2_{L^2(\Omega,X)} &= \|xJ_\epsilon * u\|^2_{L^2(\Omega)} \\
&= \int_\Omega \int_{\mathbb{R}} x^2 J_\epsilon(x-y)|u(y)|^2 \mathrm{d}y \ \mathrm{d}x \\
&\leq \int_{\mathbb{R}} \int_{\mathbb{R}} x^2 J_\epsilon(x-y)|u(y)|^2 \mathrm{d}y \ \mathrm{d}x \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}} x^2 J_\epsilon(x-y)|u(y)|^2 \mathrm{d}x \ \mathrm{d}y \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}} x^2 J_\epsilon(y-x) \mathrm{d}x \ |u(y)|^2 \mathrm{d}y \\
&= \int_{\mathbb{R}} [J_\epsilon * (y^2)]|u(y)|^2 \mathrm{d}y \\
&= \int_{\Omega} [J_\epsilon * (y^2)]|u(y)|^2 \mathrm{d}y \quad . \qquad (1.7)
\end{aligned}$$

It now follows from Theorem 1.14, that given $\delta > 0$, there exists $v \in C_0(\Omega)$ such that

$$\|u - v\|_{L^2(\Omega,X)} < \frac{\delta}{2} \quad . \qquad (1.8)$$

Hence, making use of (1.7) and (1.8), we have that for $n \in \mathbb{N}$,

$$\begin{aligned}
\|J_{1/n} * u - u\|_{L^2(\Omega,X)} &\leq \|J_{1/n} * u - J_{1/n} * v\|_{L^2(\Omega,X)} \\
&\quad + \|J_{1/n} * v - v\|_{L^2(\Omega,X)} + \|u - v\|_{L^2(\Omega,X)} \\
&= \|J_{1/n} * (u - v)\|_{L^2(\Omega,X)} \\
&\quad + \|J_{1/n} * v - v\|_{L^2(\Omega,X)} + \|u - v\|_{L^2(\Omega,X)} \\
&\leq \int_\Omega J_{1/n} * (x^2)|u(x) - v(x)|^2 \mathrm{d}x \\
&\quad + \|J_{1/n} * v - v\|_{L^2(\Omega,X)} + \|u - v\|_{L^2(\Omega,X)} \\
&< \int_\Omega J_{1/n} * (x^2)|u(x) - v(x)|^2 \mathrm{d}x
\end{aligned}$$

$$
\begin{aligned}
&+ \left(\int_\Omega x^2 |J_{1/n} * v(x) - v(x)|^2 \mathrm{d}x \right)^{1/2} + \frac{\delta}{2} \\
&< \int_\Omega J_{1/n} * (x^2) |u(x) - v(x)|^2 \mathrm{d}x \\
&+ \|x\|_{\infty,\Omega} \|J_{1/n} * v - v\|_{L^2(\Omega)} + \frac{\delta}{2} \quad . \qquad (1.9)
\end{aligned}
$$

Now, since $x^2 \in L^1_{loc}(\mathbb{R})$, Theorem A.26 implies that for each $n \in \mathbb{N}$, $J_{1/n} * x^2 \in C^\infty(\mathbb{R})$ and is therefore bounded on Ω. That is, for each $n \in \mathbb{N}$, there exists a $M_n > 0$ such that

$$
|J_{1/n} * x^2| \leq M_n \quad .
$$

Furthermore, since $x^2 \in C(\overline{\Omega})$, Theorem A.26 implies that

$$
\lim_{n \to \infty} J_{1/n} * (x^2) = x^2 \text{ uniformly on } \Omega \ .
$$

It therefore follows that there exists a $N \in \mathbb{N}$ such that for each $n \geq N$ and every $x \in \Omega$

$$
|J_{1/N} * x^2 - x^2| \ , \ |J_{1/N} * x^2 - J_{1/n} * x^2| < 1 \quad . \qquad (1.10)
$$

Thus, for $n \geq N$, due to (1.10), we have that

$$
\begin{aligned}
|J_{1/n} * x^2| &\leq |J_{1/N} * x^2 - x^2| + |J_{1/N} * x^2 - J_{1/n} * x^2| + x^2 \\
&< 2 + x^2 \quad .
\end{aligned}
$$

and hence

$$
\begin{aligned}
|J_{1/n} * x^2| &< max\{M_1, M_2, ..., M_N, 2 + x^2\} \\
&< C \quad , \qquad (1.11)
\end{aligned}
$$

where

$$
C = max\{M_1, M_2, ..., M_N, 2 + \sup_{x \in \Omega}(x^2)\} \quad .
$$

We then notice that the supports of both u and v are compactly contained within Ω. We therefore have that $u - v$ vanishes within some neighbourhood of the boundary of Ω and thus within some neighbourhood of 0. It therefore follows that if we set $B = supp(u) \cup supp(v)$, we have that

$$\begin{aligned}(u(x) - v(x))^2 &= \frac{1}{x^2}\left[x^2(u(x) - v(x))^2\right] \\ &\leq \max_{x \in B}\left(\frac{1}{x^2}\right)\left[x^2(u(x) - v(x))^2\right] \quad . \end{aligned} \tag{1.12}$$

Combining (1.11) and (1.12), we therefore have that for $x \in \Omega$

$$|J_{1/n} * x^2(u(x) - v(x))^2| \leq C\left[x^2(u(x) - v(x))^2\right] \quad . \tag{1.13}$$

Since $u \in L^2(\Omega, X)$ and $v \in C_0(\Omega)$, the right-hand side of (1.13) is clearly integrable over Ω and hence we may apply the Lebesgue dominated convergence theorem and (1.8) to obtain

$$\begin{aligned}\lim_{n\to\infty}\int_\Omega J_{1/n} * (x^2)|u(x) - v(x)|^2 \mathrm{d}x &= \int_\Omega \lim_{n\to\infty} J_{1/n} * (x^2)|u(x) - v(x)|^2 \mathrm{d}x \\ &= \int_\Omega x^2|u(x) - v(x)|^2 \mathrm{d}x \\ &< \frac{\delta}{2}\end{aligned}$$

Taking the limit as $n \to \infty$ in (1.9), we therefore have

$$\lim_{n\to\infty}\|J_{1/n} * u - u\|_{L^2(\Omega,X)} < \delta + \lim_{n\to\infty}\|x\|_{\infty,\Omega}\|J_{1/n} * v - v\|_{L^2(\Omega)} \quad .$$

The result now follows due to Theorem A.26, by noting that $v \in C_0(\Omega)$ and thus $v \in L^2(\Omega)$.

□

1.4 Cut-off Functions

We now shift our attention to a second important class of smooth functions, which we will term the cut-off functions. These functions behave similarly to the standard indicator function in that, when multiplied with some function u over the domain Ω, they reduce the support of u to some subset of Ω. Unlike indicator functions however, cut-off functions do not introduce discontinuities and hence preserve the level of differentiability of u. We shall now introduce two examples of such functions that will prove useful in later sections.

Definition 1.16. Given $a > 0$, then let $\chi_{[a]}$ be a member of $C^\infty(\mathbb{R}^+)$ which satisfies

1.
$$\chi_{[a]}(x) = \begin{cases} 1 & \text{if} \quad x \leq \dfrac{a}{3} \\ \\ 0 & \text{if} \quad x \geq \dfrac{2a}{3} \end{cases} .$$

2.
$$\chi_{[a]}(x) \leq 1 \quad \text{for all} x \in \mathbb{R}^+$$

3.
$$|\chi'_{[a]}(x)| \leq \frac{C}{a} \quad \text{for some } C > 0 \quad .$$

Definition 1.17. For each $n \in \mathbb{N}$ we denote by χ_n a member of $C^\infty(\mathbb{R}^+)$ which satisfies

1.
$$\chi_n(x) = \begin{cases} 1 & \text{if} \quad x \in [\frac{1}{n}, \infty) \\ 0 & \text{if} \quad x \in [0, \frac{1}{2n}] \end{cases} .$$

2.
$$\chi'_n(x) \geq 0 \quad \text{for every } x \in \mathbb{R}^+ .$$

3.
$$\chi_n'(x) \leq Cn \quad \text{for every } x \in \mathbb{R}^+ \text{ and some } C > 0 \ .$$

We note that the cut-off function χ_n also has the following useful property.

Lemma 1.18. *For each $n \in \mathbb{N}$, the function $x\chi_n'(x)$ is bounded on $\mathbb{R}^+$.*

Proof. We begin by noting that since χ_n is constant outside of $[\frac{1}{2n}, \frac{1}{n}]$, its derivative vanishes outside of this interval. It therefore follows that there exists a constant $C > 0$ such that

$$\sup_{x\in\mathbb{R}^+} (x\chi_n'(x)) = \sup_{x\in[1/2n,1/n]} [x\chi_n'(x)] \leq \frac{1}{n}Cn = C \quad .$$

□

The construction of cut-off functions that merely reduce the support of a given function u is fairly straightforward and follows from Urysohn's Lemma (see [25] p. 177). The construction of such functions with boundedness constraints on the derivative (as required here) is however not as straightforward and hence, for illustrative purposes, we provide the following examples.

Example 1.19. For fixed $n \in \mathbb{N}$, we begin by setting

$$A_n = (0, \frac{1}{2n} + \frac{1}{8n}] \cup [\frac{1}{n} - \frac{1}{8n}, \infty)$$

and

$$\tilde{A}_n = (0, \frac{1}{2n}] \cup [\frac{1}{n}, \infty)$$

and then define

$$f_n(x) = \begin{cases} 0 & \text{if } \ x \in A_n \\ 2n & \text{if } \ x \in A_n^c = \mathbb{R}^+ \setminus A_n \end{cases}$$

and

$$\begin{aligned}
\psi_n(x) &= J_{\epsilon_n} * f_n(x) \\
&= \int_{\mathbb{R}^+} J_{\epsilon_n}(x-y) f_n(y)\mathrm{d}y \\
&= \int_{x-\epsilon_n}^{x+\epsilon_n} J_{\epsilon_n}(x-y) f_n(y)\mathrm{d}y \ ,
\end{aligned}$$

where $\epsilon_n = \dfrac{1}{16n}$.

It then follows immediately from Theorem (A.26) that $\psi_n \in C^\infty(\mathbb{R}^+)$. Moreover, we note that ψ_n has the following properties

1. For every $x \in \mathbb{R}^+$,

$$\begin{aligned}
\inf_{y\in\mathbb{R}^+}[f_n(y)] \int_{x-\epsilon_n}^{x+\epsilon_n} J_\epsilon(x-y)\mathrm{d}y &\leq \psi_n(x) \\
&\leq \sup_{y\in\mathbb{R}^+}[f_n(y)] \int_{x-\epsilon_n}^{x+\epsilon_n} J_{\epsilon_n}(x-y)\mathrm{d}y \ ,
\end{aligned}$$

   and hence

$$0 \leq \psi_n(x) \leq 2n \ . \tag{1.14}$$

2. If $x \in (0, \dfrac{1}{2n}]$, then

$$\begin{aligned}
\psi_n(x) &= \int_{x-\epsilon_n}^{x+\epsilon_n} J_{\epsilon_n}(x-y) f_n(y)\mathrm{d}y \\
&\leq \int_{x-\epsilon_n}^{x+1/8n} J_{\epsilon_n}(x-y) f_n(y)\mathrm{d}y \\
&= 0 \ . \qquad (1.15)
\end{aligned}$$

3. If $x \in [\dfrac{1}{n}, \infty)$, then

$$\psi_n(x) = \int_{x-\epsilon_n}^{x+\epsilon_n} J_{\epsilon_n}(x-y) f_n(y)\mathrm{d}y$$

$$\leq \int_{x-1/8n}^{x+\epsilon_n} J_{\epsilon_n}(x-y) f_n(y) \mathrm{d}y$$
$$= 0 \ . \tag{1.16}$$

4. If $x \in A_0 = (\frac{1}{2n} + \frac{1}{8n} + \epsilon_n, \frac{1}{n} - \frac{1}{8n} - \epsilon_n)$, then since

   $(x - \epsilon_n, x + \epsilon_n) \subset A_n^c$ it follows that

$$\psi_n(x) = \int_{x-\epsilon_n}^{x+\epsilon_n} J_{\epsilon_n}(x-y) f_n(y) \mathrm{d}y = 2n \ .$$

We now define

$$\psi_n^*(x) = \psi_n(x) \left(\int_{\tilde{A}_n^c} \psi_n(y) \mathrm{d}y \right)^{-1} ,$$

it follows from (1.14) and property (4) above, that

$$\begin{aligned} \psi_n^*(x) &\leq 2n \left(\int_{\tilde{A}_n^c} \psi_n(y) \mathrm{d}y \right)^{-1} \\ &\leq 2n \left(\int_{A_0} \psi_n(y) \mathrm{d}y \right)^{-1} \\ &= 2n(2n\frac{1}{8n})^{-1} \\ &= 8n \ . \end{aligned}$$

Furthermore, it follows from (1.15) and (1.16) that

$$\psi_n^*(x) = 0 \quad \text{for every } x \in \tilde{A}_n \ . \tag{1.17}$$

Finally, for each $x \in \mathbb{R}^+$, we define

$$\chi_n(x) = \int_0^x \psi_n^*(t) \mathrm{d}t \ .$$

It follows from (1.17) that, if $x \in [0, \frac{1}{2n}]$, then

$$\begin{aligned}\chi_n(x) &= \int_0^x \psi_n^*(t)\mathrm{d}t \\ &= 0\end{aligned}$$

and if $x \in [\frac{1}{n}, \infty)$,

$$\begin{aligned}\chi_n(x) &= \int_0^x \psi_n^*(t) \\ &= \int_{\tilde{A}_n^c} \psi_n^*(t)\mathrm{d}t \\ &= 1\ .\end{aligned}$$

Noting that $\chi_n'(x) = \psi_n^*(x)$, it is clear that we have constructed a smooth function that satisfies the conditions of Definition 1.17.

Example 1.20. Setting

$$A = (0, \frac{1}{3a} + \frac{1}{9a}] \cup [\frac{2}{3a} - \frac{1}{9a}, \infty) \quad , \quad \tilde{A} = (0, \frac{1}{3a}] \cup [\frac{2}{3a}, \infty)$$

and

$$f(x) = \begin{cases} 0 & \text{if} \quad x \in A \\ \frac{2}{a} & \text{if} \quad x \in A^c = \mathbb{R}^+ \setminus A_n \end{cases} ,$$

the construction of a smooth function that satisfies Definition 1.16 follows by a similar procedure to the previous example.

Making use of the above cut-off functions, we now define the following products.

Definition 1.21. Given any function u, we define

$$\tilde{u} = \chi_{[a]} u$$

and for $n \in \mathbb{N}$,

$$u_{(n)} = \chi_n u \quad , \quad \tilde{u}_{(n)} = \chi_{[a]}\chi_n u \quad .$$

Since both of the cut-off functions $\chi_{[a]}$ and χ_n are bounded above by 1, it is clear that the products $\tilde{u}$ and $u_{(n)}$ will satisfy the same integrability properties as u.

In later chapters it will prove useful to know in which spaces the sequence $(u_{(n)})$ converges to u and in this regard we conclude this section by demonstrating two cases in which this convergence occurs.

Lemma 1.22. *Given $u \in L^2(\Omega)$, it follows that*

$$\lim_{n\to\infty} \|u - u_{(n)}\|_{L^2(\Omega)} = 0$$

Proof.

$$\begin{aligned}
\|u - u_{(n)}\|_{L^2(\Omega)} &= \int_0^a |u(x) - \chi_n(x)u(x)|^2 \mathrm{d}x \\
&= \int_0^{1/2n} |u(x)|^2 \mathrm{d}x + \int_{1/2n}^{1/n} |u(x) - \chi_n(x)u(x)|^2 \mathrm{d}x \\
&\le \int_0^{1/2n} |u(x)|^2 + \int_{1/2n}^{1/n} |1 - \chi_n|^2 |u(x)|^2 \mathrm{d}x \quad ,
\end{aligned}$$

recalling that χ_n is bounded above by 1 and below by 0,

$$\le \int_0^{1/n} |u(x)|^2 \mathrm{d}x \quad ,$$

and hence, as required, we have that

$$\begin{aligned}
\lim_{n\to\infty} \|u - \chi_n u\|_{L^2(\Omega)} &\le \lim_{n\to\infty} \int_0^{1/n} |u(x)|^2 \mathrm{d}x \\
&= 0 \quad .
\end{aligned}$$

□

Lemma 1.23. *Given* $u \in W^{1,2}(\Omega, X) \cap L^2(\Omega)$*, if follows that*

$$\lim_{n\to\infty} |u - u_{(n)}|_{W^{1,2}(\Omega,X)} = 0$$

Proof. Recalling the definition of the function χ_n, it follows that for $u \in W^{1,2}(\Omega, X)$,

$$\begin{aligned} |u - u_{(n)}|_{W^{1,2}(\Omega,X)} &= \int_\Omega x^2 |D^{(1)}[u(x)] - D^{(1)}[\chi_n(x)u(x)]|^2 \mathrm{d}x \\ &= \int_0^{1/2n} x^2 |D^{(1)}[u(x)]|^2 \mathrm{d}x \\ &\quad + \int_{1/2n}^{1/n} x^2 |D^{(1)}[u(x)] - D^{(1)}[\chi_n(x)u(x)]|^2 \mathrm{d}x \quad , \end{aligned}$$

making use of the product rule for weak derivatives and recalling that the weak and classical derivatives coincide a.e,

$$\begin{aligned} &= \int_0^{1/2n} x^2 |D^{(1)}[u(x)]|^2 \mathrm{d}x \\ &\quad + \int_{1/2n}^{1/n} x^2 |\chi_n'(x)u(x) + \chi_n(x)D^{(1)}[u(x)] - D^{(1)}[u(x)]|^2 \mathrm{d}x \\ &= \int_0^{1/2n} x^2 |D^{(1)}[u(x)]|^2 \mathrm{d}x + \int_{1/2n}^{1/n} x^2 |\chi_n'(x)u(x)|^2 \mathrm{d}x \\ &\quad + \int_{1/2n}^{1/n} x^2 |\chi_n(x)D^{(1)}[u(x)]|^2 \mathrm{d}x + \int_{1/2n}^{1/n} x^2 |D^{(1)}[u(x)]|^2 \mathrm{d}x \\ &\quad + 2\int_{1/2n}^{1/n} x^2 |\chi_n'(x)u(x)||\chi_n(x)D^{(1)}[u(x)]| \mathrm{d}x \\ &\quad - 2\int_{1/2n}^{1/n} x^2 |\chi_n'(x)u(x)||D^{(1)}[u(x)]| \mathrm{d}x \\ &\quad - 2\int_{1/2n}^{1/n} x^2 |\chi_n(x)D^{(1)}[u(x)]||D^{(1)}[u(x)]| \mathrm{d}x \quad , \end{aligned}$$

noting that χ_n is bounded above by 1 and that that last two integrals above are negative,

$$\begin{aligned}
&\leq \int_0^{1/2n} x^2|D^{(1)}[u(x)]|^2\mathrm{d}x \\
&\quad + \int_{1/2n}^{1/n} x^2|\chi_n'(x)u(x)|^2\mathrm{d}x + 2\int_{1/2n}^{1/n} x^2|D^{(1)}[u(x)]|^2\mathrm{d}x \\
&\quad + 2\int_{1/2n}^{1/n} x^2|\chi_n'(x)u(x)||D^{(1)}[u(x)]|\mathrm{d}x \\
&\leq \int_0^{1/2n} x^2|D^{(1)}[u(x)]|^2\mathrm{d}x \\
&\quad + \int_{1/2n}^{1/n} x^2|\chi_n'(x)u(x)|^2\mathrm{d}x + 2\int_0^{1/n} x^2|D^{(1)}[u(x)]|^2\mathrm{d}x \\
&\quad + 2\int_0^{1/n} x^2|\chi_n'(x)u(x)||D^{(1)}[u(x)]|\mathrm{d}x \quad .
\end{aligned}$$

Taking the limit as $n \to \infty$ on both sides of the above inequality then yields

$$\lim_{n\to\infty} |u - \chi_n u|_{W^{1,2}(\Omega,X)} \leq \lim_{n\to\infty} \int_{1/2n}^{1/n} x^2|\chi_n'(x)u(x)|^2\mathrm{d}x \quad ,$$

recalling the boundedness condition on χ_n,

$$\begin{aligned}
&\leq \lim_{n\to\infty} Cn^2 \sup_{x\in[1/2n,1/n]}(x^2) \int_{1/2n}^{1/n} |u(x)|^2\mathrm{d}x \\
&\leq \lim_{n\to\infty} C\int_0^{1/n} |u(x)|^2\mathrm{d}x \\
&= 0 \quad .
\end{aligned}$$

□

Chapter 2

The Weak Formulation of the Valuation Problem

We recall that the aim of this dissertation is to follow the work of Sanfelici [33] and obtain price estimates for discretely monitored barrier options by making use of a modified Galerkin finite element method. It is well known that such methods rely on reformulating the original partial differential equation (the strong form of the problem) as an integral identity (the weak form of the problem). We will begin this chapter by formally introducing the strong form of the valuation problem at hand and then proceed to consider the selection of suitable spaces in which to conduct our analysis. Making use of these spaces, we will then derive the required weak formulation of the problem and conclude by demonstrating the existence and uniqueness of its solution.

Let us consider a double knock-out European call option, with lower barrier $l > 0$, upper barrier $u > l$ and monitoring dates $0 \leq \tau_1 < \tau_2 < ... < \tau_M < \tau_{M+1} = T$. We recall from our initial discussion of option pricing that the price of this derivative will satisfy the Black-Scholes partial differential equation, with terminal condition

$$V(x,T) = \begin{cases} [x(T) - K]^+ & \text{if } \; l < x(\tau_m) < u \; \text{ for } m = 1, 2, ...M+1 \\ 0 & \text{otherwise} \end{cases} .$$

Where, as noted in the Appendix, $x(t) \in \mathbb{R}^+$ denotes the value of the underlying asset at time t.

The theory of partial differential equations is generally geared toward the study of initial value problems and hence it will prove convenient for us to reverse time and reformulate the above problem as an initial value problem. To this end, we define t to be the time until expiry and set $t = T - \tau$ for $\tau \in [0, T]$. We furthermore denote the transformed monitoring dates by $t_m = T - \tau_{M-m+1}$ for $m = 0, 1, 2, ...M$. Under this transformation we obtain the strong formulation of the valuation problem.

Problem 1. *Find $V(x,t)$ that satisfies the reverse-time Black-Scholes partial differential equation,*

$$-\frac{\partial V}{\partial t} + rx\frac{\partial V}{\partial x} + \frac{\sigma^2}{2}x^2\frac{\partial^2 V}{\partial x^2} = rV \quad ,$$

with the initial condition

$$V(x,0) = (x-K)^+ \chi_{l,u}(x) \quad ,$$

that is updated at each monitoring date t_m by

$$V(x,t_m) = V(x,t_m^-)\chi_{l,u}(x) \quad .$$

Here t_m^- denotes the instant just before time t_m and $\chi_{l,u}$ is the indicator function

$$\chi_{l,u}(x) = \begin{cases} 1 & \text{if } \; l < x < u \\ 0 & \textit{otherwise} \end{cases} \quad .$$

We note that for the remainder of this work the option value V will be a function of both the underlying x and time t, however for notational convenience we will often omit one or both of the variables.

2.1 Suitably weighted Sobolev Spaces

Recalling Problem 1 above, we have that the option value V satisfies the reverse time Black-Scholes partial differential equation

$$-\frac{\partial V}{\partial t} + rx\frac{\partial V}{\partial x} + \frac{\sigma^2}{2}x^2\frac{\partial^2 V}{\partial x^2} = rV \quad .$$

It will now be convenient for us to shift our notation and rather denote classical partial derivatives via subscripts. Under this notation, the option value V satisfies

$$-V_t + rxV_x + \frac{\sigma^2}{2}x^2V_{xx} = rV \quad . \tag{2.1}$$

We notice that we may rewrite equation (2.1) as

$$-V_t + \frac{1}{2}\sigma^2\left[x^2V_x\right]_x + (r-\sigma^2)xV_x - rV = \mathcal{L}(V) = 0 \quad . \tag{2.2}$$

Following Guermond and Ern [18] (page 112), the operator $\mathcal{L}$ is called elliptic if there exits $c > 0$ such that

$$x^2 \geq c \quad .$$

It is clear that this identity holds on any interval of the form (a, ∞) for $a > 0$, but fails as x approaches 0. We therefore say that the PDE is degenerate at $x = 0$. It therefore follows from Kufner [27] that the natural norm that arises during the derivation of weak formulation of Problem 1 is a weighted Sobolev norm. Classical Sobolev spaces will therefore be of no use in our investigation and we must rather choose a suitably weighted alternative. It follows from Kufner [27] that a natural space to consider for our analysis is given by $W^{1,2}(\mathbb{R}^+, \omega^*)$, with $\omega^* = \{1, x^2\}$ (the reason that this space is suitable will become apparent further in the Chapter). We note that the work of Achdou and Pironneau [1] presents an analysis of the Black-Scholes PDE in terms of this space.

We now recall that the norm associated with this space is given by

$$\|u\|^2_{W^{1,2}(\mathbb{R}^+,\omega^*)} = \int_{\mathbb{R}^+} [u(x)]^2 \mathrm{d}x + \int_{\mathbb{R}^+} x^2 \left[D^{(1)}[u(x)]\right]^2 \mathrm{d}x \quad .$$

Due to the unbounded nature of the domain, this norm places strict restrictions on the behaviour of functions within the space $W^{1,2}(\mathbb{R}^+, \omega^*)$ as x becomes large. In particular, given $u \in W^{1,2}(\mathbb{R}^+, \omega^*)$, we require that if the limits

$$\lim_{x\to\infty} u(x) \quad \text{and} \quad \lim_{x\to\infty} xD^{(1)}[u(x)]$$

exist, then they must be equal to 0. To demonstrate these conditions, suppose for example that there $\lim_{x\to\infty} u(x) = C$ for some $C > 0$, then there exits $a \in \mathbb{R}^+$ such that

$$|u(x) - C| < \frac{1}{2}C \quad \forall\ x > a$$

or

$$u(x) > \frac{1}{2}C \quad \forall\ x > a$$

and hence

$$\int_0^\infty [u(x)]^2 \mathrm{d}x > \ell(a,\infty)[\frac{1}{2}C]^2 \quad .$$

The right hand side of this equation is clearly infinite and hence $u \notin W^{1,2}(\mathbb{R}^+)$. A similar example may be used to demonstrate the second condition.

These restrictions are clearly very unsuitable for financial applications as they exclude many of the most common financial derivatives, in particular the European call option whose value increases with that of the underlying. Following Kufner [27], this problem may be addressed by the addition of a well chosen secondary weight function. While there are potentially many functions that decrease sufficiently rapidly as x increases to be used as the secondary weight function, we however choose to follow Sanfelici [33] and consider the function

$$\omega_\mu(x) = \begin{cases} 1 & \text{if } \ 0 \leq x \leq x_{max} \\ \\ \left(\dfrac{x_{max}}{x}\right)^\mu & \text{if } \ x > x_{max} \end{cases} ,$$

where $\mu \geq 2$ may be selected large enough to ensure that the required financial contract is admissible and $x_{max} > 0$ is a 'large' value of x . In the case of call options, $\mu = 2$ is sufficient, as for large x the option value increases linearly with the underlying.

Making use of this function, we now define the following weighted spaces that may be used in the derivation of the weak formulation of Problem 1

$$\mathcal{W}_\mu = W^{1,2}(\mathbb{R}^+, \bar{\omega}_\mu) \ \text{ where } \ \bar{\omega}_\mu = \{\omega_\mu^2(x), \omega_\mu^2(x)x^2\}$$

and

$$\mathcal{L}_\mu = L^2(\mathbb{R}^+, \omega_\mu^2) \quad .$$

While the treatment of the valuation problem within the context of the above weighted Sobolev spaces will form the basis for this dissertation, it is not the only manner in which to deal with the degeneracy of the Black-Scholes PDE. As noted previously, a far more common approach is to consider a logarithmic transformation of the reverse time Black-Scholes PDE. Setting

$X = ln(x)$ equation 2.1 becomes

$$-V_t + (r - \frac{1}{2}\sigma^2)V_X + \frac{\sigma^2}{2}V_{XX} = rV \quad .$$

Clearly, since x is defined on $\mathbb{R}^+$, the transformed problem is defined for $X \in \mathbb{R}$. The point of degeneracy ($x = 0$) has therefore been shifted to $X = -\infty$ and may be dealt with through domain truncation. Weighted spaces are therefore not required.

To conclude this section we now present a number of useful properties of these spaces. One of the main aims of these results will be to show that $C_0^\infty(\mathbb{R}^+)$ is dense in $\mathcal{W}_\mu$. The author notes that Achdou and Pironneau [1] demonstrate this result for the space $W^{1,2}(\mathbb{R}^+, X)$ and that the method employed may be investigated as a possible alternative to that presented here.

Lemma 2.1. *The weight function ω_μ is weakly differentiable on $\mathbb{R}^+$ and satisfies*

$$|D^{(1)}[\omega_\mu(x)]\omega_\mu(x)x| \leq \mu\omega_\mu^2(x) \quad \text{for almost all } x \in \mathbb{R}^+ \quad .$$

Proof. Let $\phi \in C_0^\infty(\mathbb{R}^+)$, then

$$\int_{\mathbb{R}^+} \omega_\mu(x)\phi'(x)\mathrm{d}x = \int_0^{x_{max}} \phi'(x)\mathrm{d}x + \int_{x_{max}}^\infty \left(\frac{x_{max}}{x}\right)^\mu \phi'(x)\mathrm{d}x \quad ,$$

applying integration by parts,

$$\begin{aligned} &= [\phi(x)]\Big|_0^{x_{max}} + [\left(\frac{x_{max}}{x}\right)^\mu \phi(x)]\Big|_{x_{max}}^\infty \\ &\quad + \int_{x_{max}}^\infty \mu\left(\frac{x_{max}}{x}\right)^\mu \frac{1}{x}\phi(x)\mathrm{d}x \quad . \\ &= \int_{x_{max}}^\infty \mu\left(\frac{x_{max}}{x}\right)^\mu \frac{1}{x}\phi(x)\mathrm{d}x \\ &= \int_0^{x_{max}} 0\phi(x)\mathrm{d}x + \int_{x_{max}}^\infty \mu\left(\frac{x_{max}}{x}\right)^\mu \frac{1}{x}\phi(x)\mathrm{d}x \quad . \end{aligned}$$

It therefore follows from the definition that ω_μ is weakly differentiable, with a weak derivative given by

$$D^{(1)}[\omega_\mu(x)] = \begin{cases} 0 & \text{if } \ 0 < x \leq x_{max} \\ -\mu \left(\frac{x_{max}}{x}\right)^\mu \frac{1}{x} & \text{if } \ x > x_{max} \end{cases} . \tag{2.3}$$

In order to demonstrate the inequality, we note that it follows from (2.3) that for $0 < x \leq x_{max}$,

$$\begin{aligned} D^{(1)}[\omega_\mu(x)]\omega_\mu(x)x &= \ 0 \\ &\leq \ \mu\omega_\mu^2(x) \end{aligned}$$

and for $x > x_{max}$,

$$\begin{aligned} D^{(1)}[\omega_\mu(x)]\omega_\mu(x)x &= -\mu \left(\frac{x_{max}}{x}\right)^\mu \frac{1}{x} \left(\frac{x_{max}}{x}\right)^\mu x \\ &= -\mu \left(\frac{x_{max}}{x}\right)^{2\mu} \\ &= \ -\mu\omega_\mu^2 \ . \end{aligned}$$

The result now follows by noting that the weak derivatives of ω_μ are equal almost everywhere. □

Lemma 2.2. *Given $u \in \mathcal{W}_\mu$, it follows that for each $n \in \mathbb{N}$, $\tilde{u}_{(n)} \in \mathcal{W}_\mu$, with $supp(\tilde{u}_{(n)}) \Subset \Omega$.*

Proof. It follows immediately from the definitions of $\chi_{[a]}$ and χ_n that $supp(\tilde{u}_{(n)}) \Subset \Omega$. Furthermore, since the derivative of $\chi_{[a]}$ may be bounded by a constant depending only on a, it will suffice to show that $u_{(n)} \in \mathcal{W}_\mu$. To this end,

$$\|u_{(n)}\|^2_{\mathcal{W}_\mu} = \int_{\mathbb{R}^+} \omega_\mu^2(x)|\chi_n(x)u(x)|^2 \mathrm{d}x + \int_{\mathbb{R}^+} \omega_\mu^2(x)x^2|D^{(1)}[\chi_n(x)u(x)]|^2 \mathrm{d}x \ \ ,$$

making use of the product rule for weak derivatives,

$$= \int_{\mathbb{R}^+} \omega_\mu^2(x)|\chi_n(x)u(x)|^2 \mathrm{d}x$$

$$+ \int_{\mathbb{R}^+} \omega_\mu^2(x)x^2|\chi_n'(x)u(x) + \chi_n(x)D^{(1)}[u(x)]|^2\mathrm{d}x \quad ,$$

recalling the boundedness of χ_n,

$$\begin{aligned}
&\leq \int_{\mathbb{R}^+} \omega_\mu^2(x)|u(x)|^2\mathrm{d}x + \int_{\mathbb{R}^+} \omega_\mu^2(x)x^2 \left[|\chi_n'(x)u(x)| + |D^{(1)}[u(x)]|\right]^2 \mathrm{d}x \\
&= \int_{\mathbb{R}^+} \omega_\mu^2(x)|u(x)|^2\mathrm{d}x + \int_{\mathbb{R}^+} \omega_\mu^2(x)x^2|\chi_n'(x)u(x)|^2\mathrm{d}x \\
&\quad + 2\int_\Omega \omega_\mu^2(x)x^2|\chi_n'(x)u(x)|\ |D^{(1)}[u(x)]|\mathrm{d}x \\
&\quad + \int_{\mathbb{R}^+} \omega_\mu^2(x)x^2|D^{(1)}[u(x)]|^2\mathrm{d}x \quad ,
\end{aligned}$$

recalling Lemma 1.18,

$$\begin{aligned}
&\leq \int_{\mathbb{R}^+} \omega_\mu^2(x)|u(x)|^2\mathrm{d}x + C\int_{\mathbb{R}^+} \omega_\mu^2(x)|u(x)|^2\mathrm{d}x \\
&\quad + C\int_{\mathbb{R}^+} \omega_\mu^2(x)x|u(x)|\ |D^{(1)}[u(x)]|\mathrm{d}x \\
&\quad + \int_{\mathbb{R}^+} \omega_\mu^2(x)x^2|D^{(1)}[u(x)]|^2\mathrm{d}x \quad .
\end{aligned}$$

Hölder's inequality then implies that

$$\begin{aligned}
C\int_{\mathbb{R}^+} \omega_\mu^2(x)x|u(x)|\ |D^{(1)}[u(x)]|\mathrm{d}x &\leq C\left(\int_{\mathbb{R}^+} \omega_\mu^2(x)|u(x)|^2\mathrm{d}x\right)^{1/2} \\
&\qquad \left(\int_{\mathbb{R}^+} \omega_\mu^2(x)x^2|D^{(1)}[u(x)]|^2\mathrm{d}x\right)^{1/2} \\
&= C\|u\|_{\mathcal{L}_\mu}|u|_{\mathcal{W}_\mu}
\end{aligned}$$

and hence, we have that

$$\begin{aligned}
\|u_{(n)}\|_{\mathcal{W}_\mu} \leq &\int_{\mathbb{R}^+} \omega_\mu^2(x)|u(x)|^2\mathrm{d}x + C\int_{\mathbb{R}^+} \omega_\mu^2(x)|u(x)|^2\mathrm{d}x \\
&+ C\|u\|_{\mathcal{L}_\mu}|u|_{\mathcal{W}_\mu} + \int_{\mathbb{R}^+} \omega_\mu^2(x)x^2|D^{(1)}[u(x)]|^2\mathrm{d}x \quad . \qquad (2.4)
\end{aligned}$$

Since $u \in \mathcal{W}_\mu$, the righthand side of (2.4) is clearly finite and hence, as required, $u_{(n)} \in \mathcal{W}_\mu$.

□

Lemma 2.3. *Let $u \in \mathcal{W}_\mu$, then given $\epsilon_0 > 0$ there exists a function $\phi \in C_0^\infty(\Omega)$ such that*

$$\|\tilde{u} - \phi\|_{\mathcal{W}_\mu} < \epsilon_0 \quad .$$

Proof. We begin by recalling that since χ is bounded on $\mathbb{R}$, we have $\tilde{u} \in \mathcal{W}_\mu \subset W^{1,2}(\Omega, \bar{\omega}_\mu)$. Furthermore, making use of Lemma 2.2, for each $n \in \mathbb{N}$ we have that $\tilde{u}_{(n)} \in \mathcal{W}_\mu \subset W^{1,2}(\Omega, \bar{\omega}_\mu)$, with $supp(\tilde{u}_{(n)}) \Subset \Omega$.

Now, since $\tilde{u}_{(n)} = \chi_n \tilde{u}$ and $\omega_\mu = 1$ on Ω, $\tilde{u}$ satisfies the conditions of Lemmas 1.22 and 1.23 and hence $(\tilde{u}_{(n)})$ converges to $\tilde{u}$ in $W^{1,2}(\Omega, \bar{\omega}_\mu)$. It therefore follows that given $\epsilon_0 > 0$, there exists $n^* \in \mathbb{N}$ such that

$$\|\tilde{u} - \tilde{u}_{(n^*)}\|_{W^{1,2}(\Omega,\bar{\omega}_\mu)} < \frac{\epsilon_0}{2} \quad .$$

We now consider the mollification

$$\phi_{\epsilon,n^*} = J_\epsilon * \tilde{u}_{(n^*)}$$

and note that Lemma A.26 implies that $\phi_{\epsilon,n^*} \in C_0^\infty(\Omega)$, provided $\epsilon < dist(supp(\tilde{u}_{n^*}), \partial\Omega) \leq max\{\frac{1}{2n^*}, \frac{1}{3}x_{max}\}$. Moreover, since $\tilde{u}_{(n^*)} \in W^{1,2}(\Omega, \bar{\omega}_\mu)$ and $\omega_\mu = 1$ on Ω, we may apply Lemmas A.26 and 1.15 to yield

$$\lim_{\epsilon \to 0^+} \|\tilde{u}_{(n^*)} - \phi_{\epsilon,n^*}\|_{W^{1,2}(\Omega,\bar{\omega}_\mu)} = 0 \quad .$$

It therefore follows that given $\epsilon_0 > 0$, there exists $0 < \epsilon^* < \max\{\frac{1}{2n^*}, \frac{1}{3}x_{max}\}$ such that

$$\|\tilde{u}_{(n^*)} - \phi_{\epsilon^*,n^*}\|_{W^{1,2}(\Omega,\bar{\omega}_\mu)} < \frac{\epsilon_0}{2} \quad .$$

Finally, setting $\phi = \phi_{\epsilon^*,n^*} \in C_0^\infty(\Omega)$, we have

$$\|\tilde{u} - \phi\|_{W^{1,2}(\Omega,\bar{\omega}_\mu)} \leq \|\tilde{u}_{(n^*)} - \phi\|_{W^{1,2}(\Omega,\bar{\omega}_\mu)} + \|\tilde{u} - \tilde{u}_{(n^*)}\|_{W^{1,2}(\Omega,\bar{\omega}_\mu)}$$

$$
\begin{aligned}
&< \frac{\epsilon_0}{2} + \frac{\epsilon_0}{2} \\
&= \epsilon_0 \quad .
\end{aligned}
$$

□

Lemma 2.4. *Let $u \in \mathcal{W}_\mu$, then*

$$\lim_{a \to \infty} \|u - \tilde{u}\|_{\mathcal{W}_\mu} = 0 \quad .$$

Proof. We begin by noting that

$$
\begin{aligned}
\|u - \tilde{u}\|_{\mathcal{W}_\mu} = &\int_0^\infty \omega_\mu^2(x)|u(x) - \tilde{u}(x)|^2 \mathrm{d}x \\
&+ \int_0^\infty \omega_\mu^2(x) x^2 |D^{(1)}[u(x) - \tilde{u}(x)]|^2 \mathrm{d}x \quad ,
\end{aligned}
$$

recalling that $u(x) = \tilde{u}(x)$ for $x \in (0, \frac{1}{3}a)$, $\tilde{u}(x) = 0$ for $x \geq \frac{2}{3}a$, $\omega_\mu(x) = 1$ for $0 < x \leq a$, we have that

$$
\begin{aligned}
= &\int_{a/3}^{2a/3} |u(x) - \tilde{u}(x)|^2 \mathrm{d}x + \int_{2a/3}^\infty \omega_\mu^2(x)|u(x)|^2 \mathrm{d}x \\
&+ \int_{a/3}^{2a/3} x^2 |D^{(1)}[u(x) - \tilde{u}(x)]|^2 \mathrm{d}x \\
&+ \int_{2a/3}^\infty \omega_\mu^2(x) x^2 |D^{(1)}[u(x)]|^2 \mathrm{d}x \quad .
\end{aligned}
$$

Making use of the product rule for weak derivatives, we have that

$$
\begin{aligned}
&\int_{a/3}^{2a/3} x^2 |D^{(1)}[u(x) - \tilde{u}(x)]|^2 \mathrm{d}x \\
&= \int_{a/3}^{2a/3} x^2 |D^{(1)}[u(x)] - \chi'_{[a]}(x)u(x) - \chi_{[a]}(x) D^{(1)}u(x)|^2 \mathrm{d}x \\
&\leq \int_{a/3}^{2a/3} x^2 \left[|(1 - \chi_{[a]}(x)) D^{(1)}u(x)| + |\chi'_{[a]}(x)u(x)| \right]^2 \mathrm{d}x \\
&= \int_{a/3}^{2a/3} x^2 |1 - \chi_{[a]}(x)|^2 |D^{(1)}u(x)|^2 \mathrm{d}x + \int_{a/3}^{2a/3} x^2 |\chi'_{[a]}(x)u(x)|^2 \mathrm{d}x
\end{aligned}
$$

$$+ 2\int_{a/3}^{2a/3} x^2|1-\chi_{[a]}(x)|\;|D^{(1)}u(x)|\;|\chi'_{[a]}(x)u(x)|\mathrm{d}x \quad ,$$

applying Hölder's inequality and recalling that $\chi_{[a]}$ is bounded below by 0,

$$\begin{aligned}&\leq \int_{a/3}^{2a/3} x^2|D^{(1)}u(x)|^2\mathrm{d}x + \int_{a/3}^{2a/3} x^2|\chi'_{[a]}(x)u(x)|^2\mathrm{d}x \\ &\quad + 2\left(\int_{a/3}^{2a/3} x^2|D^{(1)}u(x)|^2\mathrm{d}x\right)^{1/2}\left(\int_{a/3}^{2a/3} x^2|\chi'_{[a]}(x)u(x)|^2\mathrm{d}x\right)^{1/2} \quad ,\end{aligned}$$

recalling the the boundedness condition of $\chi'_{[a]}$,

$$\begin{aligned}&\leq \int_{a/3}^{2a/3} x^2|D^{(1)}u(x)|^2\mathrm{d}x + 4\int_{a/3}^{2a/3} |u(x)|^2\mathrm{d}x \\ &\quad + 2\left(\int_{a/3}^{2a/3} x^2|D^{(1)}u(x)|^2\mathrm{d}x\right)^{1/2}\left(4\int_{a/3}^{2a/3} |u(x)|^2\mathrm{d}x\right)^{1/2} \quad .\end{aligned}$$

Similarly,

$$\begin{aligned}\int_{a/3}^{2a/3} |u(x)-\tilde{u}(x)|^2\mathrm{d}x &= \int_{a/3}^{2a/3} |[1-\chi_{[a]}(x)]u(x)|^2\mathrm{d}x \\ &\leq \int_{a/3}^{2a/3} |u(x)|^2\mathrm{d}x \quad ,\end{aligned}$$

and thus we have that,

$$\begin{aligned}\|u-\tilde{u}\|_{\mathcal{W}_\mu} \leq & \int_{a/3}^{2a/3} |u(x)|^2\mathrm{d}x + \int_{2a/3}^{\infty} \omega_\mu^2(x)|u(x)|^2\mathrm{d}x \\ & + \int_{2a/3}^{\infty} \omega_\mu^2(x)x^2|D^{(1)}[u(x)]|^2\mathrm{d}x \\ & + \int_{a/3}^{2a/3} x^2|D^{(1)}u(x)|^2\mathrm{d}x + 4\int_{a/3}^{2a/3} |u(x)|^2\mathrm{d}x \\ & + 2\left(\int_{a/3}^{2a/3} x^2|D^{(1)}u(x)|^2\mathrm{d}x\right)^{1/2}\left(4\int_{a/3}^{2a/3} |u(x)|^2\mathrm{d}x\right)^{1/2} \quad .\end{aligned}$$

The result now follows by taking the limit as $a \to \infty$. □

Theorem 2.5. *The space $C_0^\infty(\mathbb{R}^+)$ is dense in $\mathcal{W}_\mu$.*

Proof. Let $u \in \mathcal{W}_\mu$, then recall from Lemma 2.4 that

$$\lim_{a\to\infty} \|u - \tilde{u}\|_{\mathcal{W}_\mu} = 0$$

and hence, given $\epsilon > 0$, there exists $x_{max} > 0$, such that

$$\|u - \tilde{u}\|_{\mathcal{W}_\mu} < \frac{\epsilon}{2}. \tag{2.5}$$

Now, for this x_{max}, Lemma 2.3 implies that there exists $\phi \in C_0^\infty(\Omega) \subset C_0^\infty(\mathbb{R}^+)$ such that

$$\|\tilde{u} - \phi\|_{W^{1,2}(\Omega,\omega_\mu)} < \frac{\epsilon}{2} \quad . \tag{2.6}$$

Noting that both $supp(\tilde{u})$ and $supp(\phi)$ are contained within Ω, it follows from (2.5) and (2.6) that

$$\begin{aligned}\|u - \phi\|_{\mathcal{W}_\mu} &\leq \|u - \tilde{u}\|_{\mathcal{W}_\mu} + \|\tilde{u} - \phi\|_{\mathcal{W}_\mu} \\ &= \|u - \tilde{u}\|_{\mathcal{W}_\mu} + \|\tilde{u} - \phi\|_{W^{1,2}(\Omega,\omega_\mu)} \\ &< \epsilon \quad .\end{aligned}$$

□

This result may seem rather surprising as the analogous result does not hold in the unweighted case(see the Appendix). This apparent discrepancy may be explained by the presence of the weight x^2 in the second term of the $\mathcal{W}_\mu$-norm. In the above proofs, we relied on the cut-off function to χ_n to ensure that our approximating functions vanish within some neighbourhood of 0. Since χ_n increases from 0 to 1 over the interval $[1/2n, 1/n]$, it is clear that as n increases, the derivative of χ_n becomes infinitely large within this interval. It is this growth that hinders convergence in unweighted norms. The growth is however controlled within the weighted $\mathcal{W}_\mu$-norm by the presence on the fast decaying x^2 term in the weight function.

Theorem 2.6. *The space $C_0^\infty(\mathbb{R}^+)$ is dense in $\mathcal{L}_\mu$ and $L^2(\mathbb{R}^+, x^2\omega_\mu^2)$.*

Proof. The proof of this result follows similarly to Theorem 2.5, by noting that results analogous to Lemmas 2.2, 2.3 and 2.4 hold for these spaces. □

Theorem 2.7. *The spaces $\mathcal{L}_\mu$ and $\mathcal{W}_\mu$ are separable Hilbert Spaces.*

Proof. We have already demonstrated that $\mathcal{L}_\mu$ and $\mathcal{W}_\mu$ are complete inner product spaces and hence it only remains to show that they are separable. To this end, we will adapt work by Adams and Fournier [2] to demonstrate the separability of two particular weighted Lebesgue spaces and then make use of these results to demonstrate the separability of $\mathcal{W}_\mu$.

Following Adams and Fournier [2], we begin by setting

$$\Omega_n = (1/n, n) \text{ for each } n \in \mathbb{N}$$

and then denote by P_n the collection of all polynomials over Ω_n with rational coefficients. It is well known that for each $n \in \mathbb{N}$, P_n is countable and hence so is $\bigcup_{n=1}^\infty P_n$.

We now consider functions $u \in \mathcal{L}_\mu$ and $v \in L^2(\mathbb{R}^+, x^2\omega_\mu^2)$ and notice that Theorem 2.6 implies that for each $\epsilon > 0$, there exist functions ϕ,$\varphi \in C_0^\infty(\mathbb{R}^+)$ such that

$$\|u - \phi\|_{\mathcal{L}_\mu} < \frac{\epsilon}{2} \tag{2.7}$$

and

$$\|v - \varphi\|_{L^2(\mathbb{R}^+} < \frac{\epsilon}{2} \quad . \tag{2.8}$$

Since ϕ and φ are continuous functions and have compact support within $\mathbb{R}^+$, it follows that there exists an $n \in \mathbb{N}$ such that $\phi, \varphi \in C(\Omega_n)$. Making use of the Stone-Weierstrass Theorem to yield the density of P_n in $C(\Omega_n)$, it therefore follows that given $\epsilon > 0$, there exist functions $\upsilon, \nu \in P_n$ such that

$$\|\phi - \upsilon\|_{\infty,\Omega_n} < \frac{\epsilon}{2}\left(\int_{\Omega_n} \omega_\mu^2(x)\mathrm{d}x\right)^{-1/2} \tag{2.9}$$

and

$$\|\varphi - \nu\|_{\infty,\Omega_n} < \frac{\epsilon}{2}\left(\int_{\Omega_n} x^2\omega_\mu^2(x)\mathrm{d}x\right)^{-1/2} \quad . \tag{2.10}$$

Making use of equations (2.7) and (2.9) it then follows that

$$\begin{aligned}
\|u-\nu\|_{\mathcal{L}_\mu} &\leq \|u-\phi\|_{\mathcal{L}_\mu} + \|\phi-\nu\|_{\mathcal{L}_\mu} \\
&= \|u-\phi\|_{\mathcal{L}_\mu} + \left(\int_{\mathbb{R}^+} \omega_\mu^2(x)|\phi(x)-\nu(x)|^2\right)^{1/2} \\
&= \|u-\phi\|_{\mathcal{L}_\mu} + \left(\int_{\Omega_n} \omega_\mu^2(x)|\phi(x)-\nu(x)|^2\right)^{1/2} \\
&\leq \|u-\phi\|_{\mathcal{L}_\mu} + \|\phi-v\|_{\infty,\Omega_n}\left(\int_{\Omega_n} \omega_\mu^2(x)\right)^{1/2} \\
&< \frac{\epsilon}{2} + \frac{\epsilon}{2}\left(\int_{\Omega_n} \omega_\mu^2(x)\mathrm{d}x\right)^{-1/2}\left(\int_{\Omega_n} \omega_\mu^2(x)\mathrm{d}x\right)^{1/2} \\
&= \epsilon \quad ,
\end{aligned}$$

so that $\bigcup_{n=1}^{\infty} P_n$ is dense in $\mathcal{L}_\mu$ and thus $\mathcal{L}_\mu$ is separable. Making use of equations (2.8) and (2.10) we similarly have that $L^2(\mathbb{R}^+, x^2\omega_\mu^2)$ is separable.

To complete the result, we follow Kufner [27] in noting that $\mathcal{W}_\mu$ can be viewed as a subspace of the Cartesian product $\mathcal{L}_\mu \times L^2(\mathbb{R}^+, x^2\omega_\mu^2)$ and hence must also be separable. □

To conclude this section, we present the following characterisation of the dual space $\mathcal{W}_\mu^*$ that is adapted from a similar result for unweighted spaces as presented by Evans [19].

Theorem 2.8. *A mapping $f : \mathcal{W}_\mu \to \mathbb{R}$ is a member of $\mathcal{W}_\mu^*$ if and only if there exist functions $f_0, f_1 \in L^2(\mathbb{R}^+)$ such that for every $u \in \mathcal{W}_\mu$*

$$f(u) = \int_{\mathbb{R}^+} \omega_\mu(x) f_0(x) u(x) \mathrm{d}x + \int_{\mathbb{R}^+} \omega_\mu(x) x f_1(x) D^{(1)}[u(x)] \mathrm{d}x \quad .$$

Proof. Suppose that $f \in \mathcal{W}_\mu^*$, it then follows from the Riesz representation theorem (see Kreyszig [26] p. 188) that there exists a unique $v \in \mathcal{W}_\mu$ that depends only on f, such that

$$\begin{aligned} f(u) &= (u, v)_{\mathcal{W}_\mu} \\ &= \int_{\mathbb{R}^+} \omega_\mu^2(x) u(x) v(x) \mathrm{d}x + \int_{\mathbb{R}^+} \omega_\mu^2(x) x^2 D^{(1)}[u(x)] D^{(1)}[v(x)] \mathrm{d}x \\ &= \int_{\mathbb{R}^+} \omega_\mu(x) u(x) \left[\omega_\mu(x) v(x)\right] \mathrm{d}x \\ &\quad + \int_{\mathbb{R}^+} \omega_\mu(x) x D^{(1)}[u(x)] \left[\omega_\mu(x) x D^{(1)}[v(x)]\right] \mathrm{d}x \quad . \end{aligned}$$

Since $v \in \mathcal{W}_\mu$, we clearly have that

$$f_0 = \omega_\mu(x) v(x) \in L^2(\mathbb{R}^+)$$

and

$$f_1 = \omega_\mu(x) x D^{(1)}[v(x)] \in L^2(\mathbb{R}^+) \quad .$$

Conversely, suppose that there exist functions $f_0, f_1 \in L^2(\mathbb{R}^+)$ such that

$$f(u) = \int_{\mathbb{R}^+} \omega_\mu(x) f_0(x) u(x) \mathrm{d}x + \int_{\mathbb{R}^+} \omega_\mu(x) x f_1(x) D^{(1)}[u(x)] \mathrm{d}x \quad .$$

Since f is clearly linear, it remains to show that it is bounded. To this end, we may apply Hölder's inequality to obtain

$$\begin{aligned} f(u) \leq &\left(\int_{\mathbb{R}^+} f_0^2(x) \mathrm{d}x\right)^{1/2} \left(\int_{\mathbb{R}^+} \omega_\mu^2(x) u^2(x) \mathrm{d}x\right)^{1/2} \\ &+ \left(\int_{\mathbb{R}^+} f_1^2(x) \mathrm{d}x\right)^{1/2} \left(\int_{\mathbb{R}^+} \omega_\mu^2(x) x^2 \left|D^{(1)}[u(x)]\right|^2 \mathrm{d}x\right)^{1/2} \end{aligned}$$

$$\leq max\{\|f_0\|_{L^2(\mathbb{R}^+)}, \|f_1\|_{L^2(\mathbb{R}^+)}\}\|u\|_{\mathcal{W}_\mu} \quad .$$

□

2.2 Weak formulation

In this section we will derive the weak formulation of Problem 1. To this end, it will prove convenient to follow the convention of Evans [19] and view the option value $V(x,t)$ not as a function of the underlying and time, but rather as a mapping of time into the space $\mathcal{W}_\mu$. In other words, for each fixed t, we view V as a function of x that lies within $\mathcal{W}_\mu$. For notational convenience, we will write $V(x)$ in instances where the statement in question does not make reference to time variable. Now, following Evans [19]

Definition 2.9. Given times $0 \leq t_1 < t_2$, $p \geq 1$ and a real valued Banach space X, we define the space $L^p(t_1, t_2, X)$ to be the collection of measurable functions

$$u(t) : [t_1, t_2] \to X$$

that satisfy

$$\|u\|_{L^p(t_1,t_2,X)} = \left(\int_{t_1}^{t_2} \|u(t)\|_X^p \mathrm{d}t\right)^{1/p} < \infty \quad .$$

Bearing this convention in mind, standard procedure indicates that we should begin by multiplying equation (2.1) by an arbitrary test function $\phi \in C_0^\infty(\mathbb{R}^+)$ and then integrate the resultant identity over $\mathbb{R}^+$. We however wish to ensure that the natural integral norm that arises during our derivation is that of the space $\mathcal{W}_\mu$ and hence rather multiply equation (2.1) by $\omega_\mu^2\phi$ and then integrate to obtain

$$r\int_{\mathbb{R}^2} \omega_\mu^2(x)V(x)\phi(x)\mathrm{d}x = -\int_{\mathbb{R}^+} \omega_\mu^2(x)V_t(x)\phi(x)\mathrm{d}x + \int_{\mathbb{R}^+} r\omega_\mu^2(x)xV_x(x)\phi(x)\mathrm{d}x$$

$$+\frac{\sigma^2}{2}\int_{\mathbb{R}^+}\omega_\mu^2(x)x^2V_{xx}(x)\phi(x)\mathrm{d}x \quad . \tag{2.11}$$

Since classical and distributional derivatives agree almost everywhere, we may instead write

$$\begin{aligned} r\int_{\mathbb{R}^2}\omega_\mu^2(x)V(x)\phi(x)\mathrm{d}x = & -\int_{\mathbb{R}^+}\omega_\mu^2(x)[V]_t(x)\phi(x)\mathrm{d}x \\ & +\int_{\mathbb{R}^+}r\omega_\mu^2(x)xV_x(x)\phi(x)\mathrm{d}x \\ & +\frac{\sigma^2}{2}\int_{\mathbb{R}^+}\omega_\mu^2(x)x^2V_{xx}(x)\phi(x)\mathrm{d}x \quad . \end{aligned} \tag{2.12}$$

where $[V]_t$ denotes the distributional derivative of the option value V with respect to time.

It is now easily verified that the functions $\omega_\mu, x, D^{(1)}[V(x)]$ and ϕ satisfy the conditions of Lemma A.31 and hence we may apply the product rule for weak derivatives to obtain

$$\begin{aligned} D^{(1)}\left[\frac{1}{2}\sigma^2\omega_\mu^2(x)x^2D^{(1)}[V(x)]\phi(x)\right] = & \ \sigma^2\omega_\mu(x)D^{(1)}[\omega_\mu(x)]x^2D^{(1)}[V(x)]\phi(x) \\ & + \ \sigma^2\omega_\mu^2(x)xD^{(1)}[V(x)]\phi(x) \\ & + \ \frac{1}{2}\sigma^2\omega_\mu^2(x)x^2D^{(2)}[V(x)]\phi(x) \\ & + \ \frac{1}{2}\sigma^2\omega_\mu^2(x)x^2D^{(1)}[V(x)]D^{(1)}[\phi(x)] \ , \end{aligned}$$

or

$$\begin{aligned} \frac{1}{2}\sigma^2\omega_\mu^2(x)x^2D^{(2)}[V(x)]\phi(x) = & \ D^{(1)}\left[\frac{1}{2}\sigma^2\omega_\mu^2(x)x^2D^{(1)}[V(x)]\phi(x)\right] \\ & - \ \sigma^2\omega_\mu(x)D^{(1)}[\omega_\mu(x)]x^2D^{(1)}[V(x)]\phi(x) \\ & - \ \sigma^2\omega_\mu^2(x)xD^{(1)}[V(x)]\phi(x) \\ & - \ \sigma^2\omega_\mu^2(x)x^2D^{(1)}[V(x)]D^{(1)}[\phi(x)] \quad . \end{aligned}$$

Substituting into (2.12), it then follows that

$$
\begin{aligned}
\int_{\mathbb{R}^+} r\omega_\mu^2(x)V(x)\phi(x)\mathrm{d}x = &-\int_{\mathbb{R}^+} \omega_\mu^2(x)[V]_t(x)\phi(x)\mathrm{d}x \\
&+\int_{\mathbb{R}^+} r\omega_\mu^2(x)xD^{(1)}[V(x)]\phi(x)\mathrm{d}x \\
&-\int_{\mathbb{R}^+} \sigma^2\omega_\mu(x)D^{(1)}[\omega_\mu(x)]x^2D^{(1)}[V(x)]\phi(x)\mathrm{d}x \\
&-\int_{\mathbb{R}^+} \sigma^2\omega_\mu^2(x)xD^{(1)}[V(x)]\phi(x)\mathrm{d}x \\
&-\int_{\mathbb{R}^+} \frac{1}{2}\sigma^2\omega_\mu^2(x)x^2D^{(1)}[V(x)]D^{(1)}[\phi(x)]\mathrm{d}x \\
&+\int_{\mathbb{R}^+} D^{(1)}\left[\frac{1}{2}\sigma^2\omega_\mu^2(x)x^2D^{(1)}[V(x)]\phi(x)\right]\mathrm{d}x \quad .
\end{aligned}
$$

and hence,

$$
\begin{aligned}
\int_{\mathbb{R}^+} \omega_\mu^2(x)[V]_t(x)\phi(x)\mathrm{d}x + &\int_{\mathbb{R}^+} \frac{1}{2}\sigma^2\omega_\mu^2(x)x^2D^{(1)}[V(x)]D^{(1)}[\phi(x)]\mathrm{d}x \\
&+\int_{\mathbb{R}^+} A(x)\omega_\mu^2(x)xD^{(1)}[V(x)]\phi(x)\mathrm{d}x \\
&+\int_{\mathbb{R}^+} r\omega_\mu^2(x)V(x)\phi(x)\mathrm{d}x \\
&-\int_{\mathbb{R}^+} D^{(1)}\left[\frac{1}{2}\sigma^2\omega_\mu^2(x)x^2D^{(1)}[V(x)]\phi(x)\right]\mathrm{d}x \\
&= 0 \quad ,
\end{aligned}
$$

where

$$
A(x) = \left[\sigma^2\left(1+\frac{D^{(1)}[\omega_\mu(x)]}{\omega_\mu(x)}x\right) - r\right] \quad .
$$

Recalling that $\frac{1}{2}\sigma^2\omega_\mu^2(x)x^2D^{(1)}[V(x)]\phi(x)$ is weakly differentiable on $\mathbb{R}^+$, Lemma A.29 implies that for each $n \in \mathbb{N}$, it has a version absolutely continuous on $[1/n, n]$. We may therefore apply the fundamental theorem of calculus to obtain

$$\begin{aligned}
&\int_{\mathbb{R}^+} D^{(1)}\left[\frac{1}{2}\sigma^2\omega_\mu^2(x)x^2D^{(1)}[V(x)]\phi(x)\right]\mathrm{d}x \\
&= \lim_{n\to\infty}\int_{1/n}^{n} D^{(1)}\left[\frac{1}{2}\sigma^2\omega_\mu^2(x)x^2D^{(1)}[V(x)]\phi(x)\right]\mathrm{d}x \\
&= \lim_{n\to\infty}\left(\frac{1}{2}\sigma^2\omega_\mu^2(n)n^2D^{(1)}[V(n)]\phi(x)\right) \\
&\quad - \lim_{n\to\infty}\left(\frac{1}{2}\sigma^2\omega_\mu^2(1/n)\left(1/n\right)^2D^{(1)}[V(1/n)]\phi(1/n)\right) \quad .
\end{aligned}$$

Since $\phi \in C_0^\infty(\mathbb{R}^+)$, it must vanish for large enough n and within some neighbourhood of 0 and hence there exists a $N \in \mathbb{N}$ such that $\phi(n) = \phi(1/n) = 0$ for each $n \geq N$. It therefore follows that

$$\begin{aligned}
&\lim_{n\to\infty}\left(\frac{1}{2}\sigma^2\omega_\mu^2(n)n^2D^{(1)}[V(n)]\phi(n)\right) \\
&= \lim_{n\to\infty}\left(\frac{1}{2}\sigma^2\omega_\mu^2(1/n)\left(1/n\right)^2D^{(1)}[V(1/n)]\phi(1/n)\right) \\
&= 0
\end{aligned}$$

and hence

$$\begin{aligned}
&\int_{\mathbb{R}^+}\omega_\mu^2(x)[V]_t(x)\phi(x)\mathrm{d}x \ + \int_{\mathbb{R}^+}\frac{1}{2}\sigma^2\omega_\mu^2(x)x^2D^{(1)}[V(x)]D^{(1)}[\phi(x)]\mathrm{d}x \\
&+ \int_{\mathbb{R}^+}A(x)\omega_\mu^2(x)xD^{(1)}[V(x)]\phi(x)\mathrm{d}x \ + \int_{\mathbb{R}^+}r\omega_\mu^2(x)V(x)\phi(x)\mathrm{d}x \\
&= 0 \quad .
\end{aligned} \tag{2.13}$$

This expression naturally leads us to define the bilinear form

$$\begin{aligned}
\mathcal{A}_\mu(u,v) = &\int_{\mathbb{R}^+}\frac{1}{2}\sigma^2\omega_\mu^2(x)x^2D^{(1)}[u(x)]D^{(1)}[v(x)]\mathrm{d}x \\
&+ \int_{\mathbb{R}^+}A(x)\omega_\mu^2(x)xD^{(1)}[u(x)]v(x)\mathrm{d}x \ + \int_{\mathbb{R}^+}r\omega_\mu^2(x)u(x)v(x)\mathrm{d}x \quad ,
\end{aligned}$$

for $u, v \in \mathcal{W}_\mu$. Under this notation, equation (2.13) becomes

$$([V(t)]_t, \phi)_{\mathcal{L}_\mu} + \mathcal{A}_\mu(V(t), \phi) = 0 \quad \forall\, \phi \in C_0^\infty(\mathbb{R}^+) \quad . \tag{2.14}$$

We now wish to show that this identity is still valid if ϕ is replaced by an arbitrary function $u \in \mathcal{W}_\mu$. To this end we give the following important property of the bilinear form $\mathcal{A}_\mu$.

Lemma 2.10. *The bilinear form $\mathcal{A}_\mu$ is continuous on $\mathcal{W}_\mu$, i.e. there exists a constant $\gamma > 0$ such that for $u, v \in \mathcal{W}_\mu$*

$$|\mathcal{A}_\mu(u, v)| \le \gamma \|u\|_{\mathcal{W}_\mu} \|v\|_{\mathcal{W}_\mu} \quad .$$

Proof. Setting

$$A(x) = \sigma^2 \left(1 + \frac{\omega_\mu'(x)}{\omega_\mu(x)} x\right) - r$$

It follows that

$$|\mathcal{A}_\mu(\phi, \varphi)| \le \frac{\sigma^2}{2} \int_{\mathbb{R}^+} \omega_\mu^2(x) x^2 |D^{(1)} u(x)| |D^{(1)} v(x)| \mathrm{d}x + \int_{\mathbb{R}^+} A(x) \omega_\mu^2(x) x |D^{(1)} u(x)| |v(x)| \mathrm{d}x + \int_{\mathbb{R}^+} r \omega_\mu^2(x) |u(x)| |v(x)| \mathrm{d}x \quad ,$$

making use of Hölder's inequality,

$$\begin{aligned}
&\le \frac{\sigma^2}{2} \left(\int_{\mathbb{R}} \omega_\mu^2(x) x^2 |D^{(1)} u(x)|^2 \mathrm{d}x \right)^{1/2} \left(\int_{\mathbb{R}} \omega_\mu^2(x) x^2 |D^{(1)} v(x)|^2 \mathrm{d}x \right)^{1/2} \\
&\quad + \|A(x)\|_\infty \left(\int_{\mathbb{R}} \omega_\mu^2(x) x^2 |D^{(1)} u(x)|^2 \mathrm{d}x \right)^{1/2} \|v\|_{\mathcal{L}_\mu} + r \|u\|_{\mathcal{L}_\mu} \|v\|_{\mathcal{L}_\mu} \\
&\le \frac{\sigma^2}{2} |u|_{\mathcal{W}_\mu} |v|_{\mathcal{W}_\mu} + \|A(x)\|_\infty |u|_{\mathcal{W}_\mu} \|v\|_{\mathcal{L}_\mu} + r \|u\|_{\mathcal{L}_\mu} \|v\|_{\mathcal{L}_\mu} \\
&\le \frac{\sigma^2}{2} \|u\|_{\mathcal{W}_\mu} \|v\|_{\mathcal{W}_\mu} + \|A(x)\|_\infty \|u\|_{\mathcal{W}_\mu} \|v\|_{\mathcal{W}_\mu} + r \|u\|_{\mathcal{W}_\mu} \|v\|_{\mathcal{W}_\mu} \quad .
\end{aligned}$$

The result now follows by setting $\gamma = 3 \max\left(\frac{\sigma^2}{2}, \|A(x)\|_\infty, r\right)$

□

We now recall that due to Theorem 2.5, each $u \in \mathcal{W}_\mu$ may be approximated in $\mathcal{W}_\mu$ (and hence in $\mathcal{L}_\mu$) by a sequence of functions in $C_0^\infty(\mathbb{R}^+)$. It therefore follows that by making use of the above lemma, equation (2.14) may be extended by continuity to yield

$$([V(t)]_t, u)_{\mathcal{L}_\mu} + \mathcal{A}_\mu(V(t), u) = 0 \quad \forall\ u \in \mathcal{W}_\mu \quad . \tag{2.15}$$

We now notice that this equation places far weaker restrictions on the option value V than does Problem 1. Indeed, while Problem 1 requires V to be twice classically differentiable with respect to the underlying, equation (2.10) only requires that $V \in \mathcal{W}_\mu$. We furthermore note that the definition of the bilinear form and the requirement for it be continuous with respect to our chosen Sobolev space clearly indicate why the original weight function $\omega^* = \{1, x^2\}$ arises naturally and is suitable for our investigation.

The requirement that $V \in \mathcal{W}_\mu$ gives us direction as to in which spaces we should search for the solution to the weak formulation of Problem 1. In continuation with this theme we present the following results.

Lemma 2.11. *There exist functions $f_1, f_2 \in L^2(\mathbb{R}^+)$ such that for every $\phi \in C_0^\infty(\mathbb{R}^+)$,*

$$([V]_t, \phi)_{\mathcal{L}_\mu} = \int_{\mathbb{R}^+} \omega_\mu(x) f_1(x)\phi(x)\mathrm{d}x + \int_{\mathbb{R}^+} \omega_\mu(x) x f_2(x) D^{(1)}[\phi(x)]\mathrm{d}x \quad .$$

Proof. We begin by noting that equation (2.1) may be reformulated to yield

$$\begin{aligned} [V]_t &= rxV_x + \frac{\sigma^2}{2}x^2V_{xx} - rV \\ &= \left(r - \sigma^2\right) xV_x + \frac{\sigma^2}{2}D^{(1)}[x^2V_x] - rV \quad . \end{aligned}$$

Making use of this fact, it follows that

$$\begin{aligned} ([V]_t, \phi)_{\mathcal{L}_\mu} &= \int_{\mathbb{R}^+} \omega_\mu^2(x)V_t(x)\phi(x)\mathrm{d}x \\ &= \int_{\mathbb{R}^+} \omega_\mu^2(x) \left[\left(r - \sigma^2\right) xV_x(x) + \frac{\sigma^2}{2}D^{(1)}[x^2V_x(x)] - rV(x)\right] \phi(x)\mathrm{d}x \end{aligned}$$

$$
\begin{aligned}
&= \int_{\mathbb{R}^+} \omega_\mu(x) \left[\omega_\mu(x)\left(r-\sigma^2\right) x V_x(x) + \omega_\mu(x) r V(x)\right] \phi(x) \mathrm{d}x \\
&\quad + \int_{\mathbb{R}^+} \frac{\sigma^2}{2} \omega_\mu^2(x) \frac{d}{dx}[x^2 V_x(x)] \phi(x) \mathrm{d}x \quad . \qquad (2.16)
\end{aligned}
$$

Noting that the functions $x^2 V_x(x), \phi(x)$ and $\omega_\mu(x)$ satisfy the conditions of A.31, we may apply the product rule for weak derivatives to yield

$$
\begin{aligned}
D^{(1)}\left[x^2 V_x(x) \omega_\mu^2(x) \phi(x)\right] &= D^{(1)}\left[x^2 V_x(x)\right] \omega_\mu^2(x) \phi(x) \\
&\quad + 2\omega_\mu(x) D^{(1)}\left[\omega_\mu(x)\right] x^2 V_x(x) \phi(x) \\
&\quad + \omega_\mu^2(x) x^2 V_x(x) \phi_x(x) \ ,
\end{aligned}
$$

and hence

$$
\begin{aligned}
\int_{\mathbb{R}^+} \frac{\sigma^2}{2} \omega_\mu^2(x) \frac{d}{dx}[x^2 V_x(x)] \phi(x) \mathrm{d}x &= \int_{\mathbb{R}^+} D^{(1)}\left[x^2 V_x(x) \omega_\mu^2(x) \phi(x)\right] \mathrm{d}x \\
&\quad - \int_{\mathbb{R}^+} 2\omega_\mu(x) D^{(1)}\left[\omega_\mu(x)\right] x^2 V_x(x) \phi(x) \\
&\quad + \omega_\mu^2(x) x^2 V_x(x) \phi_x(x) \mathrm{d}x \quad . \qquad (2.17)
\end{aligned}
$$

We now notice that since the function $x^2 V_x(x) \omega_\mu^2(x) \phi(x)$ is weakly differentiable on $\mathbb{R}^+$, it follows from Lemma A.29 that for each $n \in \mathbb{N}$ it has a version which is absolutely continuous on $[1/n, n]$ and hence

$$
\begin{aligned}
\int_{\mathbb{R}^+} D^{(1)}\left[x^2 V_x(x) \omega_\mu^2(x) \phi(x)\right] \mathrm{d}x &= \lim_{n\to\infty} \int_{1/n}^{n} D^{(1)}\left[x^2 V_x(x) \omega_\mu^2(x) \phi(x)\right] \mathrm{d}x \\
&= \lim_{n\to\infty} \left[x^2 V_x(x) \omega_\mu^2(x) \phi(x)\right]_{1/n}^{n} \ .
\end{aligned}
$$

This limit clearly vanishes since $\phi \in C_0^\infty(\mathbb{R}^+)$. Equation (2.17) therefore becomes

$$
\begin{aligned}
&\int_{\mathbb{R}^+} \frac{\sigma^2}{2} \omega_\mu^2(x) \frac{d}{dx}[x^2 V_x(x)] \phi(x) \mathrm{d}x \\
&= -\int_{\mathbb{R}^+} 2\omega_\mu(x) D^{(1)}\left[\omega_\mu(x)\right] x^2 V_x(x) \phi(x) \\
&\quad + \omega_\mu^2(x) x^2 V_x(x) \phi_x(x) \mathrm{d}x \ ,
\end{aligned}
$$

applying Lemma 2.1

$$
\begin{aligned}
&\leq C \int_{\mathbb{R}^+} \omega_\mu^2(x) x V_x(x)\phi(x) + \omega_\mu^2(x) x^2 V_x(x)\phi_x(x)\mathrm{d}x \\
&= C \int_{\mathbb{R}^+} \omega_\mu(x)\left[\omega_\mu(x) x V_x(x)\right]\phi(x) \\
&\quad + \int_{\mathbb{R}^+} \omega_\mu(x) x \left[\omega_\mu(x) x V_x(x)\right]\phi_x(x)\mathrm{d}x \quad .
\end{aligned}
$$

Combining this equation with (2.16), we have that

$$
\begin{aligned}
([V]_t, \phi)_{\mathcal{L}_\mu} = &\int_{\mathbb{R}^+} \omega_\mu(x)\left[\left(r - \sigma^2 + 1\right)\omega_\mu(x) x V_x(x) + r\omega_\mu(x) V(x)\right]\phi(x)\mathrm{d}x \\
&+ \int_{\mathbb{R}^+} \omega_\mu(x) x \left[\omega_\mu(x) x V_x(x)\right]\phi_x(x)\mathrm{d}x \quad .
\end{aligned}
$$

The result now follows by setting

$$
f_1(x) = \left(r - \sigma^2 + 1\right)\omega_\mu(x) x V_x(x) + \omega_\mu(x) r V(x)
$$

and

$$
f_2(x) = \omega_\mu(x) x V_x(x) \quad .
$$

□

Lemma 2.12. *There exist functions $f_1, f_2 \in L^2(\mathbb{R}^+)$ such that for each $u \in \mathcal{W}_\mu$ we may write*

$$
([V]_t, u)_{\mathcal{L}_\mu} = \int_{\mathbb{R}^+} \omega_\mu(x) f_1(x) u(x)\mathrm{d}x + \int_{\mathbb{R}^+} \omega_\mu(x) x f_2(x) D^{(1)}[u(x)]\mathrm{d}x \quad .
$$

Proof. We begin by recalling that each for each $u \in \mathcal{W}_\mu$, there exists a sequence (ϕ_n) in $C_0^\infty(\mathbb{R}^+)$ such that $\phi_n \to u$ in $\mathcal{W}_\mu$ (and hence also in$\mathcal{L}_\mu$). Making use of the fact that the inner product $(.,.)_{\mathcal{L}_\mu}$ is continuous (see Kreyszig [26] p.138) we therefore have that

$$
([V]_t, u)_{\mathcal{L}_\mu} = \lim_{n\to\infty} ([V]_t, \phi_n)_{\mathcal{L}_\mu} \quad .
$$

It then follows from Lemma 2.11 that there exist functions $f_1, f_2 \in L^2(\mathbb{R}^+)$ such that

$$([V]_t, u)_{\mathcal{L}_\mu} = \lim_{n\to\infty} \left[\int_{\mathbb{R}^+} \omega_\mu(x) f_1(x)\phi_n(x)\mathrm{d}x + \int_{\mathbb{R}^+} \omega_\mu(x) x f_2(x) D^{(1)}[\phi_n(x)]\mathrm{d}x \right] \ . \tag{2.18}$$

In order to evaluate the limit on the right-hand side of this expression, we note that due to Hölder's inequality we have that

$$\begin{aligned} &\lim_{n\to\infty} \left| \int_{\mathbb{R}^+} \omega_\mu(x) f_1(x)\phi_n(x)\mathrm{d}x - \int_{\mathbb{R}^+} \omega_\mu(x) f_1(x) u(x)\mathrm{d}x \right| \\ &\leq \lim_{n\to\infty} \int_{\mathbb{R}^+} |\omega_\mu(x) f_1(x)| \, |\phi_n(x) - u(x)| \, \mathrm{d}x \\ &\leq \lim_{n\to\infty} \|f_1\|_{L^2(\mathbb{R}^+)} \|\phi_n - u\|_{\mathcal{L}_\mu} \\ &= 0 \ . \end{aligned}$$

Similarly

$$\begin{aligned} &\lim_{n\to\infty} \left| \int_{\mathbb{R}^+} \omega_\mu(x) x f_2(x) D^{(1)}[\phi_n(x)]\mathrm{d}x - \int_{\mathbb{R}^+} \omega_\mu(x) x f_2(x) D^{(1)}[u(x)]\mathrm{d}x \right| \\ &\leq \lim_{n\to\infty} \int_{\mathbb{R}^+} |\omega_\mu(x) x f_2(x)| \left| D^{(1)}[\phi_n(x) - u(x)] \right| \mathrm{d}x \\ &\leq \lim_{n\to\infty} \|f_2\|_{L^2(\mathbb{R}^+)} |\phi_n - u|_{\mathcal{W}_\mu} \\ &= 0 \ . \end{aligned}$$

Applying these limits to the right-hand side of equation (2.18) then yields the desired result. □

We now notice that the inner product $(.,.)_{\mathcal{L}_\mu}$ may be viewed as a mapping from the space $\mathcal{W}_\mu$ into $\mathbb{R}$ and hence due to Lemma 2.12, Theorem 2.8 implies that $([V]_t, .)_{\mathcal{L}_\mu}$ is a member of the dual space $\mathcal{W}_\mu^*$. The Riesz representation theorem implies that $[V]_t$ is unique to this member of $\mathcal{W}_\mu^*$ and hence it follows that it would be natural that the weak solution to Problem 1 should be such that $[V]_t \in \mathcal{W}_\mu^*$. Bearing this in mind, we have the following weak formulation of Problem 1.

Problem 2. *For each* $m = 0, 1, ..., N-1$*, find* $V \in L^2(t_m, t_{m+1}, \mathcal{W}_\mu) \cap C(t_m, t_{m+1}, \mathcal{L}_\mu)$ *with* $[V]_t \in L^2(t_m, t_{m+1}, \mathcal{W}^*_\mu)$*, such that*

$$([V]_t, u)_{\mathcal{L}_\mu} + \mathcal{A}_\mu(V, u) = 0 \quad \forall\, u \in \mathcal{W}_\mu \tag{2.19}$$

and

$$V(t_m) = \begin{cases} (x-K)^+ \chi_{l,u}(x) & \text{if} \quad m = 0 \\ \\ V(x, t_m^-)\chi_{l,u}(x) & \text{if} \quad m = 1, 2, ..., N-1 \end{cases} .$$

We note that the initial conditions $V(t_m) = V(x, t_m^-)\chi_{l,u}(x)$ for $m = 1, 2, ..., N-1$ ensure continuity of the solution in all cases except when the barrier is it. In this event, the solution immediately becomes zero and remains so for the remainder of the time intervals.

2.3 Existence and Uniqueness

In this section we will make use of a well known result due to Zeidler [40] (similar results are also presented by Brezis [11] and Evans [19]) to demonstrate that Problem 2 has a unique solution.

Definition 2.13. (Zeidler [40])
Let V,H be separable, real Hilbert spaces such that V is dense in H and we have the following chain of continuous embeddings

$$V \hookrightarrow H \hookrightarrow V^* \quad .$$

We then call (V, H, V^*) an evolution triple.

Theorem 2.14. *(Zeidler [40])*
Given $u_0 \in H$ *and* $0 < T < \infty$*, suppose that the following conditions are satisfied*

1. (V, H, V^*) *is an evolution triple, with* $dim(V) = \infty$.

2. *The mapping* $A : V \times V \to \mathbb{R}$ *is bilinear, continuous and coercive.*

3. *There exists a basis* $(w_1, w_2, ...)$ *in* V *and a sequence* $(u_{0,n})$ *in* H*, such that*

$$u_{0,n} \in span\{w_1, w_2, ..., w_n\} \quad \forall\, n \in \mathbb{N}$$

and

$$u_{0,n} \to u_0 \ \text{ in } H \quad .$$

Then there exists a unique function $u \in L^2(0,T,V)$ *with* $[u]_t \in L^2(0,T,V^*)$ *that satisfies*

$$\begin{aligned} ([u]_t, v)_H + A(u,v) = 0 \quad \text{for all} \quad v \in V \qquad (2.20)\\ u(.,0) = u_0(.) \quad . \end{aligned}$$

Proposition 2.15. *(Zeidler [40]) The condition that the bilinear mapping* $A : V \times V \to \mathbb{R}$ *satisfies the Gårding inequality is sufficient to replace the coercivity requirements in Theorem 2.14.*

Proof. Suppose that the bilinear mapping A satisfies the Gårding inequality, that is, that there exist constants $\alpha > 0$ and $\lambda \in \mathbb{R}$, such that for every $u \in V$

$$A(u,u) + \lambda \|u\|_H^2 \geq \alpha \|u\|_V^2 \quad .$$

We now set $w = e^{-\lambda t} u$ and notice that

$$\begin{aligned} ([w]_t, v)_H + A(w,v) &= (e^{-\lambda t}[u]_t, v)_H - \lambda(e^{-\lambda t}u, v)_H + A(e^{-\lambda t}u, v) \\ &= e^{-\lambda t}\left(([u]_t, v)_H + A(u,v)\right) - \lambda(w,v)_H \quad ., \end{aligned}$$

and hence, recalling (2.20),

$$([w]_t, v)_H + A(w,v) + \lambda(w,v)_H = 0 \ \text{ for all } v \in V \quad .$$

Setting $\tilde{A}(w,v) = A(w,v) + \lambda(w,v)_H$ then yields the transformed problem:

$$([w]_t, v)_H + \tilde{A}(w,v) = 0 \ \text{ for all } v \in V \quad , \qquad (2.21)$$

with

$$w(0) = u_0 \quad .$$

The mapping $\tilde{A} : V \times V \to \mathbb{R}$ is clearly bilinear and is coercive since

$$\begin{aligned}\tilde{A}(w,w) &= A(w,w) + \lambda(w,w)_H \\ &\geq \alpha\|w\|_V^2 - \lambda\|w\|_H^2 + \lambda\|w\|_H^2 \\ &= \alpha\|w\|_V^2 \quad .\end{aligned}$$

We may therefore apply Theorem 2.14 to the transformed problem (2.21) as required. □

Proposition 2.16. *Condition* 3 *in Theorem 2.14 follows from Condition* 1.

Proof. We begin by recalling that since V is separable, the Gram-Schmidt process (see Kreyszig [26] p.157) may be employed to obtain a total orthonormal sequence in V. That is, a sequence $(w_1, w_2, ...)$ such that

$$V = \overline{\text{span}(w_1, w_2, ...)} \quad . \tag{2.22}$$

We will now demonstrate that this sequence may be used to construct a suitable $(u_{0,n})$ that satisfies Condition 3.

To this end, we notice that since V is dense in in H, for every $u_0 \in H$ there must exists a sequence $(u_0^{(n)})$ in V such that for each $n \in \mathbb{N}$

$$\|u_0^{(n)} - u_0\|_H < \frac{1}{2n} \quad . \tag{2.23}$$

Now, since the sequence $(u_0^{(n)})$ is contained within V, it follows from (2.22) that each element therein may be approximated by elements within $\text{span}(\mathrm{w}_1, \mathrm{w}_2, ...)$. Hence, for each $n \in \mathbb{N}$, there exists $v_n \in \text{span}(w_1, w_2, ...)$ such that

$$\|v_n - u_0^{(n)}\| < \frac{1}{2n} \tag{2.24}$$

We now notice that , since the sequence (v_n) is contained within $\text{span}(w_1, w_2, ...)$, for each $n \in \mathbb{N}$ there exists an $m_n \in \mathbb{N}$ and scalars $a_{1,n}, a_{2,n}, ..., a_{m_n,n}$ such that

$$v_n = a_{1,n}w_1 + a_{2,n}w_2 + ... + a_{m_n,n}w_{m_n} \quad .$$

Making use of this fact, we now construct the sequence $(u_{0,n})$ as follows:

For $n = 1, 2, ...m_1$, we set

$$u_{0,n} = \sum_{i=1}^{n} a_{i,1}w_i \quad ,$$

so that $u_{0,n} \in span\{w_1, w_2, ...w_n\}$ and $u_{0,m_1} = v_1$.
Then, iteratively for each $i = 2, 3, ...$, we set

$$u_{0,m_1+m_2+...m_{i-1}+j} = u_{0,m_1+m_2+...m_{i-1}} \text{ for } j = 1, 2, ...m_i - 1$$

and

$$u_{0,m_1+m_2+...m_{i-1}+m_i} = v_{(m_i)}.$$

Clearly this sequence satisfies the requirement that $u_{0,n} \in span\{w_1, w_2, ...w_n\}$ and hence it only remains to show that it has the limit u_0.

To this end, we begin by noting that the constructed sequence is of the form

$$(u_{0,n}) = (a_{1,1}w_1, a_{1,1}w_1 + a_{2,1}w_2, ..., v_1, ..., v_1, v_2, ...v_2, ..., v_n, ...v_n, ...)$$

and hence, given $n \in \mathbb{N}$, there exists $N \in \mathbb{N}$ such that

$$\|u_{0,n} - u_0\|_H = \|v_n - u_0\|_H \leq \|v_n - u_0^{(n)}\|_H + \|u_0^{(n)} - u_0\|_H \quad ,$$

recalling that V is continuously embedded in H

$$\leq \|v_n - u_0^{(n)}\|_V + \|u_0^{(n)} - u_0\|_H \quad ,$$

recalling 2.23 and 2.24,

$$< \frac{1}{2n} + \frac{1}{2n} = \frac{1}{n} \quad .$$

Thus, as required $\lim_{n\to\infty} \|u_{0,n} - u_0\|_H = 0$ and hence Condition 3 is satisfied.

□

We will now proceed to demonstrate that Theorem 2.14 may be applied to Problem 2.

Lemma 2.17. $(\mathcal{W}_\mu, \mathcal{L}_\mu, \mathcal{W}_\mu^*)$ *is an evolution triple.*

Proof. It follows from Theorem 2.7 that $\mathcal{W}_\mu$ and $\mathcal{L}_\mu$ are separable Hilbert spaces. Furthermore, we notice that the definitions of the spaces $\mathcal{W}_\mu$ and $\mathcal{L}_\mu$ imply that for each $u \in \mathcal{W}_\mu$

$$\|u\|_{\mathcal{L}_\mu} \leq \|u\|_{\mathcal{W}_\mu} \quad ,$$

so that

$$\mathcal{W}_\mu \hookrightarrow \mathcal{L}_\mu \quad .$$

Making use of this fact and Theorems 2.5 and 2.6 in the previous section, it follows that $\mathcal{W}_\mu$ is dense in $\mathcal{L}_\mu$ and hence (see Brezis [11] p.136)

$$\mathcal{W}_\mu \hookrightarrow \mathcal{L}_\mu \hookrightarrow \mathcal{W}_\mu^* \quad .$$

□

Lemma 2.18. *(Cauchy inequality)*
For $\alpha, \beta \in \mathbb{R}$ and any $\epsilon > 0$

$$\alpha\beta \leq \epsilon\alpha^2 + \frac{1}{4\epsilon}\beta^2$$

Proof. Given $\epsilon > 0$, it follows that that

$$0 \leq \left(\sqrt{\epsilon}\alpha - \frac{1}{2\sqrt{\epsilon}}\beta\right)^2$$

$$= \epsilon\alpha^2 - \alpha\beta + \frac{1}{4\epsilon}\beta^2 \quad ,$$

from which the result follows immediately. □

Lemma 2.19. *(Sanfelici [33])*
The bilinear form $\mathcal{A}_\mu$ satisfies the Gårding inequality. That is, there exist constants $\alpha > 0$ and $\lambda \in \mathbb{R}$ such that for each $u \in \mathcal{W}_\mu$

$$\mathcal{A}_\mu(u,u) + \lambda\|u\|^2_{\mathcal{L}_\mu} \geq \alpha\|u\|^2_{\mathcal{W}_\mu} \quad .$$

Proof. We begin by setting

$$\zeta = \min\{\frac{\sigma^2}{2}, r\} \quad ,$$

so that,

$$\begin{aligned} \mathcal{A}_\mu(u,u) &\geq \zeta\|u\|_{\mathcal{W}_\mu} + \int_{\mathbb{R}^+} A(x)\omega^2_\mu(x)xD^{(1)}[u(x)]u(x)\mathrm{d}x \\ &\geq \zeta\|u\|_{\mathcal{W}_\mu} - \|A(x)\|_\infty \int_{\mathbb{R}^+} \omega^2_\mu(x)xD^{(1)}[u(x)]u(x)\mathrm{d}x \quad , \end{aligned}$$

applying Hölders inequality,

$$\begin{aligned} &\geq \zeta\|u\|_{\mathcal{W}_\mu} - \|A\|_\infty|u|_{\mathcal{W}_\mu}\|u\|_{\mathcal{L}_\mu} \\ &\geq \zeta\|u\|_{\mathcal{W}_\mu} - \|A\|_\infty\|u\|_{\mathcal{W}_\mu}\|u\|_{\mathcal{L}_\mu} \quad , \end{aligned}$$

applying the Cauchy inequality,

$$\geq \zeta\|u\|_{\mathcal{W}_\mu} - \epsilon\|A\|_\infty\|u\|^2_{\mathcal{W}_\mu} - \frac{1}{4\epsilon}\|A\|_\infty\|u\|^2_{\mathcal{L}_\mu} \text{ for any } \epsilon > 0 \quad .$$

It therefore follows that

$$\mathcal{A}_\mu(u,u) + \frac{1}{4\epsilon}\|A\|_\infty\|u\|^2_{\mathcal{L}_\mu} \geq (\zeta - \epsilon\|A\|_\infty)\|u\|^2_{\mathcal{W}_\mu} \quad .$$

The result now follows by selecting $0 < \epsilon < \dfrac{\zeta}{\|u\|_\infty}$. □

Recalling Lemma 2.10 from the previous section, we have therefore shown that Problem 2 satisfies the conditions of Theorem 2.14 and hence has a unique solution.

Chapter 3

The Galerkin Approximation Method

In the previous chapter we completed the first step required for the application of finite element type methods - reformulating the valuation PDE as an integral identity over a well chosen Hilbert space (the weak formulation). Our aim is now to obtain easily calculable numerical approximations for the solution of this problem. To this end we will consider the well known Galerkin approximation method. Simplistically, this method will involve deriving approximating versions (termed semi-discrete versions) of Problem 2 over carefully constructed finite dimensional subspaces of $\mathcal{W}_\mu$ (the approximation spaces).

An investigation into the construction of these subspaces will form the bulk of this chapter. In this regard we begin by introducing classical methods for such constructions in the context of a generic variational problem. We will then proceed to follow the work of Sanfelici [33] and Bettess [5] and consider the various alterations that are required in order to derive suitable semi-discrete versions of Problem 2.

3.1 The Finite Element Method

Suppose for the moment that instead of Problem 2, we consider the following generic, time independent problem defined over a bounded domain $\Omega = [a, b] \subset \mathbb{R}$:

Problem 3. *Find* $u \in W^{1,2}(\Omega)$ *such that*

$$\mathcal{A}(u, v) = f(v) \quad \text{for all } v \in W^{1,2}(\Omega) \quad ,$$

where $W^{1,2}(\Omega)$ *is the classical Sobolev space (see Definition A.13 in the appendix) and* $\mathcal{A} : W^{1,2}(\Omega) \times W^{1,2}(\Omega) \to \mathbb{R}$ *is a continuous and coercive bilinear form.*

We now wish to consider the application of Galerkin's method to this problem. Similarly to Problem 2, this requires the selection of suitable finite dimensional subspaces of $W^{1,2}(\Omega)$ over which we may define approximating versions of the above problem. The most common choice within literature for these subspaces are the well-known finite element spaces (a Galerkin's method that makes use of finite element spaces is referred to as the Galerkin finite element method). In this section we will therefore follow the work of Ciarlet [13], as well as that of Guermond and Ern [18] and provide a brief introduction to the construction of finite element spaces and their application within Galerkin's method. To this end, we begin by recalling the following well known definitions:

Definition 3.1. (see Guermond and Ern [18] p.3)
A triangulation of the domain Ω is a finite collection of closed, non-overlapping sub-intervals $\{I_i\}_{i=1}^{i=N}$ (termed elements) that satisfy $\Omega = \cup_{i=1}^{N} I_i$. We write: $\mathcal{T}_h = \{I_i\}_{i=1}^{N}$, where $h_i < \infty$ is the length of the ith element and $h = \max_{i=1:N}(h_i)$.

A triangulation may be uniquely described by a finite set of points $a = a_0 \leq a_1 \leq ... \leq a_N = b$ such that for each $i = 1, 2, ..., N$, $I_i = [a_{i-1}, a_i]$.

Definition 3.2. For each $k \in \mathbb{N}$, we denote by P_k the collection of all polynomials of degree less or equal to k.

Definition 3.3. (see Guermond and Ern [18] p.4)
Given a triangulation $\mathcal{T}_h$ of Ω, we define the associated space of piecewise polynomials of degree k by,

$$P_h^k(\Omega) = \{u \in C(\Omega) |\ \forall i = 1, 2, ..., N\ , u_{|_{I_i}} \in P_k\} \quad .$$

Making use of these definitions, we may now proceed to detail the construction of the finite element spaces for use within Galerkin's method. In keeping with standard notation, we will denote these spaces by $W_{h>0}$ (Functions within these spaces are also written with subscript h. The meaning of this subscript will be given presently). Following Ciarlet [13], we begin by noting that the constructions rely on the following three key aspects.

1. Associated with each finite element space W_h should be a triangulation $\mathcal{T}_h$ of Ω.
2. The finite element spaces W_h should consist of piecewise polynomials. That is, we require that $W_h = P_h^k$ for some $k \in \mathbb{N}$.
3. There should exist a basis of W_h that consists of functions with "small" supports in Ω.

Owing to the simplicity of the domain in question, $\Omega = [a, b]$, it will suffice for us to consider the trivial triangulations $\mathcal{T}_h = \{I_i\}_{i=1}^N$, where $h = \dfrac{b-a}{N}$ and $I_i = [a + (i-1)h, a + ih]$. Property 2 above then dictates that we should choose

$$W_h = P_h^k = \{u \in C(\Omega) |\ \forall i = 1, 2, ..., N\ , u_{|_{I_i}} \in P_k\} \quad ,$$

for some $k \in \mathbb{N}$. For simplicity, we will restrict ourselves to the case in which $k = 1$.

It now remains to show that these spaces have bases that consist of functions with "small" support in Ω. To this end, we follow standard theory and introduce the well known Lagrange polynomials of degree 1.

Definition 3.4. (see Guermond and Ern [18] p.4)
Let $\mathcal{T}_h$ be a triangulation of $\Omega = [a, b]$. We then define, for $i = 1, 2, ...N-1$, the functions

$$\varphi_i = \begin{cases} (1/h)(x - a_{i-1}) & \text{if} \quad x \in I_i \\ (1/h)(a_{i+1} - x) & \text{if} \quad x \in I_{i+1} \\ 0 & \text{otherwise} \end{cases}$$

and

$$\varphi_0 = \begin{cases} (1/h)(h - x) & \text{if} \quad x \in I_1 \\ 0 & \text{otherwise} \end{cases}$$

$$\varphi_N = \begin{cases} (1/h)(x - a_{N-1}) & \text{if} \quad x \in I_N \\ 0 & \text{otherwise} \end{cases} .$$

It should be noted that within literature these functions are often termed "hat functions" due to their characteristic shape (see Figure 3.1 below).

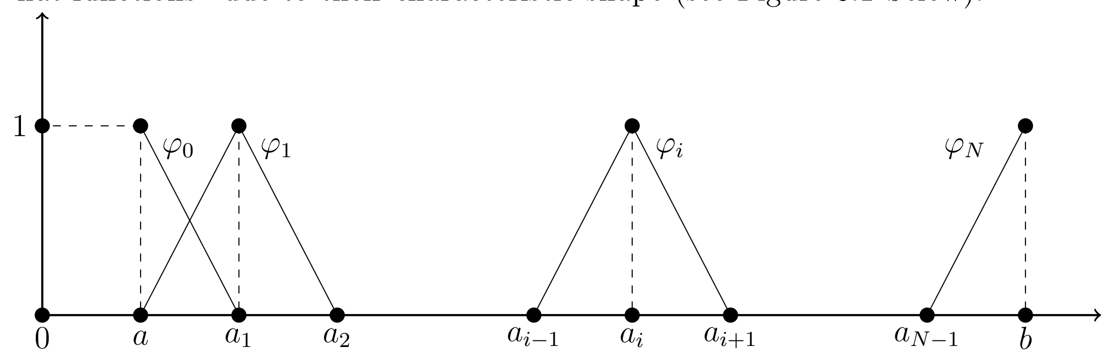

Figure 3.1: Lagrange polynomials of degree 1.

It is easy to show (see Guermond and Ern [18] p.4) that the collection of Lagrangian polynomials, $\{\varphi_0, \varphi_1, ..., \varphi_N\}$, forms a basis for W_h and hence that W_h is finite dimensional. Furthermore, since the support of each of the φ_i's is restricted to at most two elements, this basis satisfies aspect 3 above (the reason for this requirement will become apparent in later chapters).

Lemma 3.5. *The finite element spaces W_h are subspaces of $W^{1,2}(\Omega)$.*

Proof. Let $v_h \in W_h$, it then follows that v_h may be written as a linear combination of the basis functions $\{\varphi_0, \varphi_1, ..., \varphi_N\}$ and thus, there exist scalars $\alpha_0, \alpha_1, ..., \alpha_N$ such that

$$v_h(x) = \Sigma_{i=0}^{N} \alpha_i \varphi_i(x) \quad .$$

We then note that an argument similar to that used in Example A.8 may be employed to show that the function v_h is weakly differentiable. Noting that

$$D^{(1)}[\varphi_i] = \begin{cases} h & \text{on } \mathring{I}_i \\ -h & \text{on } \mathring{I}_{i+1} \\ 0 & \text{otherwise} \end{cases} \quad ,$$

the distributional derivative of v_h is given by

$$D^{(1)}[v_h] = \begin{cases} h[\alpha_i - \alpha_{i-1}] & \text{on } \mathring{I}_i, \text{for } i = 1, 2, ..., N \\ 0 & \text{otherwise} \end{cases} \quad .$$

Noting that Ω is bounded, it then follows immediately from the fact that both v_h and its distributional derivative are bounded that

$$\int_\Omega \varphi_i^2(x) \mathrm{d}x + \int_\Omega \left(D^{(1)}[\varphi_i(x)]\right)^2 \mathrm{d}x < \infty \quad .$$

Thus, as required $v_h \in W^{1,2}(\Omega)$. □

We have thus shown that the finite element spaces W_h are indeed finite dimensional subspaces of $W^{1,2}(\Omega)$ and owing to the way in which these spaces were constructed, they may now be used within Galerkin's method to define approximating versions of Problem 3 (termed discrete problems).

Problem 4. *Find* $u_h \in W_h$ *such that*

$$\mathcal{A}(u_h, v_h) = f(v_h) \text{ for all } v_h \in W_h \quad .$$

For a discussion of the manner in which these problems approximate Problem 3, we direct the reader to the work of Ciarlet [13] as well as that of Guermond and Ern [18]. We will however present a similar discussion in the context of Problem 2 in the next chapter.

3.2 The Infinite Element Method

We now recall that the ultimate aim of this chapter is the application of Galerkin's method to Problem 2 and as such, we must now consider the construction of suitable finite dimensional subspaces of $\mathcal{W}_\mu$ to serve as approximation spaces. In the previous section we illustrated the construction of finite element spaces - the standard choice in this regard. These spaces can however not be constructed within the context of Problem 2, as the unbounded nature of the spacial domain $\mathbb{R}^+$ does not allow for the construction of the required triangulations $\mathcal{T}_h$. In this section, we will therefore follow the work of Sanfelici [33] and Bettess [5], and consider an extension to the finite element theory of the previous section, in the form of an infinite element. This extension will allow us to circumvent the problem of the unbounded domain and construct suitable approximation spaces for use within Galerkin's method.

The construction of these spaces (which we will term the infinite element spaces and denote by $\mathcal{W}_{h>0}$) begins by considering the decomposition of the spacial domain $\mathbb{R}^+$ into the bounded sub-domain $\Omega = [0, x_{max}]$ and an unbounded interval $I_{inf} = [x_{max}, \infty)$ (the infinite element). Following Sanfelici [33], we wish the construction over Ω to mirror that of the finite element spaces in the previous section. Consequently, considering the trivial triangulations $\mathcal{T}_h$ of Ω, we select the associated Lagrangian polynomials of degree 1 as basis functions over this set. $\mathcal{W}_h$ is then defined such that its restriction to Ω consists of piecewise polynomials of degree 1.

It now remains for us to select suitable basis functions over the infinite element. To this end we will follow the work of Bettess [5] and introduce the method of "mapped infinite elements". Under this method, suitable basis functions are obtained by mapping the standard basis functions over the interval $\hat{I} = [0, 1)$ (termed the reference element) onto the infinite element. In this regard, we begin by recalling (see Definition 3.4) that the Lagrangian basis functions of degree 1 over the reference element are given by

$$\hat{\varphi}_1(\hat{x}) = (1 - \hat{x})$$

and

$$\hat{\varphi}_2(\hat{x}) = \hat{x} \quad .$$

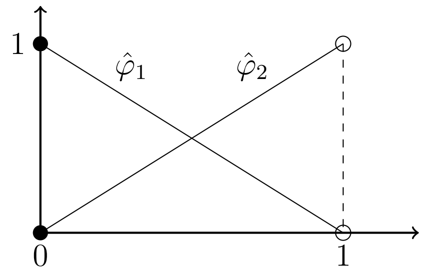

Figure 3.2: Lagrangian basis functions of degree 1 over the reference element.

Adapting the work of Bettess [5], we now introduce the following mapping:

$$F : [0, 1) \rightarrow [x_{max}, \infty) \quad ,$$

where

$$x = F(\hat{x}) = \left(\frac{-\hat{x}}{1-\hat{x}}\right)(x_{max} - h) + \left(1 + \frac{\hat{x}}{1-\hat{x}}\right)(x_{max}) \quad .$$

Clearly we have that

$$F(0) = (0)(x_{max} - h) + (1 - 0)\,(x_{max}) = x_{max}$$

and

$$\lim_{\hat{x}\to 1^-} F(\hat{x}) = \lim_{\hat{x}\to 1^-} \left(\frac{-\hat{x}}{1-\hat{x}}\right)(-h) + x_{max} = \infty \quad .$$

Furthermore,

$$\begin{aligned} x =& \left(\frac{-\hat{x}}{1-\hat{x}}\right)(x_{max} - h) + \left(1 + \frac{\hat{x}}{1-\hat{x}}\right)(x_{max}) \\ =& \left(\frac{-\hat{x}}{1-\hat{x}}\right)(x_{max} - h) + \left(1 + \frac{\hat{x}}{1-\hat{x}}\right)(x_{max} - h) + h\left(1 + \frac{\hat{x}}{1-\hat{x}}\right) \\ =& (x_{max} - h) + \left(h + \frac{h\hat{x}}{1-\hat{x}}\right) = (x_{max} - h) + \left(\frac{h}{1-\hat{x}}\right) \quad . \end{aligned}$$

Thus, F is invertible, with

$$\hat{x} = F^{-1}(x) = 1 - \frac{h}{x - (x_{max} - h)} \quad .$$

Under this mapping, the images of $\hat{\varphi}_1$ and $\hat{\varphi}_2$ on the infinite element are therefore given by

$$\varphi_{inf_1}(x) = \hat{\varphi}_1(F^{-1}(x)) = \frac{h}{x - (x_{max} - h)}$$

and

$$\varphi_{inf_2}(x) = \hat{\varphi}_2(F^{-1}(x)) = 1 - \frac{h}{x - (x_{max} - h)} \quad .$$

Setting

$$\varphi_N^* = \begin{cases} \varphi_N & \text{on } I_N \\ \varphi_{inf_1} & \text{on } I_{inf} \end{cases} ,$$

we may then define the infinite element spaces as

$$\mathcal{W}_h = \text{span}\{\varphi_0, \varphi_1, ..., \varphi_{N-1}, \varphi_N^*, \varphi_{inf_2}\} \quad .$$

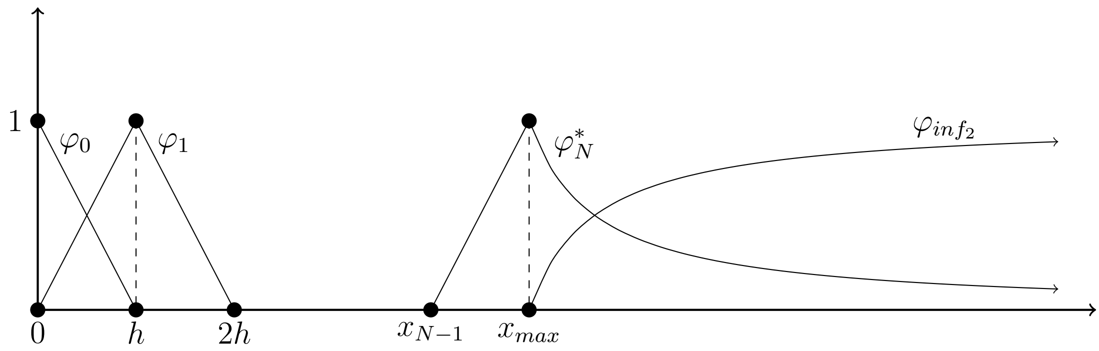

Figure 3.3: Infinite element basis functions.

Lemma 3.6. *The infinite element spaces $\mathcal{W}_h$ are contained within the weighted Sobolev space $\mathcal{W}_\mu$.*

Proof. Let $v_h \in \mathcal{W}_h$, it then follows that v_h may be written as a linear combination of the basis functions $\{\varphi_0, \varphi_1, ..., \varphi_{N-1}, \varphi_N^*, \varphi_{inf_2}\}$ and hence, there exist scalars $\alpha_0, \alpha_1, ...\alpha_{N-1}, \alpha_N, \alpha_{N+1}$ such that

$$v_h(x) = \Sigma_{i=0}^{N-1}(\alpha_i \varphi_i(x)) + \alpha_N \varphi_N^* + \alpha_{N+1}\varphi_{inf_2} \quad .$$

Since v_h is continuous and piecewise classically differentiable, an argument similar to that used in Example A.8 may be employed to show that v_h is weakly differentiable, with

$$D^{(1)}[v_h] = \begin{cases} h[\alpha_i - \alpha_{i-1}] & \text{on } \mathring{I}_i \text{ , for } i = 1, 2, ...N \\ h(x - (x_{max} - h))^{-2}[\alpha_{inf_2} - \alpha_N^*] & \text{on } \mathring{I}_{inf} \\ 0 & \text{otherwise} \end{cases} .$$

It remains for us to show that $\|v_h\|_{\mathcal{W}_\mu} < \infty$. To this end we note that since each of the basis functions is bounded above by 1, we have that

$$\int_{\mathbb{R}+} \omega_\mu^2(x)|v_h(x)|^2 \mathrm{d}x \leq (N+1)^2 \max_{i=0,1,\ldots,N+1} |\alpha_i|^2 \int_{\mathbb{R}+} \omega_\mu^2(x)\mathrm{d}x < \infty \quad .$$

Furthermore,

$$\begin{aligned}\int_{\mathbb{R}+} \omega_\mu^2(x) x^2 |D^{(1)}[v_h(x)]|^2 \mathrm{d}x = & \sum_{i=1}^{N} \int_{I_i} h|\alpha_i - \alpha_{i-1}|^2 \omega_\mu^2(x) x^2 \mathrm{d}x \\ & + \int_{x_{max}}^{\infty} \left| \frac{h(\alpha_{N+1} - \alpha_N)}{[x - (x_{max} - h)]^2} \right|^2 \omega_\mu^2(x) x^2 \mathrm{d}x \quad .\end{aligned}$$

It is clear that the first term is finite and hence, the result follows by noting that

$$\begin{aligned}& \int_{x_{max}}^{\infty} \left| \frac{h(\alpha_{N+1} - \alpha_N)}{[x - (x_{max} - h)]^2} \right|^2 \omega_\mu^2(x) x^2 \mathrm{d}x \\ \leq & \, C \int_{x_{max}}^{\infty} \left| \frac{1}{[x - (x_{max} - h)]} \right|^4 \left(\frac{x_{max}}{x^2} \right)^\mu x^2 \mathrm{d}x \\ \leq & \, C \int_{x_{max}}^{\infty} \left| \frac{1}{[x - (x_{max} - h)]} \right|^4 x^{2-2\mu} \mathrm{d}x \\ \leq & \, C \int_{x_{max}}^{\infty} \left| \frac{1}{[x - (x_{max} - h)]} \right|^4 \mathrm{d}x < \infty \quad .\end{aligned}$$

□

The infinite element spaces $\mathcal{W}_h$ are therefore finite dimensional subspaces of $\mathcal{W}_\mu$. Before we may proceed to make use of these spaces to formulate semi-discrete versions of Problem 2, we note that due to their discontinuity, the initial conditions at each monitoring time t_m do not lie within the spaces $\mathcal{W}_h$. These conditions should therefore also be approximated in order to be included within the infinite element spaces. Following Sanfelici [33], for this purpose, we consider the $\mathcal{L}_\mu$-projection into $\mathcal{W}_h$.

Definition 3.7. The $\mathcal{L}_\mu$-projection of a function $u \in \mathcal{L}_\mu$ into $\mathcal{W}_h$, is the function $u_h \in \mathcal{W}_h$ that satisfies

$$(u - u_h, v_h)_{\mathcal{L}_\mu} = 0 \;\; \forall \; v_h \in \mathcal{W}_h \quad .$$

Bearing this projection in mind, we may then formulate the semi-discrete approximation of Problem 2 as:

Problem 5. *For each $m = 0, 1, ..., M-1$, find $V_h \in C^1([t_m, t_{m+1}], \mathcal{W}_h)$ such that*

$$\begin{aligned} ([V_h]_t, u_h)_{\mathcal{L}_\mu} + \mathcal{A}_\mu(V_h, u_h) = 0 \quad \forall \; u_h \in \mathcal{W}_h \\ V_h(t_m) = V_{t_m,h} \quad , \end{aligned} \tag{3.1}$$

where $V_{t_m,h}$ is the $\mathcal{L}_\mu$-projection of V_{t_m} into $\mathcal{W}_h$.

Recalling that the solution of the above problem may be written as

$$V_h(t) = \sum_{i=0}^{N-1} \alpha_i(t)\varphi_i + \alpha_N(t)\varphi_N^* + \alpha_{N+1}(t)\varphi_{inf_2} \quad ,$$

or, setting $\hat{\varphi}_i := \varphi_i$ for $i = 0, 1, 2, ..., N-1$ and $\hat{\varphi}_N := \varphi_N^*$, $\hat{\varphi}_{N+1} := \varphi_{inf_2}$,

$$V_h(t) = \sum_{i=0}^{N+1} \alpha_i(t)\hat{\varphi}_i \; , \tag{3.2}$$

it follows that the coefficients $a_i(t)$ must satisfy the following system of ordinary differential equations (ODEs):

$$\sum_{i=0}^{N+1}[\alpha_i(t)]_t(\hat{\varphi}_i, \hat{\varphi}_j)_{\mathcal{L}_\mu} + \sum_{i=0}^{N+1} \alpha_i(t)\mathcal{A}_\mu(\hat{\varphi}_i, \hat{\varphi}_j) = 0 \;\; \text{for } j = 0, 1, 2, ..., N+1 \quad . \tag{3.3}$$

The Cauchy-Lipschitz-Picard Theorem (see Brezis [11] p. 184) implies that the system of ODEs 3.3 has a unique solution. Noting that the solution of Problem 5 is uniquely determined by the coefficients $a_i(t)$, this in turn implies that this solution exists and is unique.

Chapter 4

Convergence of the Galerkin Infinite Element Method

In the previous chapter we introduced the Galerkin infinite element method and then proceeded to derive semi-discrete versions of the valuation problem defined over the infinite element spaces $\mathcal{W}_h$. We now wish to show that the solutions of these problems, V_h, may be used to approximate the option value V. To this end, we begin this chapter by considering a weighted analog of a well known interpolation estimate. We will then proceed to characterise the sense in which the spaces $\mathcal{W}_h$ approximate the weighted Sobolev space $\mathcal{W}_\mu$. Finally, we present a number of stability estimates for the solution of the semi-discrete problem and then, making use of these estimates, demonstrate the convergence of V_h to the option value V.

4.1 Interpolation Estimates

As in the previous chapter, we denote by Ω the bounded interval $[0, x_{max}]$ and by $\mathcal{T}_{h>0}$ the trivial triangulations of Ω. We now proceed to introduce an important class of operators, termed the linear interpolation operators, that map continuous functions into the spaces of piecewise polynomials.

Definition 4.1. (See Guermond and Ern [18] p. 5)
The piecewise linear interpolation operator associated with a triangulation $\mathcal{T}_h$ of Ω is the linear operator $\Pi_h : C(\Omega) \to P_h^1(\Omega)$ defined by

$$\Pi_h u = \sum_{i=0}^{N} u(a_i)\varphi_i \quad ,$$

where $\{\varphi_0, \varphi_1, ..., \varphi_N\}$ is a basis for $P_h^1(\Omega)$.

For the remainder of this dissertation we will assume that the basis functions $\varphi_0, \varphi_1, ..., \varphi_N$ used to define the piece-wise linear interpolation operator Π_h are the Lagrangian basis functions of degree 1, as introduced in Chapter 4.

It is well known (see Guermond and Ern [18] or Oden and Reddy [31]) that the linear interpolant $\Pi_h u$ of a function $u \in W^{2,2}(\Omega)$, approximates u and satisfies

$$\|u - \Pi_h u\|_{W^{1,2}(\Omega)} \leq Ch\|u\|_{W^{2,2}(\Omega)} \;\; \forall \; u \in W^{2,2}(\Omega) \quad , \tag{4.1}$$

for some constant C that is independent of h.

This estimate is essential in classical theory of finite element spaces and hence, before we may proceed investigate the manner in which the infinite element spaces $\mathcal{W}_h$ approximate $\mathcal{W}_\mu$, we must obtain an analogous estimate within the weighted Sobolev space $W^{2,2}(\Omega, \omega_0)$.

To this end, we begin by noting that Definition 4.1 implies that in order for the linear interpolant $\Pi_h u$ to be well defined, we require that the function u is essentially continuous on Ω (in the case of (4.1), this follows due to the well known compact embedding of $W^{2,2}(\Omega)$ into $C(\Omega)$). It therefore follows that in order to derive a weighted version of estimate (4.1), we must first demonstrate that the functions within the weighted Sobolev space $W^{2,2}(\Omega, \omega_0)$ are essentially continuous. In this regard, we will first show that the space $W^{2,2}(\Omega, \omega_0)$ is continuously embedded into $W^{2,2}(\Omega, X)$ and then, by making use of Theorem 1.12, that the compact embedding $W^{2,2}(\Omega, \omega_0) \subset\subset C(\Omega)$ holds.

Lemma 4.2. *The weighted Sobolev space $W^{2,2}(\Omega, \omega_0)$ is continuously embedded into $W^{2,2}(\Omega, X)$.*

Proof. Let $u \in W^{2,2}(\Omega, \omega_0)$, then

$$\begin{aligned}
\|u\|^2_{W^{2,2}(\Omega,X)} &= \int_\Omega x^2 u^2(x)\mathrm{d}x \\
&+ \int_\Omega x^2 \left[D^{(1)}u(x)\right]^2 \mathrm{d}x + \int_\Omega x^2 \left[D^{(2)}u(x)\right]^2 \mathrm{d}x \\
&\leq \|x^2\|_{\infty,\Omega} \int_\Omega u^2(x)\mathrm{d}x + \int_\Omega x^2 \left[D^{(1)}u(x)\right]^2 \mathrm{d}x \\
&+ \int_\Omega x^2 \left[D^{(2)}u(x)\right]^2 \mathrm{d}x \\
&\leq C\left[\int_\Omega u^2(x)\mathrm{d}x + \int_\Omega x^2 \left[D^{(1)}u(x)\right]^2 \mathrm{d}x + \int_\Omega x^2 \left[D^{(2)}u(x)\right]^2 \mathrm{d}x\right] \\
&= C\|u\|_{W^{2,2}(\Omega,\omega_0)} \quad ,
\end{aligned}$$

where $C = \max\{\|x^2\|_{\infty,\Omega}, 1\}$. The result now follows due to Definition A.20 in the appendix. □

Theorem 4.3. *The space $W^{2,2}(\Omega, \omega_0)$ is compactly embedded into $C(\Omega)$.*

Proof. Recalling the above lemma, it follows immediately from Definition A.20 in the appendix that the identity operator

$$\mathbb{I}_1 : W^{2,2}(\Omega, \omega_0) \to W^{2,2}(\Omega, X)$$

is bounded and linear.

Similarly, Theorem 1.12 and Definition A.22 imply that the identity operator

$$\mathbb{I}_2 : W^{2,2}(\Omega, X) \to C(\Omega)$$

is compact.

Suppose now that we consider the operator

$$\mathbb{I}_3 = \mathbb{I}_2 \circ \mathbb{I}_1 : W^{2,2}(\Omega, \omega_0) \to C(\Omega) \quad .$$

Recalling that the collection of compact operators between Banach spaces forms an operator ideal within the space of bounded linear operators (see Conway [14] p. 174), we have that since $W^{2,2}(\Omega, \omega_0), W^{2,2}(\Omega, X)$ and $C(\Omega)$ are Banach spaces, $\mathbb{I}_3$ is compact. The result then follows by nothing that a compact operator between Banach spaces is bounded and hence that for each $u \in W^{2,2}(\Omega, \omega_0)$,

$$\|u\|_{\infty,\Omega} \leq \|\mathbb{I}_3\| \|u\|_{W^{2,2}(\Omega,\omega_0)} \quad . \tag{4.2}$$

□

As in the unweighted case, an embedding of this form implies that the piecewise linear interpolant $\Pi_h u$ is well defined for functions $u \in W^{2,2}(\Omega, \omega_0)$. We may therefore proceed to derive an estimate analogous to (4.1) for functions $u \in W^{2,2}(\Omega, \omega_0)$.

Sanfelici [33] suggests that a starting point for this estimate is the work of French [21], Lyashko and Timerbaev [29], as well as that of their references. The author of this dissertation could however not find a reference to this result in the work of French [21]. Lyashko and Timerbaev [29] do however cite a paper due to Timerbaev [36], in which such an estimate is discussed. This paper, entitled "Оценки погрешности n-мерной сплайн-ихтерполяции в весовых нормах" or "Error estimates for n-dimenstional spline interpolation in weighted norms", does however not appear to have been translated into English and presents results in a more general setting than that which is required here. The following work therefore makes use of author's own translation and adaptation of [36].

We begin by noting that local versions of the required estimate are easily obtained over the elements I_i, for $i = 2, 3, ...N$, as the weight functions may be bounded above and below by non-zero constants.

Lemma 4.4. *There exists a constant $C > 0$ such that for each $i = 2, ...N$ and $u \in W^{2,2}(\Omega, \omega_0)$, we have that*

$$|u - \Pi_h u|_{W^{m,2}(I_i,\omega_0)} \leq Ch|u|_{W^{2,2}(I_i,\omega_0)} \text{ for } m = 0, 1 \quad .$$

Proof. We begin by noting that since the elements I_i are mutually disjoint and $0 \in I_1$, we have that for each $i = 2, ...N$; $I_i \cap \{0\} = \emptyset$. We then recall that it follows from Ciarlet [13] (p.121) that there exists a constant $C > 0$ such that for each $u \in W^{2,2}(I_i)$

$$|u - \Pi_h u|_{W^{m,2}(I_i)} \leq Ch^{2-m}|u|_{W^{2,2}(I_i)} \quad \text{for } m = 0, 1 \ .$$

Now, for $u \in W^{2,2}(I_i, \omega_0)$ and $m = 0, 1$, we have that

$$\begin{aligned}
|u - \Pi_h u|_{W^{m,2}(I_i,\omega_0)} &\leq \left[\max_{x \in I_i}(|x|^{2m})\right]^{1/2} |u - \Pi_i u|_{W^{m,2}(I_i)} \\
&= (ih)^m |u - \Pi_h u|_{W^{m,2}(I_i)} \\
&\leq [2h(i-1)]^m |u - \Pi_h u|_{W^{m,2}(I_i)} \\
&\leq C[(i-1)h]^m h^{2-m} |u|_{W^{2,2}(I_i)} \\
&= C[(i-1)h]^m h^{2-m} \left(\int_{I_i} \left(\frac{x}{x}\right)^2 |D^{(2)}[u(x)]|^2 \mathrm{d}x\right)^{1/2} \\
&\leq C(i-1)^m h^2 \left[\min_{x \in I_i}(x^2)\right]^{-1/2} \left(\int_{I_i} x^2 |D^{(2)}[u(x)]|^2 \mathrm{d}x\right)^{1/2} \\
&= C(i-1)^m h^2 \left[\frac{1}{(i-1)h}\right] \left(\int_{I_i} x^2 |D^{(2)}[u(x)]|^2 \mathrm{d}x\right)^{1/2} \\
&= C(i-1)^{m-1} h \left(\int_{I_i} x^2 |D^{(2)}[u(x)]|^2 \mathrm{d}x\right)^{1/2}
\end{aligned}$$

noting that $i - 1 \geq 1$ and $m = 0, 1$;

$$\leq \ Ch|u|_{W^{2,2}(I_i,\omega_0)} \ .$$

$\square$

The method presented in the above lemma is however not applicable in the case of the first element I_1, as the presence of the x^2 term in the weight functions causes them to become degenerate ($=0$) at $x=0$. We are therefore unable to bound the weight functions below by a non-zero constant. In order to obtain the local estimate over the element I_1, we must therefore proceed by following the method presented within the work Timerbaev [36]. To this end, we begin by introducing the following mapping from the reference element $\hat{I}=[0,1]$ onto I_1.

Definition 4.5. Let $F:\hat{I}\to I_1$ be the mapping defined by

$$F(\hat{x}) = h\hat{x} \quad , \quad \hat{x}\in\hat{I} \quad .$$

We then define the mapped functions

$$\hat{u}(\hat{x}) = \ u(F(\hat{x})) \text{ where } u\in W^{2,2}(I_1,\omega_0)$$

and denote by $\hat{\Pi}$ the interpolation operator (see Definition 4.1 above) over $\hat{I}$, with respect to the trivial decomposition $\mathcal{T}_1=\{\hat{I}\}$. Finally, if we give the obvious meaning to the weighted Sobolev spaces $W^{m,p}(\hat{I},\hat{\omega}_0)$ and $W^{m,p}(\hat{I},\hat{X})$ it is clear that these spaces satisfy embeddings analogous to those given in Theorem 1.12, Theorem 1.13, Lemma 4.2 and Theorem 4.3.

We now present a number of Lemmas due to Timerbaev [36] and then derive weighted versions of two well known results.

Lemma 4.6. *For $m=0,1$, each $u\in W^{m,2}(I_1,\omega_0)$ satisfies*

$$h^{-1/2}|u|_{W^{m,2}(I_1,\omega_0)} = \left(\int_{\hat{I}}(\hat{x})^{2m}\,|D^{(m)}\hat{u}(\hat{x})|^2\mathrm{d}\hat{x}\right)^{1/2} = |\hat{u}|_{W^{m,2}(\hat{I},\hat{\omega}_0)} \quad .$$

Proof. We begin by noting that for any $\phi\in C_0^\infty(\hat{I})$ we have that

$$\frac{d^m}{dx^m}\phi(\hat{x}) = \phi^{(m)}(\hat{x})\left[\frac{d^m}{dx^m}\hat{x}\right]^m$$

$$= [1/h]^m \frac{d^m}{d\hat{x}^m}\phi(\hat{x})$$

so that

$$\frac{d^m}{d\hat{x}^m}\phi(\hat{x}) = h^m \frac{d^m}{dx^m}\phi(\hat{x}) \quad .$$

Thus

$$\begin{aligned}
-\int_{\hat{I}} \hat{u}(\hat{x}) \frac{\mathrm{d}^m}{\mathrm{d}\hat{x}^m}\left[\phi(\hat{x})\right] \mathrm{d}\hat{x} =& -\frac{1}{h}\int_{I_1} u(x) h^m \frac{\mathrm{d}^m}{\mathrm{d}x^m}\left[\phi(\hat{x})\right] \mathrm{d}x \\
=& \frac{1}{h}\int_{I_1} D^{(m)}\left[u(x)\right] h^m \phi(\frac{1}{h}x)\mathrm{d}x \\
=& h^m \int_{\hat{I}} D^{(m)}\left[u(F(\hat{x}))\right] \phi(\hat{x})\mathrm{d}\hat{x} \quad .
\end{aligned}$$

The uniqueness of the weak derivative then implies that

$$|D^{(m)}\hat{u}(\hat{x})|^2 = |h|^{2m}|D^{(m)}u(F(\hat{x}))|^2 \quad . \tag{4.3}$$

Multiplying both sides of this equation by $(\hat{x})^{2m}$ and integrating over $\hat{I}$ yields

$$\begin{aligned}
\int_{\hat{I}} (\hat{x})^{2m}\, |D^{(m)}\hat{u}(\hat{x})|^2 \mathrm{d}\hat{x} =& |h|^{2m}\int_{\hat{I}} (\hat{x})^{2m}\, |D^{(m)}u(F(\hat{x}))|^2 \mathrm{d}\hat{x} \\
=& \int_{\hat{I}} (h\hat{x})^{2m}\, |D^{(m)}u(F(\hat{x}))|^2 \mathrm{d}\hat{x} \quad .
\end{aligned}$$

Recalling that each $x \in I_1$ may be written as $x = F(\hat{x}) = h\hat{x}$ for some $\hat{x} \in \hat{I}$, a change of variables yields

$$\begin{aligned}
\int_{\hat{I}} (\hat{x})^{2m}\, |D^{(m)}\hat{u}(\hat{x})|^2 \mathrm{d}\hat{x} =& \int_{\hat{I}} (h\hat{x})^{2m}\, |D^{(m)}u(F(\hat{x}))|^2 \mathrm{d}\hat{x} \\
=& h^{-1}\int_I x^{2m}|D^{(m)}u(x)|^2 \mathrm{d}x
\end{aligned}$$

and hence,

$$\left(\int_{\hat{I}} (\hat{x})^{2m}\, |D^{(m)}\hat{u}(\hat{x})|^2 \mathrm{d}\hat{x}\right)^{1/2} = h^{-1/2}\left(\int_I x^{2m}|D^{(m)}u(x)|^2 \mathrm{d}x\right)^{1/2}$$

□

Lemma 4.7. *Every* $u \in W^{2,2}(I_1, \omega_0)$ *satisfies*

$$h^{1/2}|u|_{W^{2,2}(I_1,\omega_0)} = \left(\int_{\hat{I}} (\hat{x})^2 \, |D^{(2)}\hat{u}(\hat{x})|^2 \mathrm{d}\hat{x}\right)^{1/2} = |\hat{u}|_{W^{2,2}(\hat{I},\hat{\omega}_0)} \quad .$$

Proof. We begin similarly to the previous lemma and for $m = 2$, multiply equation (4.3) by $\hat{x}^2$ and then integrate over $\hat{I}$ to obtain

$$\begin{aligned}\int_{\hat{I}} (\hat{x})^2 \, |D^{(2)}\hat{u}(\hat{x})|^2 \mathrm{d}\hat{x} &= h^2 \int_{\hat{I}} (h\hat{x})^2 \, |D^{(2)}u(F(\hat{x}))|^2 \mathrm{d}\hat{x} \\ &= h \int_{I_1} x^2 |D^{(2)}u(x)|^2 \mathrm{d}x\end{aligned}$$

and hence,

$$\left(\int_{\hat{I}} (\hat{x})^2 \, |D^{(2)}\hat{u}(\hat{x})|^2 \mathrm{d}\hat{x}\right)^{1/2} = h^{1/2} \left(\int_{I_1} x^m |D^{(2)}u(x)|^2 \mathrm{d}x\right)^{1/2} \quad .$$

□

The proofs of the following results are adapted from the analogous unweighted versions as presented by Ciarlet [13].

Lemma 4.8. *For* $m = 0, 1,$ *the interpolation operator*

$$\hat{\Pi} : W^{2,2}(\hat{I}, \hat{\omega}_0) \to W^{m,2}(\hat{I}, \hat{\omega}_0)$$

is bounded and linear.

Proof. We begin by noting that the operator $\hat{\Pi} : W^{2,2}(\hat{I}, \hat{\omega}_0) \to W^{m,2}(\hat{I}, \hat{\omega}_0)$ is well defined due the fact that $P_1^1(\hat{I}) \subset W^{m,2}(\hat{I}, \hat{\omega}_0)$ and that linearity follows immediately from the definition.

It therefore remains for us to demonstrate boundedness, to this end we note that for each $\hat{u} \in W^{2,2}(\hat{I}, \hat{\omega}_0)$ we have that

$$\begin{aligned}\|\hat{\Pi}\hat{u}\|_{W^{m,2}(\hat{I},\hat{\omega}_0)} &= \|\hat{u}(0)\hat{\varphi}_0 + \hat{u}(1)\hat{\varphi}_1\|_{W^{m,2}(\hat{I},\hat{\omega}_0)} \\ &\leq |\hat{u}(0)|\|\hat{\varphi}_0\|_{W^{m,2}(\hat{I},\hat{\omega}_0)} + |\hat{u}(1)|\|\hat{\varphi}_1\|_{W^{m,2}(\hat{I},\hat{\omega}_0)}\end{aligned}$$

$$
\begin{aligned}
&\leq \max\left\{\|\hat{\varphi}_0\|_{W^{m,2}(\hat{I},\hat{\omega}_0)}, \|\hat{\varphi}_1\|_{W^{m,2}(\hat{I},\hat{\omega}_0)}\right\}[\hat{u}(0)+\hat{u}(1)] \\
&\leq C\|\hat{u}\|_{\infty,\hat{I}}
\end{aligned}
$$

We now recall that due to the embedding $W^{2,2}(\hat{I},\hat{\omega}_0) \subset\subset C(\hat{I})$, there exists a constant $C > 0$ such that for every $\hat{u} \in W^{2,2}(\hat{I},\hat{\omega}_0)$

$$
\begin{aligned}
\|\hat{u}\|_{\infty,\hat{I}} &= \sup_{\hat{x}\in\hat{I}}[\hat{u}(\hat{x})] \\
&\leq C\|\hat{u}\|_{W^{2,2}(\hat{I},\hat{\omega}_0)}
\end{aligned}
$$

and hence, as required

$$
\|\hat{\Pi}\hat{u}\|_{W^{m,2}(\hat{I},\hat{\omega}_0)} \leq C\|\hat{u}\|_{W^{2,2}(\hat{I},\hat{\omega}_0)} \quad .
$$

□

Lemma 4.9. *(Weighted Deny-Lions Lemma)*
For every bounded domain $\Omega \subset \mathbb{R}$, there exists a constant $C > 0$ such that for each $u \in W^{2,2}(\Omega,\omega_0)$

$$
\inf_{p\in P_1(\Omega)} \|u-p\|_{W^{2,2}(\Omega,\omega_0)} \leq C|u|_{W^{2,2}(\Omega,\omega_0)} \quad .
$$

Proof. We begin by noting that the space $P_1(\Omega)$ has dimension 2 and hence so does its dual space $P_1(\Omega)^*$ (see Kreyszig [26] p.114). Now, suppose that f_1 and f_2 are basis functions for $P_1(\Omega)^*$. Then, since
$P_1(\Omega) \subset W^{2,2}(\Omega,\omega_0)$, the Hahn-Banach extension theorem (see Kreyszig [26] p.221) implies that f_1 and f_2 may be extended to bounded linear functionals $\tilde{f}_1$ and $\tilde{f}_2$ in the dual of $W^{2,2}(\Omega,\omega_0)$.

We then recall (see Kreyszig [26] p.115) that since $P_1(\Omega)$ is finite dimensional, we have that for $p \in P_1(\Omega)$,

$$
f_1(p) = f_2(p) = 0 \text{ if and only if } p = 0
$$

and hence, for $p \in P_1(\Omega)$,

$$\tilde{f}_1(p) = \tilde{f}_2(p) = 0 \text{ if and only if } p = 0 \quad . \tag{4.4}$$

Now, suppose that there exists a constant $C > 0$ such that for any $u \in W^{2,2}(\Omega, \omega_0)$

$$\|u\|_{W^{2,2}(\Omega,\omega_0)} \leq C\left[|u|_{W^{2,2}(\Omega,\omega_0)} + |\tilde{f}_1(u)| + |\tilde{f}_2(u)|\right] \quad . \tag{4.5}$$

We then define the mapping $T(u) = (f_1(u), f_2(u)) : P_1 \to \mathbb{R}^2$. Since $\{f_1, f_2\}$ is a basis for the finite dimensional space P_1, the mapping T is an isomorphism and hence for each $u \in W^{2,2}(\Omega, \omega_0)$ there exists a $q \in P_1$ such that

$$T(q) = (f_1(q), f_2(q)) = (\tilde{f}_1(q), \tilde{f}_2(q)) = (\tilde{f}_1(u), \tilde{f}_2(u))$$

and hence that

$$\tilde{f}_1(u-q) = \tilde{f}_2(u-q) = 0 \ . \tag{4.6}$$

Recalling that $q \in P_1(\Omega)$ and hence $D^{(2)}[q] = 0$, it follows from (4.5) and (4.6) that

$$\begin{aligned} \inf_{p\in P_1(\Omega)} \|u-p\|_{W^{2,2}(\Omega,\omega_0)} \leq & \ \|u-q\|_{W^{2,2}(\Omega,\omega_0)} \\ \leq & \ C\left[|u-q|_{W^{2,2}(\Omega,\omega_0)} + |\tilde{f}_1(u-q)| + |\tilde{f}_2(u-q)|\right] \\ = & \ C|u|_{W^{2,2}(\Omega,\omega_0)} \quad , \end{aligned}$$

as required. It now remains to show that inequality (4.5) does indeed hold. To this end, suppose that (4.5) is false, it then follows that there exists a sequence (u_n) in $W^{2,2}(\Omega, \omega_0)$ such that

$$\|u_n\|_{W^{2,2}(\Omega,\omega_0)} = 1 \quad \text{for each } n \in \mathbb{N}$$

and

$$\lim_{n\to\infty}\left(|u_n|_{W^{2,2}(\Omega,\omega_0)} + |\tilde{f}_1(u_n)| + |\tilde{f}_2(u_n)|\right) = 0 \quad . \tag{4.7}$$

Since (u_n) is bounded in $W^{2,2}(\Omega,\omega_0)$, it follows from the compact embedding $W^{2,2}(\Omega,\omega_0) \subset\subset C(\bar{\Omega})$ that there exists a subsequence (u_{n_k}) of (u_n) that converges in $C(\bar{\Omega})$. It therefore follows that (u_{n_k}) is Cauchy in $C(\bar{\Omega})$, i.e. given $\epsilon > 0$ there exists $M \in \mathbb{N}$ such that

$$\|u_{n_j} - u_{n_l}\|_{\infty,\bar{\Omega}} < \frac{\epsilon}{3} \quad \forall j,l \geq M \quad .$$

Now, if ℓ denotes the Lebesgue measure and we set
$C = \left[\sqrt{\ell(\Omega)} + \|x\|_{L^1(\Omega)}\right]$, we have that for $j,l \geq M$.

$$\begin{aligned}
\|u_{n_j} - u_{n_l}\|_{W^{1,2}}(\Omega,\omega_0) &\leq |u_{n_j} - u_{n_l}|_{W^{0,2}(\Omega,\omega_0)} + |u_{n_j} - u_{n_l}|_{W^{1,2}(\Omega,\omega_0)} \\
&\leq \|u_{n_j} - u_{n_l}\|_{\infty,\bar{\Omega}} \left[\sqrt{\mu(\Omega)} + \|x\|_{L^1(\Omega)}\right] \\
&= C\|u_{n_j} - u_{n_l}\|_{\infty,\bar{\Omega}} \\
&< \epsilon/3
\end{aligned}$$

Furthermore, it follows from (4.7) that there exists $M^* \in \mathbb{N}$ such that

$$|u_{n_k}|_{W^{2,2}(\Omega,\omega_0)} < \frac{\epsilon}{3} \quad \forall\ k \geq M^* \quad .$$

Now, provided $j,l \geq \max\{M,M^*\}$, we have that

$$\begin{aligned}
\|u_{n_j} - u_{n_l}\|_{W^{2,2}(\Omega,\omega_0)} \leq&\ \|u_{n_j} - u_{n_l}\|_{W^{1,2}(\Omega,\omega_0)} + |u_{n_j} - u_{n_l}|_{W^{2,2}(\Omega,\omega_0)} \\
\leq&\ \|u_{n_j} - u_{n_l}\|_{W^{1,2}(\Omega,\omega_0)} + |u_{n_j}|_{W^{2,2}(\Omega,\omega_0)} + |u_{n_l}|_{W^{2,2}(\Omega,\omega_0)} \\
<&\ \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} \\
=&\ \epsilon \quad .
\end{aligned}$$

It therefore follows that (u_{n_k}) is Cauchy in $W^{2,2}(\Omega,\omega_0)$ and hence, since $W^{2,2}(\Omega,\omega_0)$ is complete, that there exists $u \in W^{2,2}(\Omega,\omega_0)$ such that

$$(u_{n_k}) \to u \quad \text{in} \quad W^{2,2}(\Omega,\omega_0) \quad .$$

We therefore have that

$$\begin{aligned}|u|_{W^{2,2}(\Omega,\omega_0)} &= \lim_{k\to\infty}\left[|u|_{W^{2,2}(\Omega,\omega_0)} - |u_{n_k}|_{W^{2,2}(\Omega,\omega_0)}\right] \\ &\leq \lim_{k\to\infty}|u-u_{n_k}|_{W^{2,2}(\Omega,\omega_0)} \\ &= 0 \ .\end{aligned}$$

It follows that $D^{(2)}[u]=0$ almost everywhere on Ω and hence that $u\in P_1(\Omega)$. Making use of this fact, as well as (4.7) we then have that

$$0=\lim_{k\to\infty}\tilde{f}_1(u_{n_k})=\tilde{f}_1(u)$$

and

$$0=\lim_{k\to\infty}\tilde{f}_2(u_{n_k})=\tilde{f}_2(u)$$

and hence $u=0$ due to (4.4). This however contradicts the fact that $\|u_{n_k}\|_{W^{2,2}(\Omega,\omega_0)}=1$ and hence (4.5) must hold. □

Making use of the above lemmas, we will now derive the local interpolation estimate over the first interval I_1. The proof of this result is due to Timerbaev [36].

Lemma 4.10. *There exists a constant $C>0$ such that for each $u\in W^{2,2}(\Omega,\omega_0)$ we have that*

$$|u-\Pi_h u|_{W^{m,2}(I_1,\omega_0)} \leq Ch|u|_{W^{2,2}(I_1,\omega_0)} \ .$$

Proof. We begin by noting that clearly $W^{2,2}(\hat{I},\hat{\omega}_0)\hookrightarrow W^{m,2}(\hat{I},\hat{\omega}_0)$ and hence the identity operator $\mathbb{I}: W^{2,2}(\hat{I},\hat{\omega}_0) : W^{m,2}(\hat{I},\hat{\omega}_0)$ exists and is bounded. Furthermore, we demonstrated in Lemma 4.8 that the interpolation operator $\hat{\Pi}: W^{2,2}(\hat{I},\hat{\omega}_0)\to W^{m,2}(\hat{I},\hat{\omega}_0)$ is also bounded. It therefore follows that there exists a constant $C>0$ such that

$$\|I-\hat{\Pi}\|\leq C \ . \tag{4.8}$$

We now also notice that for each $\hat{q}\in P_1(\hat{I})$, we have that

$$\hat{\Pi}(\hat{q})=\hat{q} \ ,$$

and hence that

$$(I - \hat{\Pi})(\hat{q}) = \hat{q} - \hat{q} = 0 \quad . \tag{4.9}$$

Making use of equations (4.8) and (4.9), as well as the weighted Deny-Lions Lemma, it then follows that for each $\hat{u} \in W^{2,2}(\hat{I}, \hat{\omega}_0)$

$$\begin{aligned}
|\hat{u} - \hat{\Pi}(\hat{u})|_{W^{m,2}(\hat{I},\hat{\omega}_0)} &= \inf_{\hat{q}\in P_1(\hat{I})} |\hat{u} - \hat{\Pi}(\hat{u}) - (\hat{q} - \hat{\Pi}(\hat{q}))|_{W^{m,2}(\hat{I},\hat{\omega}_0)} \\
&= \inf_{\hat{q}\in P_1(\hat{I})} |(I - \hat{\Pi})(\hat{u} - \hat{q})|_{W^{m,2}(\hat{I},\hat{\omega}_0)} \\
&\leq \|I - \hat{\Pi}\| \inf_{\hat{q}\in P_1(\hat{I})} \|\hat{u} - \hat{q}\|_{W^{m,2}(\hat{I},\hat{\omega}_0)} \\
&\leq C \inf_{\hat{q}\in P_1(\hat{I})} \|\hat{u} - \hat{q}\|_{W^{m,2}(\hat{I},\hat{\omega}_0)} \\
&\leq C \inf_{\hat{q}\in P_1(\hat{I})} \|\hat{u} - \hat{q}\|_{W^{2,2}(\hat{I},\hat{\omega}_0)} \\
&\leq C|\hat{u}|_{W^{2,2}(\hat{I},\hat{\omega}_0)} \quad .
\end{aligned} \tag{4.10}$$

Making use of Lemmas 4.6 and 4.7, as well as (4.10), we then obtain the required result as follows

$$\begin{aligned}
|u - \Pi(u)|_{W^{m,2}(I_1,\omega_0)} &= h^{1/2}|\hat{u} - \hat{\Pi}(\hat{u})|_{W^{m,2}(\hat{I},\hat{\omega}_0)} \\
&\leq Ch^{1/2}|\hat{u}|_{W^{2,2}(\hat{I},\hat{\omega}_0)} \\
&= Ch|u|_{W^{2,2}(I_1,\omega_0)} \quad .
\end{aligned}$$

□

To conclude this section we will now make use of the local estimates derived in Lemma 4.4 and Lemma 4.10 to illustrate the required interpolation estimate over Ω.

Theorem 4.11. *There exists a constant $C > 0$ such that for each $u \in W^{2,2}(\Omega, \omega_0)$, the following interpolation estimate holds*

$$\|u - \Pi_h(u)\|_{W^{1,2}(\Omega,\omega_0)} \leq Ch\|u\|_{W^{2,2}(\Omega,\omega_0)} \quad .$$

Proof. It follows from Lemma 4.4 and Lemma 4.10, that for $i = 1, 2, ...N$ and $m = 0, 1$,

$$|u - \Pi_h u|_{W^{m,2}(I_i,\omega_0)} \leq \ Ch|u|_{W^{2,2}(I_i,\omega_0)}$$

Making use of these estimates, the result is derived as follows

$$\begin{aligned}
\|u - \Pi_h(u)\|_{W^{1,2}(\Omega,\omega_0)} &\leq \ |u - \Pi_h(u)|_{W^{0,2}(\Omega,\omega_0)} + |u - \Pi_h(u)|_{W^{1,2}(\Omega,\omega_0)} \\
&= \ \sum_{i=1}^{N} \left[|u - \Pi_h(u)|_{W^{0,2}(I_i,\omega_0)} + |u - \Pi_h(u)|_{W^{1,2}(I_i,\omega_0)}\right] \\
&\leq \ \sum_{i=1}^{N} \left[Ch|u|_{W^{2,2}(I_i,\omega_0)}\right] \\
&= \ Ch^2|u|^2_{W^{2,2}(\Omega,\omega_0)} \\
&\leq \ Ch\|u\|^2_{W^{2,2}(\Omega,\omega_0)} \quad .
\end{aligned}$$

□

Corollary 4.12. *Given $\phi \in C_0^\infty(\Omega)$, it follows that*

$$\lim_{h\to 0} \|\phi - \Pi_h\phi\|_{W^{2,2}(\Omega,\omega_0)} = 0$$

Proof. This result follows immediately from Theorem 4.11 by noting that the space $C_0^\infty(\Omega)$ is contained within $W^{2,2}(\Omega, \omega_0)$. □

4.2 Error Estimates

Making use of the interpolation estimates derived in the previous section, we may now proceed to investigate the convergence of the solution of the semi-discrete problem V_h to the option value V .

We begin by investigating the manner in which the weighted Sobolev space $\mathcal{W}_\mu$ is approximated by the infinite element spaces $\mathcal{W}_h$. To this end, we present the following result due to Sanfelici [33].

Theorem 4.13. *Let $x_{max} = h^{-c}$ for some $c > 0$, then for each $u \in \mathcal{W}_\mu$, we have that*

$$\lim_{h\to 0^+} \inf_{u_h \in \mathcal{W}_h} \|u - u_h\|_{\mathcal{W}_\mu} = 0 \quad .$$

Proof. Given $u \in \mathcal{W}_\mu$, we begin by setting $\tilde{u} = u\chi_{[x_{max}]}$, where $\chi_{[a]}$ is the cut-off function as defined in Definition 1.16. Lemma 2.3 then implies that for each $\epsilon > 0$, there exists a function $\phi \in C_0^\infty(\Omega)$ such that

$$\|\tilde{u} - \phi\|_{W^{1,2}(\Omega,\omega_\mu)} < \epsilon \quad .$$

Since both $\tilde{u}$ and ϕ vanish within some neighbourhood of x_{max}, they may be extended by 0 outside of Ω to obtain

$$\|\tilde{u} - \phi\|_{\mathcal{W}_\mu} < \epsilon \quad .$$

Furthermore, we notice that since $\phi \in C_0^\infty(\Omega)$, Corollary 4.12 implies that

$$\lim_{h\to 0^+} \|\phi - \Pi_h\phi\|_{W^{1,2}(\Omega,\omega_0)} = \lim_{h\to 0^+} \|\phi - \Pi_h\phi\|_{\mathcal{W}_\mu} = 0 \quad . \tag{4.11}$$

Now, since $\Pi_h\phi \in \mathcal{W}_h$, it follows that

$$\begin{aligned}
\inf_{u_h \in \mathcal{W}_h} \|u - u_h\|_{\mathcal{W}_\mu} \leq\ & \|u - \Pi_h\phi\|_{\mathcal{W}_\mu} \\
\leq\ & \|u - \tilde{u}\|_{\mathcal{W}_\mu} + \|\tilde{u} - \phi\|_{\mathcal{W}_\mu} \\
+\ & \|\phi - \Pi_h\phi\|_{\mathcal{W}_\mu} \\
<\ & \|u - \tilde{u}\|_{\mathcal{W}_\mu} + \epsilon + \|\phi - \Pi_h\phi\|_{\mathcal{W}_\mu} \quad .
\end{aligned} \tag{4.12}$$

Since this holds for every $\epsilon > 0$, the result now follows by taking the limit as $h \to 0^+$ and considering (4.11) and Lemma 2.4 □

For completeness, we now present an estimate due to Sanfelici [33]. We do however note that this estimate is not used in the proof of further results.

Lemma 4.14. *There exists a constant $C > 0$ such that for $u \in \mathcal{W}_\mu$, we have that*

$$\inf_{u_h \in \mathcal{W}_h} \|u - u_h\|_{W^{1,2}(\Omega,\omega_0)} \leq Ch\|u\|_{W^{2,2}(\mathbb{R}^+,\omega_\mu)} \quad ,$$

for every compact domain $\Omega = [0, x_{max}] = [0, h^{-c}] \subset \mathbb{R}^+$.

Proof. Similarly to the previous theorem, we begin by setting $\tilde{u} = u\chi_{[x_{max}]}$ and then note that since $supp(\tilde{u}) \subset \Omega$, $\tilde{u} \in W^{2,2}(\Omega, \omega_\mu)$. The piecewise linear interpolant $\Pi_h\tilde{u}$ is therefore well defined and is a member of $\mathcal{W}_h$ with support in Ω.

Following Sanfelici [33] (this method was originally used by Babuška [4]), we then define the set of weight functions

$$\omega_{\mu,\hat{\mu}} = \{\omega_\mu^2(x)e^{-2\hat{\mu}x}, \omega_\mu^2(x)e^{-2\hat{\mu}x}x^2\}, \text{ for some } \hat{\mu} > 0.$$

Clearly we then have that $e^{-2\hat{\mu}x}$ may be bounded above by 1 on $\mathbb{R}^+$. Furthermore, it follows that since $u(x) - \tilde{u}(x) = 0$ for $0 \leq x \leq \dfrac{1}{3}x_{max}$, we have that

$$\|u - \tilde{u}\|_{W^{1,2}(\mathbb{R}^+,\omega_{\mu,\hat{\mu}})} \leq e^{-\hat{\mu}\frac{1}{3}x_{max}}\|u - \tilde{u}\|_{\mathcal{W}_\mu} \ .$$

Making use of the above facts, as well as Lemma 4.11, we therefore have that

$$\begin{aligned}
\inf_{u_h \in \mathcal{W}_h} \|u - u_h\|_{W^{1,2}(\Omega,\omega_{\mu,\hat{\mu}})} &\leq \|u - \Pi_h\tilde{u}\|_{W^{1,2}(\Omega,\omega_{\mu,\hat{\mu}})} \\
&\leq \|u - \tilde{u}\|_{W^{1,2}(\Omega,\omega_{\mu,\hat{\mu}})} + \|\tilde{u} - \Pi_h\tilde{u}\|_{W^{1,2}(\Omega,\omega_{\mu,\hat{\mu}})} \\
&\leq \|u - \tilde{u}\|_{W^{1,2}(\mathbb{R}^+,\omega_{\mu,\hat{\mu}})} + \|\tilde{u} - \Pi_h\tilde{u}\|_{W^{1,2}(\Omega,\omega_{\mu,\hat{\mu}})} \\
&\leq e^{-\hat{\mu}\frac{1}{3}x_{max}}\|u - \tilde{u}\|_{\mathcal{W}_\mu} + \|\tilde{u} - \Pi_h\tilde{u}\|_{W^{1,2}(\Omega,\omega_\mu)} \\
&\leq e^{-\hat{\mu}\frac{1}{3}x_{max}}\|u - \tilde{u}\|_{\mathcal{W}_\mu} + Ch\|\tilde{u}\|_{W^{2,2}(\Omega,\omega_\mu)} \\
&\leq e^{-\hat{\mu}\frac{1}{3}x_{max}}\left[\|u\|_{\mathcal{W}_\mu} + \|\tilde{u}\|_{\mathcal{W}_\mu}\right] + Ch\|\tilde{u}\|_{W^{2,2}(\mathbb{R}^+,\omega_\mu)} \\
&\leq 2e^{-\hat{\mu}\frac{1}{3}x_{max}}\|u\|_{\mathcal{W}_\mu} + Ch\|\tilde{u}\|_{W^{2,2}(\mathbb{R}^+,\omega_\mu)}
\end{aligned}$$

$$\leq 2e^{-\hat{\mu}\frac{1}{3}x_{max}}\|u\|_{W^{2,2}(\mathbb{R}^+,\omega_\mu)} + Ch\|u\|_{W^{2,2}(\mathbb{R}^+,\omega_\mu)} \quad .$$

The result now follows by noting that there exists a constant $C > 0$ such that $e^{-\hat{\mu}\frac{1}{3}x_{max}} = e^{-\hat{\mu}\frac{1}{3}h^{-1}} \leq Ch.$ □

We now turn our attention to demonstrating the convergence of V_h to the option value V. To this end, we begin by presenting a number of introductory results that will prove useful within the main results of this section.

Lemma 4.15. *Given $u, v \in \mathcal{W}_\mu$, we have that*

$$\lim_{x\to 0^+} \omega_\mu^2(x)xu(x)v(x) = 0$$

Proof. We begin by noting that since $u, v \in \mathcal{W}_\mu$, we have $u, v \in W^{1,2}(\Omega, X)$ and hence, due to the embedding $W^{1,2}(\Omega, X) \subset\subset C(\Omega)$, both u and v are bounded on Ω. We furthermore recall that ω_μ is bounded above by 1 and hence the result follows immediately. □

Lemma 4.16. *Given $u, v \in \mathcal{W}_\mu$, we have that*

$$\lim_{x\to\infty} \omega_\mu^2(x)xu(x)v(x)$$

Proof. We begin by noting that since $\omega_\mu(x)^2u(x)v(x)$ satisfies the conditions of Lemma A.31, it is weakly differentiable over $\mathbb{R}^+$ and hence Lemma A.29 implies that it has a version that is absolutely continuous on $[\frac{1}{n}, n]$, for each $n \in \mathbb{N}$. We therefore have that

$$\int_{1/n}^{n} \omega_\mu(x)^2u(x)v(x)\mathrm{d}x = \left[\omega_\mu(x)^2xu(x)v(x)\right]\Big|_{1/n}^{n} - \int_{1/n}^{n} xD^{(1)}\left[\omega_\mu(x)^2u(x)v(x)\right]\mathrm{d}x \quad .$$

Noting that $\omega_\mu(x)^2u(x)v(x)$ also satisfies the conditions of Lemma A.31, we may apply the product rule for weak derivatives to yield

$$\begin{aligned} \int_{1/n}^{n} \omega_\mu(x)^2 u(x)v(x)\mathrm{d}x = & \left[\omega_\mu(x)^2 x u(x)v(x)\right]\Big|_{1/n}^{n} \\ & - \int_{1/n}^{n} 2\omega_\mu(x) D^{(1)}\left[\omega_\mu(x)\right] x u(x)v(x)\mathrm{d}x \\ & - \int_{1/n}^{n} \omega_\mu^2(x) x D^{(1)}\left[u(x)\right] v(x)\mathrm{d}x \\ & - \int_{1/n}^{n} \omega_\mu^2(x) x u(x) D^{(1)}\left[v(x)\right]\mathrm{d}x \quad . \end{aligned} \tag{4.13}$$

Taking the limit as $n \to \infty$ in (4.13) and applying Lemma 4.15, we have that

$$\begin{aligned} \int_{0}^{\infty} \omega_\mu(x)^2 u(x)v(x)\mathrm{d}x = & \lim_{x\to\infty}\left[\omega_\mu(x)^2 x u(x)v(x)\right] \\ & - \int_{0}^{\infty} 2\omega_\mu(x) D^{(1)}\left[\omega_\mu(x)\right] x u(x)v(x)\mathrm{d}x \\ & - \int_{0}^{\infty} \omega_\mu^2(x) x D^{(1)}\left[u(x)\right] v(x)\mathrm{d}x \\ & - \int_{0}^{\infty} \omega_\mu^2(x) x u(x) D^{(1)}\left[v(x)\right]\mathrm{d}x \quad . \end{aligned} \tag{4.14}$$

Making use of Hölder's inequality, it is easy to show that the the integrals in (4.14) are finite, and hence we must have that $\lim_{x\to\infty}[\omega_\mu(x)^2 x u(x)v(x)]$ exists and is finite.

Suppose now that for some $C > 0$,

$$\lim_{x\to\infty}\left[\omega_\mu(x)^2 x u(x)v(x)\right] = C \quad .$$

It then follows that for every $\epsilon > 0$, there exists $X > 0$ such that

$$\left|\omega_\mu^2(x) x u(x)v(x) - C\right| < \epsilon \quad \forall\, x \geq X \quad ,$$

or,

$$\left[\frac{C-\epsilon}{x_{max}^{2\mu}}\right] x^{2\mu-1} < u(x)v(x) < \left[\frac{C+\epsilon}{x_{max}^{2\mu}}\right] x^{2\mu-1} \quad \forall\, x \geq X \quad .$$

Multiplying this inequality by ω_μ^2 and integrating over $[X, \infty)$, Hölder's inequality implies that

$$\begin{aligned}\int_X^\infty \omega_\mu^2(x)\left[\frac{C-\epsilon}{x_{max}^{2\mu}}\right]x^{2\mu-1}\mathrm{d}x < & \int_X^\infty \omega_\mu^2(x)u(x)v(x)\mathrm{d}x \\ \le & \left(\int_X^\infty \omega_\mu^2(x)u^2(x)\mathrm{d}x\right)^{1/2}\left(\int_X^\infty \omega_\mu^2(x)v^2(x)\mathrm{d}x\right)^{1/2} .\end{aligned} \tag{4.15}$$

Furthermore, we note that

$$\int_X^\infty \omega_\mu^2(x)\left[\frac{C-\epsilon}{x_{max}^{2\mu}}\right]x^{2\mu-1}\mathrm{d}x = \int_X^\infty [C-\epsilon]\, x^{-1}\mathrm{d}x \quad . \tag{4.16}$$

The integral on the right-hand side of 4.16 is clearly not finite and hence neither is the right-hand side of 4.15. This however contradicts the fact that $u, v \in \mathcal{W}_\mu$ and thus, as required,

$$\lim_{x\to\infty}[\omega_\mu(x)^2 x u(x)v(x)] = 0 \quad .$$

□

Lemma 4.17. *There exists a constant* $C > 0$*, such that for all* $u, v \in \mathcal{W}_\mu$*, the bilinear form* $\mathcal{A}_\mu$ *satisfies*

$$|\mathcal{A}_\mu(u,v) - \mathcal{A}_\mu(v,u)| \le C\|u\|_{\mathcal{W}_\mu}\|v\|_{\mathcal{L}_\mu} \quad .$$

Proof. We begin by noting that

$$\begin{aligned}|\mathcal{A}_\mu(u,v) - \mathcal{A}_\mu(v,u)| = & \left|\int_{\mathbb{R}^+} A(x)\omega_\mu^2(x)x\left[D^{(1)}[u(x)]\,v(x) - u(x)D^{(1)}[v(x)]\right]\mathrm{d}x\right| \\ \le & \ \|A\|_\infty\left|\int_{\mathbb{R}^+}\omega_\mu^2(x)xD^{(1)}[u(x)]\,v(x)\mathrm{d}x\right| \\ & + \|A\|_\infty\left|\int_{\mathbb{R}^+}\omega_\mu^2(x)xu(x)D^{(1)}[v(x)]\right| \quad .\end{aligned} \tag{4.17}$$

Now, applying Hölder's inequality we have that

$$\begin{aligned}\left|\int_{\mathbb{R}^+} \omega_\mu^2(x) x D^{(1)}\left[u(x)\right] v(x) \mathrm{d}x\right| &\leq \int_{\mathbb{R}^+} \left|\omega_\mu^2(x) x D^{(1)}\left[u(x)\right] v(x)\right| \mathrm{d}x \\ &\leq \left(\int_{\mathbb{R}^+} \omega_\mu^2(x) x^2 |D^{(1)} u(x)|^2 \mathrm{d}x\right)^{1/2} \\ &\quad \left(\int_{\mathbb{R}^+} \omega_\mu^2(x) |v(x)|^2 \mathrm{d}x\right)^{1/2} \\ &\leq \|u\|_{\mathcal{W}_\mu} \|v\|_{\mathcal{L}_\mu} \quad . \end{aligned} \tag{4.18}$$

Furthermore, since $\omega_\mu^2(x) x u(x)$ satisfies the conditions of Lemma A.31 and $v(x)$ is weakly differentiable, we may make use of Lemma A.29 and apply integration by parts to obtain

$$\begin{aligned}\left|\int_{\mathbb{R}^+} \omega_\mu^2(x) x u(x) D^{(1)}\left[v(x)\right]\right| &\leq \lim_{x\to\infty} \left|\omega_\mu^2(x) x u(x) v(x)\right| + \lim_{x\to 0^+} \left|\omega_\mu^2(x) x u(x) v(x)\right| \\ &\quad + \left|\int_{\mathbb{R}^+} D^{(1)}\left[\omega_\mu^2(x) x u(x)\right] v(x) \mathrm{d}x\right| \quad , \end{aligned}$$

making use of the product rule for weak derivatives, as well as Lemma 4.15 and Lemma 4.16,

$$\begin{aligned} &\leq \left|\int_{\mathbb{R}^+} D^{(1)}\left[\omega_\mu^2(x)\right] u(x) v(x) \mathrm{d}x\right| \\ &\quad + \left|\int_{\mathbb{R}^+} \omega_\mu^2(x) u(x) v(x) \mathrm{d}x\right| \\ &\quad + \left|\int_{\mathbb{R}^+} \omega_\mu^2(x) x D^{(1)}\left[u(x)\right] v(x) \mathrm{d}x\right| \\ &\leq \int_{\mathbb{R}^+} \left|D^{(1)}\left[\omega_\mu^2(x)\right] x u(x) v(x)\right| \mathrm{d}x \\ &\quad + \int_{\mathbb{R}^+} \left|\omega_\mu^2(x) u(x) v(x)\right| \mathrm{d}x \\ &\quad + \int_{\mathbb{R}^+} \left|\omega_\mu^2(x) x D^{(1)}\left[u(x)\right] v(x)\right| \mathrm{d}x \quad . \end{aligned}$$

Recalling Lemma 2.1, it then follows that

$$\int_{\mathbb{R}^+} \left|D^{(1)}\left[\omega_\mu^2(x)\right] xu(x)v(x)\right| \mathrm{d}x = \mu \int_{\mathbb{R}^+} \left|\omega_\mu^2(x)u(x)v(x)\right| \mathrm{d}x \quad .$$

Thus

$$\begin{aligned}\left|\int_{\mathbb{R}^+} \omega_\mu^2(x)xu(x)D^{(1)}\left[v(x)\right]\right| \leq\ & C\int_{\mathbb{R}^+} \left|\omega_\mu^2(x)u(x)v(x)\right| \mathrm{d}x \\ & + \int_{\mathbb{R}^+} \left|\omega_\mu^2(x)u(x)v(x)\right| \mathrm{d}x \\ & + \int_{\mathbb{R}^+} \left|\omega_\mu^2(x)xD^{(1)}\left[u(x)\right]v(x)\right| \mathrm{d}x \quad ,\end{aligned}$$

applying Hölder's inequality,

$$\begin{aligned}& \leq C\|u\|_{\mathcal{W}_\mu}\|v\|_{\mathcal{L}_\mu} + \|u\|_{\mathcal{W}_\mu}\|v\|_{\mathcal{L}_\mu} + \|u\|_{\mathcal{W}_\mu}\|v\|_{\mathcal{L}_\mu} \\ & \leq C\|u\|_{\mathcal{W}_\mu}\|v\|_{\mathcal{L}_\mu} \quad . \end{aligned} \tag{4.19}$$

The result now follows by combining (4.17), (4.18) and (4.19). □

Following the work of Sanfelici [33], we now present a number of stability results for V_h, over the first time interval $[0, t_1]$. We note that similar results hold for each time interval $[t_{m-1}, t_m]$. We begin by recalling that in Chapter 3 we demonstrated that the bilinear form $\mathcal{A}_\mu$ is continuous and satisfies the Gårding inequality and hence that there exist constants $\alpha > 0$, $\lambda \in \mathbb{R}$ and $\gamma > 0$ such that

$$\mathcal{A}_\mu(u, u) + \lambda\|u\|^2_{\mathcal{L}_\mu} \geq \alpha\|u\|^2_{\mathcal{W}_\mu} \tag{4.20}$$

and

$$|\mathcal{A}_\mu(u, v)| \leq \gamma\|u\|_{\mathcal{W}_\mu}\|v\|_{\mathcal{W}_\mu} \quad . \tag{4.21}$$

Proposition 2.15 implies that we may suppose that $\lambda = 0$ in (4.20).

We now recall that we have $V_h(0) = V_{h,0}$, where $V_{h,0}$ is the $\mathcal{L}_\mu$-projection of $V_h(0)$ into $\mathcal{W}_h$ and satisfies

$$(V_h(0), u_h)_{\mathcal{L}_\mu} = (V_{h,0}, u_h)_{\mathcal{L}_\mu} \quad \forall u_h \in \mathcal{W}_\mu \quad . \tag{4.22}$$

Lemma 4.18. *The semi-discrete solution V_h satisfies*

$$2\alpha \int_0^{t_1} \|V_h(t)\|^2_{\mathcal{W}_\mu} \mathrm{d}t \leq \|V_0\|^2_{\mathcal{L}_\mu} \quad ,$$

where α is the coercivity constant given in (4.20).

Proof. We begin by setting $u_h = V_h$ in (3.1) to obtain

$$([V_h]_t, V_h)_{\mathcal{L}_\mu} + \mathcal{A}_\mu(V_h, V_h) = 0 \quad .$$

The product rule (applied to the time derivative) then implies that

$$\frac{d}{dt}(V_h, V_h)_{\mathcal{L}_\mu} + 2\mathcal{A}_\mu(V_h, V_h) = \frac{d}{dt}\|V_h\|^2_{\mathcal{L}_\mu} + 2\mathcal{A}_\mu(V_h, V_h) = 0$$

and hence, since $\mathcal{A}_\mu$ is coercive,

$$\frac{d}{dt}\|V_h\|^2_{\mathcal{L}_\mu} + 2\alpha\|V_h\|^2_{\mathcal{W}_h} \leq 0 \quad .$$

Integrating this inequality over $(0, t_1)$, we obtain

$$\|V_h(t)\|^2_{\mathcal{L}_\mu}\Big|_0^{t_1} + 2\alpha \int_0^{t_1} \|V_h(t)\|^2_{\mathcal{W}_\mu} \mathrm{d}t \leq 0$$

and hence

$$\|V_h(t_1)\|^2_{\mathcal{L}_\mu} + 2\alpha \int_0^{t_1} \|V_h(t)\|^2_{\mathcal{W}_\mu} \mathrm{d}t \leq \|V_h(0)\|^2_{\mathcal{L}_\mu}$$

Clearly then

$$2\alpha \int_0^{t_1} \|V_h(t)\|^2_{\mathcal{W}_\mu} \mathrm{d}t \leq \|V_{h,0}\|^2_{\mathcal{L}_\mu} \tag{4.23}$$

Making use of (4.22), as well as Lemma 2.18 and the Cauchy-Schwarz inequality we then have that

$$\begin{aligned} \|V_{h,0}\|^2_{\mathcal{L}_\mu} &= (V_{h,0}, V_{h,0})_{\mathcal{L}_\mu} \\ &= (V_0, V_{h,0})_{\mathcal{L}_\mu} \\ &\leq \|V_0\|_{\mathcal{L}_\mu}\|V_{0,h}\|_{\mathcal{L}_\mu} \end{aligned}$$

$$\leq \frac{1}{2}\|V_0\|^2_{\mathcal{L}_\mu} + \frac{1}{2}\|V_{0,h}\|_{\mathcal{L}_\mu}$$

and hence

$$\|V_{h,0}\|^2_{\mathcal{L}_\mu} \leq \|V_0\|^2_{\mathcal{L}_\mu} \quad .$$

The result then follows by combining this result with (4.23). □

Lemma 4.19. *If γ and α are the continuity and coercivity constants as given in* (4.20) *and* (4.21)*, then for each $\hat{t} \in (0, t_1]$, the semi-discrete solution V_h satisfies*

$$\frac{1}{2}\hat{t}^2\|[V_h]_t\|^2_{\mathcal{L}_\mu} + \int_0^{\hat{t}} t^2\| \ [V_h(t)]_t \ \|^2_{\mathcal{W}_\mu}\mathrm{d}t \leq c_{\alpha,\gamma}\|V_0\|^2_{\mathcal{L}_\mu} \quad ,$$

where $c_{\alpha,\gamma} = \dfrac{\gamma}{2\alpha}\left(\dfrac{\gamma}{4\alpha} + 1\right)$.

Proof. Differentiating (3.1) with respect to time and then setting $u_h = t^2[V_h]_t$, we have that

$$([V_h]_{tt}, t^2[V_h]_t) + \mathcal{A}_\mu([V_h]_t, t^2[V_h]_t) = 0 \quad ,$$

and hence that

$$\frac{1}{2}\frac{d}{dt}\left[t^2 \ \| \ [V_h]_t \ \|^2_{\mathcal{L}_\mu}\right] - t\| \ [V_h]_t \ \|^2_{\mathcal{L}_\mu} + t^2\mathcal{A}_\mu([V_h]_t, [V_h]_t) = 0 \quad .$$

Making use of (4.20) and then integrating over $[0, \hat{t}]$, we obtain

$$\hat{t}^2\| \ [V_h]_t \ \|^2_{\mathcal{L}_\mu} + 2\alpha\int_0^{\hat{t}} t^2\| \ [V_h(t)]_t \ \|_{\mathcal{W}_\mu}\mathrm{d}t \leq 2\int_0^{\hat{t}} t\| \ [V_h(t)]_t \ \|^2_{\mathcal{L}_\mu}\mathrm{d}t \qquad (4.24)$$

Following Sanfelici [33], we now denote by $\mathcal{A}^s_\mu$ the symmetric part of the bilinear form, i.e.

$$\mathcal{A}^s_\mu(u,v) = \int_{\mathbb{R}^+}\frac{1}{2}\sigma^2\omega^2_\mu(x)x^2D^{(1)}[u(x)]D^{(1)}[v(x)]\mathrm{d}x + \int_{\mathbb{R}^+} r\omega^2_\mu(x)u(x)v(x)\mathrm{d}x \quad .$$

We then note that

$$\begin{aligned}\frac{d}{dt}\left[t\mathcal{A}_\mu^s(V_h,V_h)\right] &= \mathcal{A}_\mu^s(V_h,V_h) + t\mathcal{A}_\mu^s([V_h]_t,V_h) + t\mathcal{A}_\mu^s(V_h,[V_h]_t)\\ &= \mathcal{A}_\mu^s(V_h,V_h) + 2t\mathcal{A}_\mu^s([V_h]_t,V_h)\\ &= \mathcal{A}_\mu^s(V_h,V_h) + 2t\mathcal{A}_\mu([V_h]_t,V_h)\\ &- 2t\int_{\mathbb{R}+} A(x)\omega_\mu^2(x)xD^{(1)}\left[\,[V_h]_t(x)\,\right]V_h(x)\mathrm{d}x \quad .\end{aligned}$$

Setting $u_h = t[V_h]_t$ in (3.1) then yields

$$t\|\,[V_h]_t\,\|_{\mathcal{L}_\mu}^2 + t\mathcal{A}_\mu(V_h,[V_h]_t) = 0 \quad ,$$

and hence,

$$\begin{aligned}&2t\|\,[V_h]_t\,\|_{\mathcal{L}_\mu}^2 + \frac{d}{dt}\left[t\mathcal{A}_\mu^s(V_h,V_h)\right]\\ &\qquad = \mathcal{A}_\mu^s(V_h,V_h) - 2t\int_{\mathbb{R}+} A(x)\omega_\mu^2(x)\ x\ D^{(1)}\left[\,[V_h]_t(x)\,\right]V_h(x)\mathrm{d}x\\ &\qquad \le \gamma\|V_h\|_{\mathcal{L}_\mu}^2 + 2t\|A\|_\infty\int_{\mathbb{R}+}\omega_\mu^2(x)\ x\left|\ D^{(1)}\left[\,[V_h]_t(x)\,\right]\ V_h(x)\ \right|\mathrm{d}x \quad ,\end{aligned}$$

applying Hölder's inequality and then making use of Lemma 2.18,

$$\begin{aligned}&\le \gamma\|V_h\|_{\mathcal{L}_\mu}^2 + \gamma t\|\,[V_h]_t\,\|_{\mathcal{W}_\mu}\|V_h\|_{\mathcal{L}_\mu}\\ &\le \gamma\|V_h\|_{\mathcal{L}_\mu}^2 + \epsilon\gamma t\|\,[V_h]_t\,\|_{\mathcal{W}_\mu}^2 + \frac{1}{4\epsilon}\gamma\|V_h\|_{\mathcal{L}_\mu}^2 \quad ,\end{aligned}$$

for each $\epsilon > 0$.

Integrating both sides of the above inequality over $[0,\hat{t}]$ then yields

$$\begin{aligned}2\int_0^{\hat{t}} t\|\,[V_h(t)]_t\,\|_{\mathcal{L}_\mu}^2\mathrm{d}t + \hat{t}\mathcal{A}_\mu^s(V_h,V_h) &\le \epsilon\gamma\int_0^{\hat{t}} t^2\|\,[V_h(t)]_t\,\|^2\mathrm{d}t\\ &\quad + \left(\frac{\gamma}{4\epsilon}+\gamma\right)\int_0^{\hat{t}}\|V_h(t)\|_{\mathcal{L}_\mu}^2\mathrm{d}t\end{aligned}$$

and thus clearly,

$$2\int_0^{\hat{t}} t\| \ [V_h(t)]_t \ \|^2_{\mathcal{L}_\mu}\mathrm{d}t \le \epsilon\gamma \int_0^{\hat{t}} t^2\| \ [V_h(t)]_t \ \|^2\mathrm{d}t + \left(\frac{\gamma}{4\epsilon}+\gamma\right)\int_0^{\hat{t}} \|V_h(t)\|^2_{\mathcal{L}_\mu}\mathrm{d}t$$

Combining this result with (4.24) then yields

$$\hat{t}^2\| \ [V_h]_t \ \|^2_{\mathcal{L}_\mu} + 2\alpha\int_0^{\hat{t}} t\| \ [V_h(t)]_t \ \|^2_{\mathcal{W}_\mu}\mathrm{d}t \le \epsilon\gamma \int_0^{\hat{t}} t^2\|[V_h(t)]_t\|^2_{\mathcal{W}_\mu}\mathrm{d}t + \left(\frac{\gamma}{4\epsilon}+\gamma\right)\int_0^{\hat{t}} \|V_h(t)\|^2_{\mathcal{L}_\mu}\mathrm{d}t \quad .$$

Since this relation holds for every $\epsilon > 0$, we may select $\epsilon = \dfrac{\alpha}{\gamma}$ to obtain

$$\hat{t}^2\|[V_h]_t\|^2_{\mathcal{L}_\mu} + \alpha\int_0^{\hat{t}} t\|[V_h(t)]_t\|^2_{\mathcal{W}_t}\mathrm{d}t \le \gamma\left(\frac{\gamma}{4\alpha}+1\right)\int_0^{\hat{t}} \|V_h(t)\|^2_{\mathcal{L}_\mu}\mathrm{d}t$$

The result now follows by making use of Lemma 4.18 .

□

Making use of the above stability results, we may now finally turn our attention to demonstrating the convergence of the solution to the semi-discrete problem, V_h, to the option value V. To this end, we define the error function

$$e_h(t) = V(t) - V_h(t)$$

and then, following Sanfelici [33], introduce the *backward auxiliary problem*, as well as its semi-discrete approximation:

Problem 6. *Find* $\upsilon \in \mathcal{W}_\mu$ *such that for any* $\hat{t} \in (0, t_1)$,

$$(u, [\upsilon]_t)_{\mathcal{L}_\mu} - \mathcal{A}_\mu(u,\upsilon) = -(u, e_h)_{\mathcal{L}_\mu} \quad \forall\ u\ \in \mathcal{W}_\mu \tag{4.25}$$

and

$$\upsilon(\hat{t}) = 0 \quad .$$

Problem 7. *Find $\upsilon_h \in \mathcal{W}_h$ such that for any $\hat{t} \in (0, t_1)$,*

$$(u_h, [\upsilon_h]_t)_{\mathcal{L}_\mu} - \mathcal{A}_\mu (u_h, \upsilon_h) = -(u_h, e_h)_{\mathcal{L}_\mu} \quad \forall\ u_h \ \in \mathcal{W}_h \tag{4.26}$$

and

$$\upsilon_h(\hat{t}) = 0 \quad .$$

Making use of the above problems, we now demonstrate a number of estimates that will prove useful later in the chapter.

Lemma 4.20. *For $\hat{t} \in (0, t_1]$, the solution to Problem 6 satisfies*

$$\int_0^{\hat{t}} \|\upsilon(t)\|^2_{\mathcal{W}_\mu} \mathrm{d}t \le \int_0^{\hat{t}} \frac{2}{\alpha^2} \|e_h(t)\|^2_{\mathcal{L}_\mu} \mathrm{d}t$$

Proof. Setting $u = \upsilon$ in equation 4.25, we have that

$$(\upsilon, [\upsilon]_t)_{\mathcal{L}_\mu} - A_\mu(\upsilon, \upsilon) = \ -(\upsilon, e_h)_{\mathcal{L}_\mu} \quad ,$$

or,

$$-\frac{1}{2}\frac{d}{dt}\|\upsilon\|^2_{\mathcal{L}_\mu} + A_\mu(\upsilon, \upsilon) = \ (\upsilon, e_h)_{\mathcal{L}_\mu} \quad .$$

Making use of the Cauchy-Schwarz inequality, 4.20 and Lemma 2.18, it then follows that for each $\epsilon > 0$, we have

$$-\frac{1}{2}\frac{d}{dt}\|\upsilon\|^2_{\mathcal{L}_\mu} + \alpha\|\upsilon\|^2_{\mathcal{W}_\mu} \le \ \epsilon\|\upsilon\|^2_{\mathcal{W}_\mu} + \frac{1}{4\epsilon}\|e_h\|^2_{\mathcal{L}_\mu} \quad .$$

Setting $\epsilon = \dfrac{1}{2}\alpha$ and integrating over $[0, \hat{t}]$, this relation becomes

$$\frac{1}{2}|\upsilon(0)|^2 + \frac{1}{2}\alpha \int_0^{\hat{t}} \|\upsilon(t)\|^2_{\mathcal{W}_\mu} \le \ \int_0^{\hat{t}} \frac{1}{2\alpha}\|e_h(t)\|^2_{\mathcal{L}_\mu} \mathrm{d}t \quad ,$$

and hence, as required,

$$\int_0^{\hat{t}} \|\upsilon(t)\|^2_{\mathcal{W}_\mu} \mathrm{d}t \le \ \int_0^{\hat{t}} \frac{2}{\alpha^2}\|e_h(t)\|^2_{\mathcal{L}_\mu} \mathrm{d}t \quad .$$

□

Lemma 4.21. *There exists a constant $C > 0$, depending on γ and μ, such that for $\hat{t} \in (0, t_1]$,*

$$\int_0^{\hat{t}} \|\upsilon'(t)\|^2_{\mathcal{L}_\mu} \mathrm{d}t \leq C \int_0^{\hat{t}} \|\upsilon(t)\|^2_{\mathcal{W}_\mu}(d)t + \int_0^{\hat{t}} \|e_h(t)\|^2_{\mathcal{L}_\mu} \mathrm{d}t \quad .$$

Proof. We begin by noting that

$$\frac{d}{dt}\mathcal{A}_\mu(\upsilon, \upsilon) = \; 2\mathcal{A}_\mu([\upsilon]_t, \upsilon) + (\mathcal{A}_\mu(\upsilon, [\upsilon]_t) - \mathcal{A}_\mu([\upsilon]_t, \upsilon)) \quad . \tag{4.27}$$

We then proceed by setting $u = [\upsilon]_t$ in equation (4.25), to obtain

$$\|[\upsilon]_t\|^2_{\mathcal{L}_\mu} - \mathcal{A}_\mu([\upsilon]_t, \upsilon) = -([\upsilon]_t, e_h)_{\mathcal{L}_\mu} \quad .$$

Making use of (4.27), we have that

$$\| \; [\upsilon]_t \; \|^2_{\mathcal{L}_\mu} - \frac{1}{2}\frac{d}{dt}\mathcal{A}_\mu(\upsilon, \upsilon) = \; \frac{1}{2}\left[\mathcal{A}_\mu([\upsilon]_t, \upsilon) - \mathcal{A}_\mu(\upsilon, [\upsilon]_t)\right] - ([\upsilon]_t, e_h)_{\mathcal{L}_\mu} \quad ,$$

applying the Cauchy-Schwarz inequality, as well as Lemma 4.17,

$$\leq C\|w\|_{\mathcal{W}_\mu}\|[\upsilon]_t\|_{\mathcal{L}_\mu} + \|[\upsilon]_t\|_{\mathcal{L}_\mu}\|e_h\|_{\mathcal{L}_\mu} \quad ,$$

where, $C > 0$ depends only on γ and μ.

Making use of Lemma 2.18, we then have that for each $\epsilon > 0$

$$\begin{aligned}\|[\upsilon]_t\|^2_{\mathcal{L}_\mu} - \frac{1}{2}\frac{d}{dt}\mathcal{A}_\mu(\upsilon, \upsilon) \leq \; & C\left[\epsilon\|[\upsilon]_t\|^2_{\mathcal{L}_\mu} + \frac{1}{4\epsilon}\|\upsilon\|^2_{\mathcal{W}_\mu}\right] \\ & + \; \epsilon\|[\upsilon]_t\|^2_{\mathcal{L}_\mu} + \frac{1}{4\epsilon}\|e_h\|^2_{\mathcal{L}_\mu} \quad ,\end{aligned}$$

so that,

$$[1 - (C+1)\epsilon]\|[\upsilon]_t\|^2_{\mathcal{L}_\mu} - \frac{1}{2}\frac{d}{dt}\mathcal{A}_\mu(\upsilon, \upsilon) \leq \; C\frac{1}{4\epsilon}\|\upsilon\|^2_{\mathcal{W}_\mu} + \frac{1}{4\epsilon}\|e_h\|^2_{\mathcal{L}_\mu} \quad .$$

Setting $\epsilon = \dfrac{1}{2(C+1)} > 0$, it follows that

$$|[\upsilon]_t|^2_{\mathcal{L}_\mu} - \frac{d}{dt}\mathcal{A}_\mu(\upsilon,\upsilon) \leq \frac{1}{2}(C+1)(C)\|\upsilon\|_{\mathcal{W}^2_\mu} + \frac{1}{2}(C+1)\|e_h\|_{\mathcal{L}_\mu} \quad .$$

The result now follows by integrating over $[0,\hat{t}]$ and recalling that $\upsilon(\hat{t}) = 0$. □

Lemma 4.22. *Consider the error function $\delta_h = \upsilon - \upsilon_h$. There then exists a constant $C > 0$, that depends only on α, γ and μ, such that for $\hat{t} \in (0, t_1]$,*

$$\int_0^{\hat{t}} \|[\delta_h]_t(t)\|^2_{\mathcal{L}_\mu} + \gamma\|\delta_h(t)\|^2_{\mathcal{W}_\mu}\mathrm{d}t \leq C\int_0^{\hat{t}} \|e_h(t)\|^2_{\mathcal{L}_\mu}\mathrm{d}t \quad .$$

Proof. We begin by noting that since $\delta_h = \upsilon - \upsilon_h$, triangle inequality implies that

$$\begin{aligned}
&\int_0^{\hat{t}} \|\delta_h'(t)\|^2_{\mathcal{L}_\mu} + \gamma\|\delta_h(t)\|^2_{\mathcal{W}_\mu}\mathrm{d}t \\
&\qquad \leq \int_0^{\hat{t}} \left[\|\upsilon'(t)\|_{\mathcal{L}_\mu} + \|\upsilon_h'(t)\|_{\mathcal{L}_\mu}\right]^2 \mathrm{d}t + \int_0^{\hat{t}} \gamma\left[\|\upsilon(t)\|_{\mathcal{W}_\mu} + \|\upsilon_h(t)\|_{\mathcal{W}_\mu}\right]^2 \mathrm{d}t \\
&\qquad = \int_0^{\hat{t}} \|\upsilon'(t)\|^2_{\mathcal{L}_\mu} + 2\|\upsilon'(t)\|_{\mathcal{L}_\mu}\|\upsilon_h'(t)\|_{\mathcal{L}_\mu} + \|\upsilon_h'(t)\|^2_{\mathcal{L}_\mu}\mathrm{d}t \\
&\qquad + \ \gamma\int_0^{\hat{t}} \|\upsilon(t)\|^2_{\mathcal{W}_\mu} + 2\|\upsilon(t)\|_{\mathcal{W}_\mu}\|\upsilon_h(t)\|_{\mathcal{W}_\mu} + \|\upsilon_h(t)\|^2_{\mathcal{W}_\mu}\mathrm{d}t \quad ,
\end{aligned}$$

making use of Lemma 2.18, with $\epsilon = \dfrac{1}{2}$,

$$\leq 2\int_0^{\hat{t}} \|\upsilon'(t)\|^2_{\mathcal{L}_\mu} + \|\upsilon_h'(t)\|^2_{\mathcal{L}_\mu}\mathrm{d}t + 2\gamma\int_0^{\hat{t}} \|\upsilon(t)\|^2_{\mathcal{W}_\mu} + \|\upsilon_h(t)\|^2_{\mathcal{W}_\mu}\mathrm{d}t \quad ,$$

the application of Lemma 4.20 and Lemma 4.21, as well as their semi-discrete analogues, then implies

$$\leq C\int_0^{\hat{t}} \|\upsilon(t)\|^2_{\mathcal{W}_\mu} + \|e_h(t)\|^2_{\mathcal{L}_\mu}\mathrm{d}t$$

$$+\int_0^{\hat{t}} \|\upsilon_h(t)\|^2_{\mathcal{W}_\mu} + \|e_h(t)\|^2_{\mathcal{L}_\mu}\mathrm{d}t + \frac{4\gamma}{\alpha^2}\int_0^{\hat{t}} \|e_h(t)\|^2_{\mathcal{L}_\mu} + \|e_h(t)\|^2_{\mathcal{L}_\mu}\mathrm{d}t$$
$$= C\int_0^{\hat{t}} \|e_h(t)\|^2_{\mathcal{L}_\mu}\mathrm{d}t + C\int_0^{\hat{h}} \|\upsilon(t)\|^2_{\mathcal{W}_\mu}\mathrm{d}t + \int_0^{\hat{t}} \|\upsilon_h(t)\|^2_{\mathcal{W}_\mu}\mathrm{d}t$$

applying Lemma 4.20 as well as its semi-discrete analogue

$$\leq C\int_0^{\hat{t}} \|e_h(t)\|^2_{\mathcal{L}_\mu}\mathrm{d}t + \int_0^{\hat{t}} \frac{2C}{\alpha^2}\|e_h(t)\|^2_{\mathcal{L}_\mu}\mathrm{d}t \quad ,$$

from which the result follows. □

Definition 4.23. Let $\upsilon \in \mathcal{W}_\mu$, then the Ritz projection $R_h\upsilon$ of υ into $\mathcal{W}_h$ is defined such that

$$\mathcal{A}_\mu\left(R_h\upsilon, u_h\right) = \mathcal{A}_\mu(\upsilon, u_h) \quad \forall\ u_h \ \in\ \mathcal{W}_h \tag{4.28}$$

We note that the existence of the Ritz projection follows from the Lax-Milgram Lemma (see Brezis [11] page 140) by noting that for each $u \in \mathcal{W}_\mu$ the bilinear form $\mathcal{A}_\mu(u, .)$ maps the space $\mathcal{W}_h$ into $\mathbb{R}$ and hence, $\mathcal{A}_\mu(u, .)$ is a member of the dual space $\mathcal{W}_h^*$.

Lemma 4.24. *Given* $\upsilon \in \mathcal{W}_\mu$

$$\|\upsilon - R_h\upsilon\|_{\mathcal{W}_\mu} \leq \frac{\gamma}{\alpha}\inf_{u_h \in \mathcal{W}_h} \|\upsilon - u_h\|_{\mathcal{W}_\mu} \quad .$$

Proof. This result follows trivially in the case that $\|\upsilon - R_h\upsilon\|_{\mathcal{W}_\mu} = 0$ and hence we may be assume that $\|\upsilon - R_h\upsilon\|_{\mathcal{W}_\mu} \neq 0$. Making use of (4.20) and (4.28), it then follows that

$$\begin{aligned}\alpha\|\upsilon - R_h\upsilon\|^2_{\mathcal{W}_\mu} &\leq \mathcal{A}_\mu(\upsilon - R_h\upsilon, \upsilon - R_h\upsilon)\\ &= \mathcal{A}_\mu(\upsilon - R_h\upsilon, \upsilon) - \mathcal{A}_\mu(\upsilon - R_h\upsilon, R_h\upsilon)\\ &= \mathcal{A}_\mu(\upsilon - R_h\upsilon, \upsilon)\end{aligned}$$

Furthermore, making use of (4.28), we have that for each $u_h \in \mathcal{W}_h$,

$$\alpha\|\upsilon - R_h\upsilon\|^2_{\mathcal{W}_\mu} \leq \mathcal{A}_\mu(\upsilon - R_h\upsilon, \upsilon) - \mathcal{A}_\mu(\upsilon - R_h\upsilon, u_h)$$

$$= \mathcal{A}_\mu(\upsilon - R_h\upsilon, \upsilon - u_h) \quad .$$

Hence, by (4.21)

$$\alpha\|\upsilon - R_h\upsilon\|^2_{\mathcal{W}_\mu} \leq \gamma\|\upsilon - R_h\upsilon\|_{\mathcal{W}_\mu}\|\upsilon - u_h\|_{\mathcal{W}_\mu} \quad \forall\ u_h \in \mathcal{W}_h$$

and hence

$$\|\upsilon - R_h\upsilon\|_{\mathcal{W}_\mu} \leq \frac{\gamma}{\alpha}\|\upsilon - u_h\|_{\mathcal{W}_\mu} \quad \forall\ u_h \in \mathcal{W}_h$$

The result now follows by taking the infimum over all $u_h \in \mathcal{W}_h$. □

Lemma 4.25. *The error function $e_h(t) = V(t) - V_h(t)$ satisfies*

$$\int_0^{\hat{t}} \|e_h(t)\|^2_{\mathcal{L}_\mu}\mathrm{d}t \leq C \inf_{w_h\in\mathcal{W}_h} \int_0^{\hat{t}} \|V(t) - w_h\|^2_{\mathcal{W}_\mu}\mathrm{d}t$$

for $\hat{t} \in (0, t_1]$ and some constant $C > 0$ that depends only on α, γ and μ.

Proof. We begin by noting that since $\mathcal{W}_h \subset \mathcal{W}_\mu$, we may subtract equation (4.26) from equation (4.25) and (3.1) from (2.19) to obtain

$$([e_h]_t, u_h)_{\mathcal{L}_\mu} + \mathcal{A}_\mu(e_h, u_h) = 0 \quad \forall\ u_h \in \mathcal{W}_h \tag{4.29}$$

and

$$(u_h, [\delta_h]_t)_{\mathcal{L}_\mu} - \mathcal{A}_\mu(u_h, \delta_h) = 0 \quad \forall\ u_h \in \mathcal{W}_h \quad . \tag{4.30}$$

We then proceed by noting that if R_h denotes the Ritz projection operator, the repeated application of (4.29) and (4.30) yields,

$$\begin{aligned}
-\frac{d}{dt}&(e_h, \upsilon_h)_{\mathcal{L}_\mu} - (V - R_hV, [\delta_h]_t)_{\mathcal{L}_\mu} + \mathcal{A}_\mu(V - R_hV, \delta_h) \\
&= -([e_h]_t, \upsilon_h)_{\mathcal{L}_\mu} - (e_h, [\upsilon_h]_t)_{\mathcal{L}_\mu} - (V, [\delta_h]_t)_{\mathcal{L}_\mu} + \mathcal{A}_\mu(V, \delta_h) \\
&\quad + \left[(R_hV, [\delta_h]_t)_{\mathcal{L}_\mu} - \mathcal{A}_\mu(R_hV, \delta_h)\right] \quad , \\
&= -([e_h]_t, \upsilon_h)_{\mathcal{L}_\mu} - (e_h, [\upsilon_h]_t)_{\mathcal{L}_\mu} - (V, [\delta_h]_t)_{\mathcal{L}_\mu} + \mathcal{A}_\mu(V, \delta_h) \\
&= -([e_h]_t, \upsilon_h)_{\mathcal{L}_\mu} - (e_h, [\upsilon_h]_t)_{\mathcal{L}_\mu} - (V, [\delta_h]_t)_{\mathcal{L}_\mu} + \mathcal{A}_\mu(V_h, \delta_h) \\
&\quad + \mathcal{A}_\mu(e_h, \delta_h)
\end{aligned}$$

$$
\begin{aligned}
&= -([e_h]_t, \upsilon_h)_{\mathcal{L}_\mu} - (e_h, [\upsilon_h]_t)_{\mathcal{L}_\mu} - (V, [\delta_h]_t)_{\mathcal{L}_\mu} + (V_h, [\delta_h]_t)_{\mathcal{L}_\mu} \\
&\quad + \mathcal{A}_\mu(e_h, \delta_h) \\
&= -([e_h]_t, \upsilon_h)_{\mathcal{L}_\mu} - (e_h, [\upsilon_h]_t)_{\mathcal{L}_\mu} - (e_h, [\delta_h]_t)_{\mathcal{L}_\mu} + \mathcal{A}_\mu(e_h, \delta_h) \\
&= -([e_h]_t, \upsilon_h)_{\mathcal{L}_\mu} - (e_h, [\upsilon]_t)_{\mathcal{L}_\mu} + \mathcal{A}_\mu(e_h, \delta_h) \\
&= -\mathcal{A}_\mu(e_h, \upsilon_h) - (e_h, [\upsilon]_t)_{\mathcal{L}_\mu} + \mathcal{A}_\mu(e_h, \delta_h) \\
&= \mathcal{A}_\mu(e_h, \upsilon) - (e_h, \upsilon)_{\mathcal{L}_\mu} \\
&= \|e_h\|^2_{\mathcal{L}_\mu} \quad .
\end{aligned}
$$

Recalling then, that $(V(0) - V_{h,0}, u_h)_{\mathcal{L}_\mu} = 0$ for every $u_h \in \mathcal{W}_h$, as well as the fact that $\upsilon(\hat{t}) = 0$, integration over $[0, \hat{t}]$ yields

$$
\begin{aligned}
\int_0^{\hat{t}} \|e_h(t)\|^2_{\mathcal{L}_\mu} \mathrm{d}t &= \int_0^{\hat{t}} -(V(t) - R_hV(t), \delta_h'(t))_{\mathcal{L}_\mu} + \mathcal{A}_\mu(V(t) - R_hV(t), \delta_h(t)) \ \mathrm{d}t \\
&\le \int_0^{\hat{t}} \left|(V(t) - R_hV(t), \delta_h'(t))_{\mathcal{L}_\mu}\right| + |\mathcal{A}_\mu(V(t) - R_hV(t), \delta_h)| \ \mathrm{d}t
\end{aligned}
$$

making use of the Cauchy-Schwarz inequality and (4.21),

$$
\begin{aligned}
&\le \int_0^{\hat{t}} \|V(t) - R_hV(t)\|_{\mathcal{L}_\mu} \|\delta_h'(t)\|_{\mathcal{L}_\mu} \\
&\quad + \gamma \|V(t) - R_hV(t)\|_{\mathcal{W}_\mu} \|\delta_h(t)\|_{\mathcal{W}_\mu} \mathrm{d}t
\end{aligned}
$$

applying Lemma 2.18, for any $\epsilon > 0$,

$$
\begin{aligned}
&\le \int_0^{\hat{t}} \epsilon \left[\|\delta_h'(t)\|^2_{\mathcal{L}_\mu} + \gamma \|\delta_h(t)\|^2_{\mathcal{W}_\mu}\right] \mathrm{d}t \\
&\quad + \int_0^{\hat{t}} \frac{1}{4\epsilon} \left[\|V(t) - R_hV(t)\|^2_{\mathcal{L}_\mu} + \gamma \|V(t) - R_hV(t)\|^2_{\mathcal{W}_\mu}\right] \mathrm{d}t \quad .
\end{aligned}
$$

Recalling Lemma 4.22 and then setting $\epsilon = \dfrac{1}{2C}$, we have that

$$
\int_0^{\hat{t}} \|e_h(t)\|^2_{\mathcal{L}_\mu} \mathrm{d}t \le \int_0^{\hat{t}} \frac{1}{2C} \left[\|\delta_h'(t)\|^2_{\mathcal{W}_\mu} + \gamma \|\delta_h(t)\|^2_{\mathcal{W}_\mu}\right] \mathrm{d}t
$$

$$
\begin{aligned}
&+ \int_0^{\hat{t}} \frac{C}{2} \left[\|V(t) - R_h V(t)\|^2_{\mathcal{L}_\mu} + \gamma \|V(t) - R_h V(t)\|^2_{\mathcal{W}_\mu} \right] \mathrm{d}t \\
\leq\ & \frac{1}{2} \int_0^{\hat{t}} \|e_h(t)\|^2_{\mathcal{L}_\mu} \mathrm{d}t \\
&+ \frac{C}{2} \int_0^{\hat{t}} \left[\|V(t) - R_h V(t)\|^2_{\mathcal{L}_\mu} + \gamma \|V(t) - R_h V(t)\|^2_{\mathcal{W}_\mu} \right] \mathrm{d}t
\end{aligned}
$$

and hence,

$$
\int_0^{\hat{t}} \|e_h(t)\|^2_{\mathcal{L}_\mu} \mathrm{d}t \leq C \int_0^{\hat{t}} \left[\|V(t) - R_h V(t)\|^2_{\mathcal{W}_\mu} + \gamma \|V(t) - R_h V(t)\|^2_{\mathcal{W}_\mu} \right] \mathrm{d}t
$$

The result now follows due to Lemma 4.24. □

It has been noted that one would expect to see the initial projection error within the error estimate presented in the above lemma. This matter is under investigation by the author.

Theorem 4.26. *The solution to the semi-discrete problem, V_h, converges to the option value V in $\mathcal{W}_\mu$ and the following error estimate holds,*

$$
\begin{aligned}
\|t_1 e_h\|^2_{\mathcal{L}_\mu} + 2\alpha \int_0^{t_1} t \|e_h(t)\|^2_{\mathcal{W}_\mu} \mathrm{d}t \leq C \inf_{v_h \in \mathcal{W}_h} \|V(t) - v_h\|_{L^2(0,t_1,\mathcal{W}_\mu)} \\
+ C \|V_0\|_{\mathcal{L}_\mu} \inf_{w_h \in \mathcal{W}_h} \|V - w_h\|_{L^2(0,t_1,\mathcal{W}_\mu)}
\end{aligned}
$$

Proof. Following Sanfelici [33], we begin by setting
$u = e_h$ in 2.19 and, for arbitrary $w_h \in \mathcal{W}_h$, $u_h = V_h - w_h$ in 3.1 to obtain

$$
([V]_t, e_h)_{\mathcal{L}_\mu} + \mathcal{A}_\mu(V, e_h) = 0
$$

and

$$
([V_h]_t, V_h - w_h)_{\mathcal{L}_\mu} + \mathcal{A}_\mu(V_h, V_h - w_h) = 0 \quad .
$$

Adding these two equations, we therefore have that

$$
\begin{aligned}
0 =\ & ([V]_t, e_h)_{\mathcal{L}_\mu} + \mathcal{A}_\mu(V, e_h) + \ ([V_h]_t, V_h - w_h)_{\mathcal{L}_\mu} + \mathcal{A}_\mu(V_h, V_h - w_h) \\
=\ & ([V]_t, e_h)_{\mathcal{L}_\mu} + \mathcal{A}_\mu(V, e_h) + ([V_h]_t, V_h - w_h)_{\mathcal{L}_\mu} + \mathcal{A}_\mu(V_h, V_h - w_h)
\end{aligned}
$$

$$
\begin{aligned}
&- ([V_h]_t, V)_{\mathcal{L}_\mu} - \mathcal{A}_\mu(V_h, V) + ([V_h]_t, V)_{\mathcal{L}_\mu} + \mathcal{A}_\mu(V_h, V) \\
&= ([V]_t, e_h)_{\mathcal{L}_\mu} + ([V_h]_t, V_h)_{\mathcal{L}_\mu} - ([V_h]_t, V)_{\mathcal{L}_\mu} + ([V_h]_t, V - w_h)_{\mathcal{L}_\mu} \\
&\quad + \mathcal{A}_\mu(V_h, V - w_h) + \mathcal{A}_\mu(V_h, V_h) - \mathcal{A}_\mu(V_h, V) + \mathcal{A}_\mu(V, e_h) \\
&= ([V]_t, e_h)_{\mathcal{L}_\mu} - ([V_h]_t, e_h)_{\mathcal{L}_\mu} + ([V_h]_t, V - w_h)_{\mathcal{L}_\mu} \\
&\quad + \mathcal{A}_\mu(V_h, V - w_h) - \mathcal{A}_\mu(V_h, e_h) + \mathcal{A}_\mu(V, e_h) \\
&= ([e_h]_t, e_h)_{\mathcal{L}_\mu} + \mathcal{A}_\mu(e_h, e_h) + ([V_h]_t, V - w_h)_{\mathcal{L}_\mu} + \mathcal{A}_\mu(V_h, V - w_h) \ .
\end{aligned}
$$

We therefore have that

$$
t([e_h]_t, e_h)_{\mathcal{L}_\mu} + t\mathcal{A}_\mu(e_h, e_h) = t([V_h]_t, w_h - V)_{\mathcal{L}_\mu} + t\mathcal{A}_\mu(V_h, w_h - V) \quad ,
$$

or,

$$
\frac{1}{2}\frac{d}{dt}\|te_h\|^2_{\mathcal{L}_\mu} + t\mathcal{A}_\mu(e_h, e_h) = \frac{1}{2}\|e_h\|_{\mathcal{L}_\mu} + t([V_h]_t, w_h - V)_{\mathcal{L}_\mu} + t\mathcal{A}_\mu(V_h, w_h - V) \quad .
$$

The application of the Cauchy-Schwarz inequality, as well as (4.20) and (4.21) then yields,

$$
\begin{aligned}
&\frac{1}{2}\frac{d}{dt}\|te_h\|^2_{\mathcal{L}_\mu} + \alpha t\|e_h\|^2_{\mathcal{W}_\mu} \\
&\qquad \leq \frac{1}{2}\|e_h\|^2_{\mathcal{L}_\mu} + t\|[V_h]_t\|_{\mathcal{L}_\mu}\|V - w_h\|_{\mathcal{L}_\mu} \\
&\qquad\quad + \gamma t\|V_h\|_{\mathcal{W}_\mu}\|V - w_h\|_{\mathcal{W}_\mu} \quad .
\end{aligned}
$$

Integrating this identity over $[0, \hat{t}]$,for any $\hat{t} \in (0, t_1]$, then implies

$$
\begin{aligned}
&\frac{1}{2}\|\hat{t}e_h\|^2_{\mathcal{L}_\mu} + \alpha \int_0^{\hat{t}} t\|e_h(t)\|^2_{\mathcal{W}_\mu}\mathrm{d}t \\
&\qquad \leq \ \frac{1}{2}\int_0^{\hat{t}} t\|e_h(t)\|^2_{\mathcal{L}_\mu}\mathrm{d}t + \int_0^{\hat{t}} t\|V_h'(t)\|_{\mathcal{L}_\mu}\|V(t) - w_h\|_{\mathcal{L}_\mu}\mathrm{d}t \\
&\qquad\quad + \gamma \int_0^{\hat{t}} t\|V_h(t)\|_{\mathcal{W}_\mu}\|V(t) - w_h\|_{\mathcal{W}_\mu}\mathrm{d}t \quad ,
\end{aligned}
$$

applying Hölder's inequality,

$$\leq \frac{1}{2}\int_0^{\hat{t}} t\|e_h(t)\|_{\mathcal{L}_\mu}\mathrm{d}t + \left(\int_0^{\hat{t}} t^2\|V_h'(t)\|^2_{\mathcal{L}_\mu}\mathrm{d}t\right)^{1/2}\left(\int_0^{\hat{t}} \|V(t)-w_h\|^2_{\mathcal{L}_\mu}\mathrm{d}t\right)^{1/2} + \gamma\left(\int_0^{\hat{t}} t^2\|V_h(t)\|^2_{\mathcal{W}_\mu}\mathrm{d}t\right)^{1/2}\left(\int_0^{\hat{t}} \|V(t)-w_h\|_{\mathcal{W}_\mu}\mathrm{d}t\right)^{1/2},$$

applying Lemmas, 4.18, 4.19 and 4.25,

$$\leq C\inf_{v_h\in\mathcal{W}_h}\int_0^{\hat{t}} \|V(t)-v_h\|_{\mathcal{W}_\mu}\mathrm{d}t + C\|V_0\|_{\mathcal{L}_\mu}\left(\int_0^{\hat{t}} \|V(t)-w_h\|^2_{\mathcal{L}_\mu}\mathrm{d}t\right)^{1/2} + Ct_1\|V_0\|_{\mathcal{L}_\mu}\left(\int_0^{\hat{t}} \|V(t)-w_h\|_{\mathcal{W}_\mu}\mathrm{d}t\right)^{1/2}.$$

Taking the infimum over all $w_h\in\mathcal{W}_h$ and selecting $\hat{t}=t_1$ then yields the required estimate,

$$\|\hat{t}e_h\|^2_{\mathcal{L}_\mu} + 2\alpha\int_0^{\hat{t}} t\|e_h(t)\|^2_{\mathcal{W}_\mu}\mathrm{d}t \leq C\inf_{v_h\in\mathcal{W}_h}\|V-v_h\|_{L^2(0,\hat{t},\mathcal{W}_\mu)} + Ct_1\|V_0\|_{\mathcal{L}_\mu}\inf_{w_h\in\mathcal{W}_h}\|V-w_h\|_{L^2(0,\hat{t},\mathcal{W}_\mu)}.$$

We now recall that due to Theorem 4.13, for any $\epsilon>0$ and $\phi\in C_0^\infty(\Omega)$, with $\Omega=[0,x_{max}]$ and $x_{max}=h^{-c}$ for some $c>0$

$$\inf_{u_h\in\mathcal{W}_h}\|u-u_h\|_{\mathcal{W}_\mu} \leq \|u-\tilde{u}\|_{\mathcal{W}_\mu} + \epsilon + \|\phi-\Pi_h\phi\|_{\mathcal{W}_\mu},$$

and hence, due to the triangle inequality we have that

$$\inf_{w_h\in\mathcal{W}_h}\|V-w_h\|_{L^2(0,\hat{t},\mathcal{W}_\mu)} \leq \|V-\tilde{V}\|_{L^2(0,\hat{t},\mathcal{W}_\mu)} + \|\epsilon\|_{L^2(0,\hat{t},\mathcal{W}_\mu)}$$

$$+ \|\phi - \Pi_h \phi\|_{L^2(0,\hat{t},\mathcal{W}_\mu)} \quad .$$

Now, since $supp(\phi) \Subset \mathbb{R}^+$ and $\omega_\mu \leq 1$, we have that there exists $C > 0$ such that

$$\begin{aligned} \|\phi - \Pi_h \phi\|_{\mathcal{W}_\mu} &\leq \|\phi - \Pi_h \phi\|_{W^{1,2}(supp(\phi),\omega_0)} \\ &\leq hC\|\phi\|_{W^{2,2}(supp(\phi),\omega_0)} \quad . \end{aligned}$$

Furthermore, recalling Lemma 2.4, it is easy to show that there exists $C > 0$ such that

$$\|V - \tilde{V}\|_{\mathcal{W}_\mu} \leq C\|V\|_{\mathcal{W}_\mu} \in L^2(0, t_1)$$

and hence, due to Lemma 2.4 and the Lebesgue dominated convergence Theorem, we have that

$$\lim_{h\to 0} \|V - \tilde{V}\|_{L^2(0,t_1,\mathcal{W}_\mu)} = 0 \quad .$$

Hence, if we set $\epsilon = h$ we have that

$$\begin{aligned} \lim_{h\to\infty} \inf_{w_h \in \mathcal{W}_h} \|V - w_h\|_{L^2(0,\hat{t},\mathcal{W}_\mu)} &\leq \lim_{h\to\infty} \|V - \tilde{V}\|_{L^2(0,\hat{t},\mathcal{W}_\mu)} + \lim_{h\to\infty} \|\epsilon\|_{L^2(0,\hat{t},\mathcal{W}_\mu)} \\ &\quad + \lim_{h\to\infty} \|\phi - \Pi_h \phi\|_{L^2(0,\hat{t},\mathcal{W}_\mu)} \\ &\leq \lim_{h\to\infty} h\|1\|_{L^2(0,\hat{t},\mathcal{W}_\mu)} + \lim_{h\to\infty} h\|\phi\|_{L^2(0,\hat{t},W^{2,2}(supp(\phi),\omega_0))} \\ &= 0 \quad . \end{aligned}$$

It therefore follows that

$$\lim_{h\to 0} \|t_1 e_h\|^2_{\mathcal{L}_\mu} + 2\alpha \int_0^{t_1} t\|e_h(t)\|^2_{\mathcal{W}_\mu} = 0 \quad ,$$

and hence, as required, V_h converges to the option value V in $\mathcal{W}_\mu$. □

Chapter 5

Numerical Methods

To conclude our investigation into the use of the infinite element method within the field of option pricing, we will now present a brief view of the implementation of this method numerically. In this regard, we begin this chapter by following the work of Sanfelici [33] and rewrite the semi-discrete problem in vector form. We then proceed to apply a finite difference scheme to discretise the time derivative and obtain a fully discrete vector problem. Finally, we will calculate the elements of the mass and stiffness matrices and thereby present the infinite element method in a form that may easily be applied to obtain a solution to the valuation problem.

In Chapter 3 we noted that Problem 5 may be reformulated so that for each $m = 0, 1, 2, ..., M-1$, we must search for the coefficients $\alpha_0(t), \alpha_1(t), ..., \alpha_{N+1}(t)$ that satisfy

$$\sum_{i=0}^{N+1} [\alpha_i(t)]_t (\hat{\varphi}_i, \hat{\varphi}_j)_{\mathcal{L}_\mu} + \sum_{i=0}^{N+1} \alpha_i(t) \mathcal{A}_\mu(\hat{\varphi}_i, \hat{\varphi}_j) = 0 \quad , \tag{5.1}$$

for $j = 0, 1, 2, ..., N+1$ and $t_m < t < t_{m+1}$.

Recalling that the basis functions $\hat{\varphi}_i$ were chosen such that for $i = 0, 1, ..., N$,

$$\hat{\varphi}_i(x_j) = \begin{cases} 1 & \text{if} \quad i = j \\ 0 & \text{otherwise} \end{cases} \quad ,$$

it then follows from equation (3.2) that we must have

$$V_h(x_i, t) = \alpha_i(t) \;\text{ for each } i = 0, 1, ...N \quad .$$

This fact motivates us making use of the vector notation

$$\mathbf{V}(t) := [\alpha_0(t), v_1(t), ..., \alpha_N(t)]^T \quad .$$

Following Sanfelici [33], equation (5.1) may then be rewritten as

$$\mathbf{M}[\mathbf{V}(t)]_t + \mathbf{A}\mathbf{V}(t) = 0 \quad , \tag{5.2}$$

where the matrix $\mathbf{M}$, termed the mass matrix, is given by $\mathbf{M} := [(\hat{\varphi}_i, \hat{\varphi}_j)_{\mathcal{L}_\mu}]_{i,j=0,1,2,...,N+1}$ and the stiffness matrix $\mathbf{A}$, is given by $\mathbf{A} := [\mathcal{A}_\mu(\hat{\varphi}_i, \hat{\varphi}_j)]_{i,j=0,1,2,...,N+1}$.
Under this notation, the initial conditions $V_h(t_m) = V_{t_m,h}$, imply that we must have

$$\mathbf{V}(t_m) = \mathbf{V}_{t_m} = [V_{t_m,h}(x_0), V_{t_m,h}(x_1), ..., V_{t_m,h}(x_N)]^T \quad .$$

Following standard practice, we will now derive a fully discrete version of this problem by discretising the time derivative via a finite difference scheme. While there are many possible finite difference schemes with which this may be achieved, we will follow the work of Sanfelici and apply the well-known Crank-Nicholson method (θ method, with $\theta = 1/2$). To this end, we begin by recalling that we may approximate the time derivative by

$$[\mathbf{V}]_t(t) \approx \frac{\mathbf{V}(t + \Delta t) - \mathbf{V}(t)}{\Delta t}$$

for small $\Delta t > 0$. Bearing this approximation in mind, we then divide each time interval $[t_m, t_{m+1}]$ into N^* subintervals of length $\Delta t = [t_{m+1} - t_m]/N^*$ and then replace equation (5.2) by the fully discrete system

$$\mathbf{M}\left[\frac{\mathbf{V}^{j+1} - \mathbf{V}^j}{\Delta t}\right] + \frac{1}{2}\mathbf{A}\left[\mathbf{V}^{j+1} + \mathbf{V}^j\right] = 0,$$

or

$$\left[\mathbf{M} + \frac{1}{2}\Delta t\mathbf{A}\right]\mathbf{V}^{j+1} = \left[\mathbf{M} - \frac{1}{2}\Delta t\mathbf{A}\right]\mathbf{V}^{j} \ , \tag{5.3}$$

for $j = 0, 1, 2, ..., N^* - 1$, and $\mathbf{V}^j = \mathbf{V}(t_j)$.

In order to solve this system numerically, we must clearly be able to calculate the matrices $\mathbf{M}$ and $\mathbf{A}$. To this end, we recall from Chapter 3 that one of the key requirements when constructing the infinite element spaces, was to ensure that there existed a basis consisting of functions with "small supports". Many readers may, in the context of Chapter 3, have questioned the relevance of this condition. We will however now demonstrate that it plays a crucial role in ensuring that the calculation of the mass and stiffness matrices is practical.

The functions $\hat{\varphi}_i$ that serve as a basis for the infinite element space $\mathcal{W}_h$, were carefully chosen to have "small" supports restricted to at most two elements,

$$\text{supp}\left[\hat{\varphi}_i\right] = I_i \cup I_{i+1} \quad .$$

It therefore follows that any product of the form $\hat{\varphi}_i\hat{\varphi}_j$ vanishes unless $j = i-1, i+1$. The selection of these carefully chosen basis functions thus reduces $\mathbf{M}$ and $\mathbf{A}$ to tri-diagonal matrices and as such, sharply reduces the computations required to solve the system of equations (5.3).

We will now proceed to derive expressions for the non-zero elements of $\mathbf{M}$ and $\mathbf{A}$. We begin by noting that since the basis functions have the same shape outside of the infinite element, we clearly have that for $i = 0, 1, ..., N-1$ and $j = 1, 2, 3..., N$

$$\mathbf{M}_{i,i+1} = (\hat{\varphi}_i, \hat{\varphi}_{i+1})_{\mathcal{L}_\mu} = \mathbf{M}_{j,j-1} = (\hat{\varphi}_j, \hat{\varphi}_{j-1})_{\mathcal{L}_\mu} = (\hat{\varphi}_0, \hat{\varphi}_1)_{\mathcal{L}_\mu}$$

and for $i = 1, 2, ..., N-1$

$$\mathbf{M}_{i,i} = (\hat{\varphi}_i, \hat{\varphi}_i)_{\mathcal{L}_\mu} = (\hat{\varphi}_1, \hat{\varphi}_1)_{\mathcal{L}_\mu} \quad ,$$

where

$$\begin{aligned}
(\hat{\varphi}_0, \hat{\varphi}_1)_{\mathcal{L}_\mu} &= \int_{\mathbb{R}^+} \hat{\varphi}_0(x)\hat{\varphi}_1(x)\mathrm{d}x \\
&= \int_0^h \left(\frac{1}{h}\right)[h-x]\left(\frac{1}{h}\right)[x]\,\mathrm{d}x \\
&= \int_0^h \left(\frac{1}{h^2}\right)\left[-x^2 + xh\right]\mathrm{d}x \\
&= \left[\frac{-x^3}{3h^2} + \frac{x^2}{2h} + \right]\Bigg|_0^h \\
&= \frac{h}{6}
\end{aligned}$$

and

$$\begin{aligned}
(\hat{\varphi}_1, \hat{\varphi}_1)_{\mathcal{L}_\mu} &= \int_{\mathbb{R}^+} \hat{\varphi}_1^2(x)\mathrm{d}x \\
&= \int_{I_1} \left(\frac{1}{h^2}\right)[x]^2\,\mathrm{d}x + \int_{I_2} \left(\frac{1}{h^2}\right)[2h-x]^2\,\mathrm{d}x \\
&= \int_0^h \left(\frac{1}{h^2}\right)[x]^2\,\mathrm{d}x + \int_h^{2h} \left(\frac{1}{h^2}\right)[2h-x]^2\,\mathrm{d}x \\
&= \left[\frac{x^3}{3h^2}\right]\Bigg|_0^h + \left[\frac{x^3}{3h^2} - \frac{2x^2}{h} + 4x\right]\Bigg|_h^{2h} \\
&= \frac{2h}{3} \quad .
\end{aligned}$$

Finally,

$$\begin{aligned}
\mathbf{M}_{0,0} &= (\hat{\varphi}_0, \hat{\varphi}_0)_{\mathcal{L}_\mu} \\
&= \int_0^h \left(\frac{x}{h}\right)^2\mathrm{d}x \\
&= \frac{h}{3} \quad .
\end{aligned}$$

We now proceed to consider the elements of the stiffness matrix $\mathcal{A}$. Under standard circumstances the consistent shape of the basis functions would cause the values of the elements along each of the three non-zero diagonal rows to coincide. The presence of weight functions within the required integrals however, implies that this is no longer the case and as such we will derive more general expressions for the required terms.

Recalling that

$$\begin{aligned}\mathcal{A}_\mu(u,v) = &\int_{\mathbb{R}^+} \frac{1}{2}\sigma^2\omega_\mu^2(x)x^2 D^{(1)}[u(x)]D^{(1)}[v(x)]\mathrm{d}x \\ &+\int_{\mathbb{R}^+} A(x)\omega_\mu^2(x)xD^{(1)}[u(x)]v(x)\mathrm{d}x \;+\int_{\mathbb{R}^+} r\omega_\mu^2(x)u(x)v(x)\mathrm{d}x \quad ,\end{aligned}$$

we then calculate:

$$\begin{aligned}\frac{\sigma^2}{2}\int_{I_i} x^2[\varphi_i'(x)]^2\mathrm{d}x &= \frac{\sigma^2}{2}\int_{h(i-1)}^{hi} x^2\left[\frac{1}{h}\right]^2\mathrm{d}x \\ &= \frac{\sigma^2}{2}\left[\frac{x^3}{3h^2}\right]\Big|_{h(i-1)}^{h} \\ &= \frac{\sigma^2}{2}\left[\frac{hi^3}{3} - \frac{h}{3}(i^3-3i^2+3i-1)\right] \\ &= \frac{\sigma^2 h}{6}\left[3i^2-3i+1\right] \\ (\sigma^2+r)\int_{I_i} x[\varphi_i'(x)][\varphi_i(x)]\mathrm{d}x &= (\sigma^2+r)\int_{h(i-1)}^{hi} x\left[\frac{1}{h}\right]\left[\frac{x-h(i-1)}{h}\right]\mathrm{d}x \\ &= (\sigma^2+r)\left[\frac{x^3}{3h} - \frac{x^2(i-1)}{2h}\right]\Big|_{h(i-1)}^{hi} \\ &= (\sigma^2+r)\left[\frac{h(3i-1)}{6}\right] \\ r\int_{I_i}[\varphi_i(x)]^2\mathrm{d}x &= r\int_{h(i-1)}^{hi}\left[\frac{x-h(i-1)}{h}\right]^2\mathrm{d}x \\ &= r\left[\frac{x^3}{3h} - \frac{x^2(i-1)}{h} + (i-1)^2x\right]\Big|_{h(i-1)}^{hi} \\ &= \frac{rh}{3}\end{aligned}$$

$$
\begin{aligned}
\frac{\sigma^2}{2}\int_{I_{i+1}} x^2[\varphi_i'(x)]^2\mathrm{d}x &= \frac{\sigma^2}{2}\int_{h(i)}^{h(i+1)} x^2\left[\frac{1}{h}\right]^2\mathrm{d}x\\
&= \frac{\sigma^2}{2}\left[\frac{x^3}{3h^2}\right]\Big|_{hi}^{h(i+1)}\\
&= \frac{\sigma^2 h}{6}\left[3i^2+3i+1\right]\\
(\sigma^2+r)\int_{I_{i+1}} x[\varphi_i'(x)][\varphi_i(x)]\mathrm{d}x &= (\sigma^2+r)\int_{hi}^{h(i+1)} x\left[\frac{1}{h}\right]\left[\frac{h(i+1)-x}{h}\right]\mathrm{d}x\\
&= (\sigma^2+r)\left[\frac{x^2(i+1)}{2h}-\frac{x^3}{3h}\right]\Big|_{hi}^{h(i+1)}\\
&= (\sigma^2+r)\left[\frac{h(3i+1)}{6}\right]
\end{aligned}
$$

$$
\begin{aligned}
r\int_{I_{i+1}}[\varphi_i(x)]^2\mathrm{d}x &= r\int_{hi}^{h(i+1)}\left[\frac{h(i+1)-x}{h}\right]^2\mathrm{d}x\\
&= r\left[(i+1)^2x-\frac{x^2(i+1)}{h}+\frac{x^3}{3h}\right]\Big|_{h(i-1)}^{hi}\\
&= \frac{rh}{3}\\
\frac{\sigma^2}{2}\int_{I_i} x^2[\varphi_i'(x)][\varphi_{i-1}'(x)]\mathrm{d}x &= \frac{\sigma^2}{2}\int_{h(i-1)}^{hi} x^2\left[\frac{1}{h}\right]\left[-\frac{1}{h}\right]\mathrm{d}x\\
&= -\frac{\sigma^2}{2}\left[\frac{x^3}{3h^2}\right]\Big|_{h(i-1)}^{h}\\
&= -\frac{\sigma^2}{2}\left[\frac{hi^3}{3}-\frac{h}{3}(i^3-3i^2+3i-1)\right]\\
&= -\frac{\sigma^2}{6}\left[3i^2-3i+1\right]\\
(\sigma^2+r)\int_{I_i} x[\varphi_i'(x)][\varphi_{i-1}(x)]\mathrm{d}x &= (\sigma^2+r)\int_{h(i-1)}^{hi} x\left[\frac{1}{h^2}\right][h(i)-x]\,\mathrm{d}x\\
&= (\sigma^2+r)\left[\frac{x^2 i}{2h}-\frac{x^3}{3h^2}\right]\Big|_{h(i-1)}^{hi}\\
&= (\sigma^2+r)\left[\frac{3hi-2h}{6}\right]
\end{aligned}
$$

$$
\begin{aligned}
r\int_{I_i}[\varphi_i(x)][\varphi_{i-1}(x)]\mathrm{d}x &= r\int_{h(i-1)}^{hi}\left[\frac{x-h(i-1)}{h}\right]\left[\frac{hi-x}{h}\right]\mathrm{d}x \\
&= r\left[\frac{x^2 i}{2h}-\frac{x^3}{3h}-(i^2-i)x+\frac{x^2(i-1)}{2h}\right]\Bigg|_{h(i-1)}^{hi} \\
&= \frac{rh}{6} \\
\frac{\sigma^2}{2}\int_{I_i}x^2[\varphi_{i-1}'(x)][\varphi_i'(x)]\mathrm{d}x &= \frac{\sigma^2}{2}\int_{h(i-1)}^{hi}x^2\left[-\frac{1}{h}\right]\left[\frac{1}{h}\right]\mathrm{d}x \\
&= -\frac{\sigma^2}{2}\left[\frac{x^3}{3h^2}\right]\Bigg|_{h(i-1)}^{h} \\
&= -\frac{\sigma^2}{2}\left[\frac{hi^3}{3}-\frac{h}{3}(i^3-3i^2+3i-1)\right] \\
&= -\frac{\sigma^2 h}{6}(3i^2-3i+1) \\
(\sigma^2+r)\int_{I_i}x[\varphi_{i-1}'(x)][\varphi_i(x)]\mathrm{d}x &= (\sigma^2+r)\int_{h(i-1)}^{hi}x\left[-\frac{1}{h^2}\right][x-h(i-1)]\,\mathrm{d}x \\
&= (\sigma^2+r)\left[-\frac{x^3}{3h^2}+\frac{x^2h(i-1)}{2h^2}\right]\Bigg|_{h(i-1)}^{hi} \\
&= (\sigma^2+r)\left[\frac{3hi-h}{6}\right] \\
r\int_{I_i}[\varphi_{i-1}(x)][\varphi_i(x)]\mathrm{d}x &= r\int_{h(i-1)}^{hi}\left[\frac{hi-x}{h}\right]\left[\frac{x-h(i-1)}{h}\right]\mathrm{d}x \\
&= r\left[\frac{x^2 i}{2h}-\frac{x^3}{3h}-(i^2-i)x+\frac{x^2(i-1)}{2h}\right]\Bigg|_{h(i-1)}^{hi} \\
&= \frac{rh}{6}\quad .
\end{aligned}
$$

Making use of these integrals, we may then calculate expressions for the non-zero elements of the stiffness matrix as follows:

$$
\begin{aligned}
\mathcal{A}_{0,0} &= \frac{\sigma^2}{2}\int_{I_1}x^2[\hat{\varphi}_0{}'(x)]^2\mathrm{d}x-(\sigma^2+r)\int_{I_1}[\varphi_0'(x)][\varphi_0(x)]\mathrm{d}x \\
&\quad +r\int_{I_1}[\varphi_0(x)]^2\mathrm{d}x \\
&= \frac{\sigma^2}{6}\left[3i^2+3i+1\right]-(\sigma^2+r)\left[\frac{h(3i+1)}{6}\right]+r\frac{h}{3}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{rh}{6} \\
\mathcal{A}_{i,i} &= \frac{\sigma^2}{2}\int_{I_i} x^2[\hat{\varphi}_i{}'(x)]^2\mathrm{d}x - (\sigma^2+r)\int_{I_i}[\varphi_i'(x)][\varphi_i(x)]\mathrm{d}x \\
&\quad + r\int_{I_i}[\varphi_i(x)]^2\mathrm{d}x + \frac{\sigma^2}{2}\int_{I_{i+1}} x^2[\hat{\varphi}_i{}'(x)]^2\mathrm{d}x \\
&\quad - (\sigma^2+r)\int_{I_{i+1}}[\varphi_i'(x)][\varphi_i(x)]\mathrm{d}x + r\int_{I_{i+1}}[\varphi_i(x)]^2\mathrm{d}x \\
&= \frac{\sigma^2 h}{6}\left[3i^2-3i+1\right] - (\sigma^2+r)\left[\frac{h(3i-1)}{6}\right] + \frac{rh}{3} \\
&\quad + \frac{\sigma^2 h}{6}\left[3i^2+3i+1\right] - (\sigma^2+r)\left[\frac{3hi-2h}{6}\right] + \frac{rh}{3} \\
\mathcal{A}_{i,i-1} &= \frac{\sigma^2}{2}\int_{I_i} x^2[\hat{\varphi}_i'(x)][\hat{\varphi}_{i-1}'(x)]\mathrm{d}x - (\sigma^2+r)\int_{I_i}[\hat{\varphi}_i'(x)][\hat{\varphi}_{i-1}(x)]\mathrm{d}x \\
&\quad + r\int_{I_i}[\hat{\varphi}_i(x)][\hat{\varphi}_{i+1}(x)]\mathrm{d}x \\
&= -\frac{\sigma^2}{6}\left[3i^2-3i+1\right] - (\sigma^2+r)\left[\frac{3hi-2h}{6}\right] + \frac{rh}{6} \\
\mathcal{A}_{i-1,i} &= \frac{\sigma^2}{2}\int_{I_i} x^2[\hat{\varphi}_{i-1}'(x)][\hat{\varphi}_i'(x)]\mathrm{d}x - (\sigma^2+r)\int_{I_i}[\hat{\varphi}_{i-1}'(x)][\hat{\varphi}_i(x)]\mathrm{d}x \\
&\quad + r\int_{I_i}[\hat{\varphi}_{i-1}(x)][\hat{\varphi}_i(x)]\mathrm{d}x \\
&= -\frac{\sigma^2 h}{6}(3i^2-3i+1) - (\sigma^2+r)\left[\frac{3hi-h}{6}\right] + \frac{rh}{6}
\end{aligned}
$$

It now remains to calculate the elements $\mathbf{M}$ and $\mathbf{A}$ that require integration over the infinite element. In these calculations we will assume that the weight parameter, $\mu = 2$. The required integration is easily performed via the method of partial fractions, however due to the number of terms required for the partial fraction decomposition (in some cases as many as 6 per integral), we have only computed one of the mass matrix elements and evaluated the remaining elements of the mass and stiffness matrices via Mathematica. We note that Sanfelici [33] computes these integrals numerically and suggests either the Gauss-Legendre method over the reference interval or the Gauss-

Laguerre method over the infinite element.

$$
\begin{aligned}
\mathbf{M}_{N,N} &= (\hat{\varphi}_N, \hat{\varphi}_N)_{\mathcal{L}_\mu} \\
&= (\varphi_{N-1}, \varphi_{N-1})_{\mathcal{L}_\mu} + (\varphi_{inf_1}, \varphi_{inf_1})_{\mathcal{L}_\mu} \\
&= \frac{h}{3} + (\varphi_{inf_1}, \varphi_{inf_1})_{\mathcal{L}_\mu}
\end{aligned}
$$

where,

$$
\begin{aligned}
(\varphi_{inf_1}, \varphi_{inf_1})_{\mathcal{L}_\mu} &= \int_{\mathbb{R}} \omega_\mu^2(x)\varphi_{inf_1}^2(x)\mathrm{d}x \\
&= \int_{x_{max}}^{\infty} \left[\frac{x_{max}}{x}\right]^4 \left[\frac{h}{x-(x_{max}-h)}\right]^2 \mathrm{d}x
\end{aligned}
$$

applying a partial fraction decomposition,

$$
\begin{aligned}
&= \int_{x_{max}}^{\infty} \frac{x_{max}^4 h^2}{x^4(x_{max}-h)^2} + \frac{2x_{max}^4 h^2}{x^3(x_{max}-h)^3} + \frac{3x_{max}^4 h^2}{x^2(x_{max}-h)^4} \\
&\quad + \frac{4x_{max}^4 h^2}{x(x_{max}-h)^5} - \frac{4x_{max}^4 h^2}{(x_{max}-h)^5(x-(x_{max}-h))} \\
&\quad + \frac{4x_{max}^4 h^2}{(x_{max}-h)^4(x-(x_{max}-h))^2}\,\mathrm{d}x \\
&= \frac{x_{max}h^2}{3(x_{max}-h)^2} + \frac{x_{max}^2 h^2}{(x_{max}-h)^3} + \frac{3x_{max}^3 h^2}{(x_{max}-h)^4} \\
&\quad + \left[\frac{4x_{max}^4 h^2}{(x_{max}-h)^5}\right][ln(x)]\Big|_{x_{max}}^{\infty} \\
&\quad - \left[\frac{4x_{max}^4 h^2}{(x_{max}-h)^5}\right][ln(x-(x_{max}-h))]\Big|_{x_{max}}^{\infty} + \frac{x_{max}^4 h}{(x_{max}-h)^4} \\
&= \frac{x_{max}h^2}{3(x_{max}-h)^2} + \frac{x_{max}^2 h^2}{(x_{max}-h)^3} + \frac{3x_{max}^3 h^2}{(x_{max}-h)^4} \\
&\quad + \left[\frac{4x_{max}^4 h^2}{(x_{max}-h)^5}\right][ln(x_{max})] - \left[\frac{4x_{max}^4 h^2}{(x_{max}-h)^5}\right][ln(h)] \\
&\quad + \frac{x_{max}^4 h}{(x_{max}-h)^4} \\
&= \frac{x_{max}h^2}{3(x_{max}-h)^2} + \frac{x_{max}^2 h^2}{(x_{max}-h)^3} + \frac{3x_{max}^3 h^2}{(x_{max}-h)^4}
\end{aligned}
$$

$$+\left[\frac{4x_{max}^4 h^2}{(x_{max}-h)^5}\right] ln\left(\frac{x_{max}}{h}\right)+\frac{x_{max}^4 h}{(x_{max}-h)^4}$$

$$=\frac{x_{max}h\left[3x_{max}^4+10x_{max}^3 h+12x_{max}^3 h\ ln\left(\frac{x_{max}}{h}\right)-18x_{max}^2 h^2+6x_{max}h^3-h^4\right]}{3(x_{max}-h)^5}$$

which agrees with the expression produced by Mathematica:

$$\frac{x_{max}^4 h^2\left(-10+\frac{x_{max}^3}{h^3}-\frac{6x_{max}^2}{h^2}+\frac{18x_{max}}{h}-\frac{3h}{x_{max}}+12ln\left(\frac{h}{x_{max}}\right)\right)}{3(x_{max}-h)^5}$$

Mathematica gives the following output for the remaining mass and stiffness element:

$$\begin{aligned}\mathbf{M}_{N,N+1} &= (\hat{\varphi}_N,\hat{\varphi}_{N+1})_{\mathcal{L}_\mu}\\ &= \int_{x_{max}}^{\infty}\omega_\mu^2(x)\phi_{\text{inf}_1}(x)\phi_{\text{inf}_2}(x)\mathrm{d}x\\ &= \int_{x_{max}}^{\infty}\left[\frac{x_{max}}{x}\right]^4\left[\frac{h}{x-(x_{max}-h)}\right]\left[1-\frac{h}{x-(x_{max}-h)}\right]\mathrm{d}x\end{aligned}$$

$$=\frac{x_{max}^2 h\left(-(x_{max}-h)\left(17x_{max}^2+8x_{max}h-h^2\right)+6x_{max}^2(x_{max}+3h)(ln\left(\frac{x_{max}}{h}\right))\right)}{6(x_{max}-h)^5}$$

$$\begin{aligned}\mathbf{M}_{N+1,N+1} &= (\hat{\varphi}_{N+1},\hat{\varphi}_{N+1})_{\mathcal{L}_\mu}\\ &= \int_{x_{max}}^{\infty}\omega_\mu^2(x)\phi_{\text{inf}_2}(x)\phi_{\text{inf}_2}(x)\mathrm{d}x\\ &= \int_{x_{max}}^{\infty}\left[\frac{x_{max}}{x}\right]^4\left[1-\frac{h}{x-(x_{max}-h)}\right]^2\mathrm{d}x\end{aligned}$$

$$=\frac{x_{max}^3\left((x_{max}-h)\left(x_{max}^2+10x_{max}h+h^2\right)-6x_{max}h(x_{max}+h)(ln\left(\frac{x_{max}}{h}\right))\right)}{3(x_{max}-h)^5}$$

$$\begin{aligned}
\mathcal{A}_{N,N+1} &= \mathcal{A}_\mu(\hat{\varphi}_N, \hat{\varphi}_{N+1}) \\
&= \frac{\sigma^2}{2}\int_{x_{max}}^{\infty} \omega_\mu^2(x) x^2 [\hat{\varphi}'_N(x)][\hat{\varphi}'_{N+1}(x)]\mathrm{d}x \\
&\quad - (\sigma^2 + r)\int_{x_{max}}^{\infty} \omega_\mu^2(x) x [\hat{\varphi}'_N(x)][\hat{\varphi}_{N+1}(x)]\mathrm{d}x \\
&\quad + r\int_{x_{max}}^{\infty} \omega_\mu^2(x) [\hat{\varphi}_N(x)][\hat{\varphi}_{N+1}(x)]\mathrm{d}x \\
\mathcal{A}_{N+1,N} &= \mathcal{A}_\mu(\hat{\varphi}_{N+1}, \hat{\varphi}_N) \\
&= \frac{\sigma^2}{2}\int_{x_{max}}^{\infty} \omega_\mu^2(x) x^2 [\hat{\varphi}'_{N+1}(x)][\hat{\varphi}'_N(x)]\mathrm{d}x \\
&\quad - (\sigma^2 + r)\int_{x_{max}}^{\infty} \omega_\mu^2(x) x [\hat{\varphi}'_{N+1}(x)][\hat{\varphi}_N(x)]\mathrm{d}x \\
&\quad + r\int_{x_{max}}^{\infty} \omega_\mu^2(x) [\hat{\varphi}_{N+1}(x)][\hat{\varphi}_N(x)]\mathrm{d}x
\end{aligned}$$

$$\begin{aligned}
\mathcal{A}_{N,N} &= \mathcal{A}_\mu(\hat{\varphi}_N, \hat{\varphi}_N) \\
&= \frac{\sigma^2}{2}\int_{I_N} x^2 [\hat{\varphi}'_N(x)]^2\mathrm{d}x + \frac{\sigma^2}{2}\int_{x_{max}}^{\infty} \omega_\mu^2(x) x^2 [\hat{\varphi}'_N(x)]^2\mathrm{d}x \\
&\quad - (\sigma^2 + r)\int_{I_N} x [\hat{\varphi}'_N(x)][\hat{\varphi}_N(x)]\mathrm{d}x \\
&\quad - (\sigma^2 + r)\int_{x_{max}}^{\infty} \omega_\mu^2(x) x [\hat{\varphi}'_N(x)][\hat{\varphi}_N(x)]\mathrm{d}x \\
&\quad + r\int_{I_N} [\hat{\varphi}_N(x)]^2\mathrm{d}x + r\int_{x_{max}}^{\infty} \omega_\mu^2(x) [\hat{\varphi}_N(x)]^2\mathrm{d}x \quad , \\
\mathcal{A}_{N+1,N+1} &= \mathcal{A}_\mu(\hat{\varphi}_{N+1}, \hat{\varphi}_{N+1}) \\
&= \frac{\sigma^2}{2}\int_{x_{max}}^{\infty} \omega_\mu^2(x) x^2 [\hat{\varphi}'_{N+1}(x)]^2\mathrm{d}x \\
&\quad - (\sigma^2 + r)\int_{x_{max}}^{\infty} \omega_\mu^2(x) x [\hat{\varphi}'_{N+1}(x)][\hat{\varphi}_{N+1}(x)]\mathrm{d}x \\
&\quad + r\int_{x_{max}}^{\infty} \omega_\mu^2(x) [\hat{\varphi}_{N+1}(x)]^2\mathrm{d}x \quad ,
\end{aligned}$$

where,

$$\int_{x_{max}}^{\infty} \omega_\mu^2(x)x^2[\hat{\varphi}_N'(x)]^2 \mathrm{d}x = \int_{x_{max}}^{\infty} \left[\frac{x_{max}^4}{x^2}\right]\left[-\frac{h}{(x-(x_{max}-h))^2}\right]^2 \mathrm{d}x$$

$$= \frac{x_{max}^4 h^2 \left(-10 + \frac{x_{max}^3}{h^3} - \frac{6x_{max}^2}{h^2} + \frac{18x_{max}}{h} - \frac{3h}{x_{max}} + 12ln\left[\frac{h}{x_{max}}\right]\right)}{3(x_{max}-h)^5}$$

$$\int_{x_{max}}^{\infty} \omega_\mu^2(x)x^2[\hat{\varphi}_{N+1}'(x)]^2 \mathrm{d}x = \int_{x_{max}}^{\infty} \left[\frac{x_{max}^4}{x^2}\right]\left[\frac{h}{(x-(x_{max}-h))^2}\right]^2 \mathrm{d}x$$

$$= \frac{x_{max}^4 h^2 \left(-10 + \frac{x_{max}^3}{h^3} - \frac{6x_{max}^2}{h^2} + \frac{18x_{max}}{h} - \frac{3h}{x_{max}} + 12ln\left[\frac{h}{x_{max}}\right]\right)}{3(x_{max}-h)^5}$$

$$\int_{x_{max}}^{\infty} \omega_\mu^2(x)x[\hat{\varphi}_N'(x)][\hat{\varphi}_N(x)]\mathrm{d}x$$

$$= \int_{x_{max}}^{\infty} \left[\frac{x_{max}^4}{x^3}\right]\left[-\frac{h}{(x-(x_{max}-h))^2}\right]\left[\frac{h}{x-(x_{max}-h)}\right] \mathrm{d}x$$

$$= -\frac{a^4 h^2 \left(\frac{(a-h)(a+h)\left(x_{max}^2 - 8x_{max}h + h^2\right)}{x_{max}^2 h^2} - 12ln\left[\frac{h}{x_{max}}\right]\right)}{2(x_{max}-h)^5}$$

$$\int_{x_{max}}^{\infty} \omega_\mu^2(x)[\hat{\varphi}_N(x)]^2 \mathrm{d}x$$

$$= \int_{x_{max}}^{\infty} \left[\frac{x_{max}^4}{x^4}\right]\left[\frac{h}{x-(x_{max}-h)}\right]^2 \mathrm{d}x$$

$$= \frac{x_{max}^4 h^2 \left(10 + \frac{3x_{max}}{h} - \frac{18h}{x_{max}} + \frac{6h^2}{x_{max}^2} - \frac{h^3}{x_{max}^3} + 12ln\left[\frac{h}{x_{max}}\right]\right)}{3(x_{max}-h)^5}$$

$$\int_{x_{max}}^{\infty} \omega_\mu^2(x)x[\hat{\varphi}_{N+1}'(x)][\hat{\varphi}_{N+1}(x)]\mathrm{d}x$$

$$= \int_{x_{max}}^{\infty} \left[\frac{x_{max}^4}{x^3}\right]\left[\frac{h}{(x-(x_{max}-h))^2}\right]\left[1 - \frac{h}{x-(x_{max}-h)}\right] \mathrm{d}x$$

$$= \frac{1}{2(x_{max}-h)^5} x_{max}^2(-x_{max}^4(-2+h) - 9x_{max}^2h^2 + x_{max}(7-8h)h^3$$
$$+ (-1+h)h^4 + x_{max}^3 h(1+8h) + 6x_{max}^2 h(x_{max} + h(-1+2h))\text{Log}\left[\frac{h}{x_{max}}\right])$$

$$\int_{x_{max}}^{\infty} \omega_\mu^2(x)[\hat{\varphi}_{N+1}(x)]^2 \mathrm{d}x$$
$$= \int_{x_{max}}^{\infty} \left[\frac{x_{max}^4}{x^4}\right]\left[1 - \frac{h}{x-(x_{max}-h)}\right]^2 \mathrm{d}x$$
$$= \frac{x_{max}^3\left((x_{max}-h)\left(x_{max}^2 + 10x_{max}h + h^2\right) - 6x_{max}h(x_{max}+h)(ln[\frac{x_{max}}{h}])\right)}{3(x_{max}-h)^5}$$

$$\int_{x_{max}}^{\infty} \omega_\mu^2(x)[\hat{\varphi}_{N+1}(x)][\hat{\varphi}_N(x)] \mathrm{d}x$$
$$= \int_{x_{max}}^{\infty} \left[\frac{x_{max}^4}{x^4}\right]\left[1 - \frac{h}{x-(x_{max}-h)}\right]\left[\frac{h}{x-(x_{max}-h)}\right] \mathrm{d}x$$
$$= \frac{x_{max}^2 h\left(-(x_{max}-h)\left(17x_{max}^2 + 8x_{max}h - h^2\right) + 6x_{max}^2(x_{max}+3h)(ln[\frac{x_{max}}{h}])\right)}{6(x_{max}-h)^5}$$

$$\int_{x_{max}}^{\infty} \omega_\mu^2(x)x^2[\hat{\varphi}'_{N+1}(x)][\hat{\varphi}'_N(x)] \mathrm{d}x$$
$$= \int_{x_{max}}^{\infty} \left[\frac{x_{max}^4}{x^2}\right]\left[-\frac{h^2}{(x-(x_{max}-h))^4}\right] \mathrm{d}x$$
$$= -\frac{x_{max}^4 h^2\left(-10 + \frac{x_{max}^3}{h^3} - \frac{6x_{max}^2}{h^2} + \frac{18x_{max}}{h} - \frac{3h}{x_{max}} + 12ln\left[\frac{h}{x_{max}}\right]\right)}{3(x_{max}-h)^5}$$

$$\int_{x_{max}}^{\infty} \omega_\mu^2(x)x[\hat{\varphi}'_{N+1}(x)][\hat{\varphi}_N(x)] \mathrm{d}x$$
$$= \int_{x_{max}}^{\infty} \left[\frac{x_{max}^4}{x^3}\right]\left[\frac{h}{(x-(x_{max}-h))^2}\right]\left[\frac{h}{x-(x_{max}-h)}\right] \mathrm{d}x$$

$$= \frac{x_{max}^4 h^2 \left(\frac{(x_{max}-h)(x_{max}+h)\left(x_{max}^2 - 8x_{max}h + h^2\right)}{x_{max}^2 h^2} - 12ln\left[\frac{h}{x_{max}}\right]\right)}{2(x_{max}-h)^5}$$

$$\int_{x_{max}}^{\infty} \omega_\mu^2(x)x[\hat{\varphi}'_N(x)][\hat{\varphi}_{N+1}(x)]\mathrm{d}x$$
$$= \int_{x_{max}}^{\infty} \left[\frac{x_{max}^4}{x^3}\right]\left[-\frac{h}{(x-(x_{max}-h))^2}\right]\left[1-\frac{h}{x-(x_{max}-h)}\right]\mathrm{d}x$$

$$= -\frac{1}{2(x_{max}-h)^5} x_{max}^2(-x_{max}^4(-2+h) - 9x_{max}^2 h^2 + x_{max}(7-8h)h^3$$
$$+ (-1+h)h^4 + x_{max}^3 h(1+8h) + 6x_{max}^2 h(x_{max} + h(-1+2h))ln\left[\frac{h}{x_{max}}\right])$$

Conclusion

In this dissertation we followed the work of Sanfelici [33] and considered the valuation of discretely monitored barrier options within the context of the infinite element method.

We began our investigation by noting that the Black-Scholes PDE displays a number of properties that make the use of finite element type methods problematic - namely the unbounded spacial domain and the degeneracy that exists in the PDE when the value of the underlying reaches zero. This degeneracy implies that the convergence of finite element type methods (including the infinite element method) should be examined within the context of weighted Sobolev spaces - an extension of standard Sobolev spaces that is not extensively treated within literature. As such, a key aim of this dissertation was to present the reader with a complete introduction to these spaces, as well as a rigorous treatment of the key results that are required to demonstrate convergence within the context of weighted spaces. To this end, we have considered and collated the key works within the field (Kufner [27] and Kufner and Opic [28]) to provide a detailed mathematical introduction within Chapter 1. We have furthermore presented weighted analogs to numerous classical results and derived a number of results that, while critical to the demonstration of convergence, are omitted from the work of Sanfelici, as well as from many similar works within literature. Of particular interest in this regard were the adaptations and translations of the weighted embedding theorems due to Timerbaev [35, 36], the treatment of mollification within weighted norms and the derivation interpolation estimates in weighted norms).

The work contained within this dissertation furthermore completes the work of Sanfelici by providing discussions with regard to the selection of a suitable weighted space in which to perform the analysis; a derivation of the weak formulation of the valuation problem; a rigorous demonstration of existence and uniqueness of the solution of the weak formulation; an introduction to the infinite element method and the selection of basis functions used to generate the associated spaces and a rigorous treatment of the estimation results required within the demonstration of convergence.

The author of this dissertation acknowledges that while we have addressed many of the gaps within the work of Sanfelici and endeavoured to present a near complete treatment of the theoretical aspects of barrier option valuation under the infinite element method, there is still scope for further development within this topic. In particular we note that we have not considered the order of the demonstrated convergence or illustrated the convergence via numerical examples. Sanfelici [33] suggests that under stronger regularity conditions on the initial condition, one is able to prove first order convergence within the $\mathcal{W}_0$ norm. The author of this dissertation notes that a rigorous investigation into this claim is the natural direction in which to extend the work contained within this dissertation and the aims to consider this topic within future research. The author also notes that the treatment of weighted Sobolev theory within the context of convergence analysis, as presented within this dissertation, may be applicable in similar contexts and allow for the rigorous demonstration the convergence of numerical methods defined on weighted spaces.

Appendix A

Sobolev Spaces and the Distributional Derivative

In this appendix we provide a brief introduction to the topics of Sobolev spaces and the distributional derivative. For a deeper treatment of these topics, we direct the reader to the books by Adams and Fournier [2] (Chapter 1 - the distributional derivative; Chapter 3, 4 and 6 - Sobolev spaces), Zeidler [40] (Section 21.1 - the distributional derivative; Sections 21.2, 21.4 and 21.4 - Sobolev Spaces) and Evans [19] (Chapter 5 - Sobolev spaces and the distributional derivative).

We begin by recalling the following well known definitions.

Definition A.1. Let u be a function defined on the domain Ω. The support of u is then defined to be the closure of the subset of Ω on which u assumes a non-zero value. We write

$$supp(u) = \overline{\{x \in \Omega | u(x) \neq 0\}} \quad .$$

Definition A.2. Given $n \in \mathbb{N}$, the space $C^n(\Omega)$ denotes the collection of all functions continuous on Ω, whose first n derivatives are also continuous on Ω.Furthermore, we set

$$C^\infty(\Omega) = \bigcap_{n=1}^{\infty} C^n(\Omega).$$

Definition A.3. Given $n \in \mathbb{N} \cup \{\infty\}$, we denote by $C_0^n(\Omega)$ the subspace of $C^n(\Omega)$ that consists of functions compactly supported in Ω.

We notice that if Ω is open, the above definition implies that functions in the space $C_0^n(\Omega)$ must vanish within some neighbourhood of the boundary of Ω.

Definition A.4. Let $1 \leq p < \infty$, we then define the Lebesgue space $L^p(\Omega)$ as the collection of all measurable functions that satisfy

$$\|u\|_{L^p(\Omega)} = \left(\int_\Omega |u(x)|^p \mathrm{d}x\right)^{1/p} < \infty \ .$$

If $p = \infty$, we require

$$\|u\|_{L^\infty(\Omega)} = \|u\|_{\infty,\Omega} = \sup_{x\in\Omega} |u(x)| < \infty \quad .$$

Definition A.5. A measurable function u defined a.e. on Ω is said to be locally integrable if

$$\int_U u(x)\mathrm{d}x < \infty$$

for every compact $U \subset \Omega$. We write $u \in L^1_{loc}(\Omega)$.

Bearing the above definitions in mind, we may now begin to examine the idea behind Sobolev spaces. Simplistically, we wish to define a Sobolev space as subspace of a Lebesgue space that includes integrability criteria not only for the function itself but also for its derivatives.

It is well known that in the context of Lebesgue spaces, functions are viewed to be equivalent if the norm of their difference is 0, or equivalently if they differ at most on a set of measure 0. We therefore make the important observation that although we treat (and even refer to) the elements of the Lebesgue spaces as functions, they are in fact equivalence classes of functions. Since we wish each Sobolev space to be a subspaces of a Lebesque space, the same should be true of elements within Sobolev spaces.

Due this fact, the classical derivative will prove unsuitable for use within the definition of Sobolev spaces, as it will impede our ability to equate functions that differ only on sets of measure 0. To illustrate this fact, consider for example the functions

$$u(x) = x \quad \text{for } x \in [0, 1]$$

and

$$\tilde{u}(x) = \begin{cases} x & x \in [0, 0.5) \cup (0.5, 1] \\ 0 & x = 0.5 \end{cases} .$$

Clearly, in the sense of the space $L^1[0, 1]$, these functions are equivalent. This would however not be the case in the sense of a Sobolev space defined via the classical derivative, as while the function u (being classically differentiable) may lie within a certain Sobolev space, $\tilde{u}$, since it is not classically differentiable could not. This indicates that a Sobolev space defined in this way would not be a subspace of a Lebesgue space as required. It therefore follows that we should rather consider a weaker version of differentiability that does not distinguish between functions that differ only on sets of measure 0. To this end, we now turn our attention to the theory of distributional (weak) derivatives.

A.1 The Distributional Derivative

Definition A.6. Let $u \in L^1_{loc}(\Omega)$, we then call $v \in L^1_{loc}(\Omega)$ an *ith* distributional (or weak) derivative of u if

$$\int_\Omega u(x)\phi^{(i)}(x)\mathrm{d}x = (-1)^i \int_\Omega v(x)\phi(x)\mathrm{d}x$$

for every $\phi \in C_0^\infty(\Omega)$.
Notationally, we will denote the *ith* weak derivative of a function $u(x)$ by $D^{(i)}[u(x)]$.

We notice that since the distributional derivative is defined by integrals, functions that differ only on sets of measure 0 will share distributional derivatives. Furthermore, the distributional derivative is not unique. In fact, if a function v satisfies the above definition, then so does every function that differs from v at most on a set of measure 0. As with Lebesgue spaces, when referring to the distributional derivative of a function, we therefore mean the equivalence class of functions differing at most on a set of measure 0.

We now demonstrate that if the classical derivative exists, it coincides with the distributional derivative.

Lemma A.7. *Suppose a function is classically differentiable. The classical derivative is then also a distributional derivative.*

Proof. This result follows immediately due to integration by parts. □

We note that making use of integration by parts, it is also easy to show that if a function is classically differentiable over a certain interval, the distributional derivative agrees with the classical derivative here.

The ability of the distributional derivative to essentially ignore the behaviour of a function on a set of measure 0 allows it to be applied to a wide range of functions that would not normally considered differentiable. We consider the following examples.

Example A.8. Let $\Omega = (-1, 1)$ and $u(x) = |x|$. The function u is clearly not differentiable on Ω in the classical sense, due to the corner at $x = 0$. The function u is however weakly differentiable as the distributional derivative will essentially allow us to ignore the point $x = 0$.

In order to construct a function v that satisfies Definition A.6, we begin by noting that u is classically differentiable on the intervals $(-1, 0)$ and $(0, 1)$ and hence the distributional derivative should coincide with the classical derivative on these intervals. It then only remains to give value to v at the point $x = 0$. This point however is a set of measure 0 and is hence ignored by the distributional derivative. We may thus choose v to have an arbitrary value at this point. Suppose that we set

$$v(x) = \begin{cases} 1 & \text{if } \ x > 0 \\ a & \text{if } \ x = 0 \quad \text{for some } a \in \mathbb{R} \ . \\ -1 & \text{if } \ x < 0 \end{cases}$$

It follows that for any $\phi \in C_0^\infty(\Omega)$ we have that

$$-\int_\Omega v(x)\phi(x)\mathrm{d}x = \int_{-1}^0 1\phi(x)\mathrm{d}x - \int_0^0 a\phi(x)\mathrm{d}x - \int_0^1 1\phi(x)\mathrm{d}x \ ,$$

making use of integration by parts,

$$\begin{aligned} &= -\int_{-1}^0 x\phi'(x)\mathrm{d}x + \int_0^1 x\phi'(x)\mathrm{d}x \\ &= \int_{-1}^1 |x|\phi'(x)\mathrm{d}x \\ &= \int_\Omega u(x)\phi'(x)\mathrm{d}x \ . \end{aligned}$$

It then follows from Definition A.6 that v is a distributional derivative of u.

Example A.9. Let $\Omega = (-1, 1)$ and

$$u(x) = \begin{cases} x & \text{if } \ x \in (-1,0) \cup (0,1) \\ 1 & \text{if } \ x = 0 \end{cases} \tag{A.1}$$

Clearly, in the Lebesgue sense, u is equivalent to the differentiable function $\tilde{u} = x$ and hence they should share distributional derivatives. Lemma A.7 thus implies that we must have $D^{(1)}[u(x)] = 1$.

Example A.10. Let $\Omega = (-1, 1)$ and

$$u(x) = \begin{cases} 1 & \text{if } \ x \geq 0 \\ -1 & \text{if } \ x < 0 \end{cases} \ .$$

The jump discontinuity present in u at $x = 0$ cannot be repaired by changing the function on a set of measure 0 and hence cannot be ignored by the distributional derivative. We therefore assert that u is not weakly differentiable on Ω. To demonstrate this claim, suppose to the contrary that there exists a locally integrable function v such that for all $\phi(x) \in C_0^\infty(\Omega)$

$$\begin{aligned}\int_\Omega v(x)\phi(x)\mathrm{d}x &= -\int_\Omega u(x)\phi'(x)\mathrm{d}x \\ &= \int_{-1}^{0} \phi'(x)\mathrm{d}x - \int_0^1 \phi'(x)\mathrm{d}x \quad .\end{aligned} \tag{A.2}$$

Since u is classically differentiable on the intervals $(-1, 0)$ and $(0, 1)$, the distributional derivative should coincide with the classical derivative on these intervals and hence we should have that

$$v(x) = 0 \text{ for } x \in (-1, 0) \cup (0, 1) \quad .$$

We thus must have that for each $\phi(x) \in C_0^\infty(\Omega)$

$$\int_{-1}^{1} v(x)\phi(x)\mathrm{d}x = 0.$$

Combining this result with equation A.2 we have that for each $\phi(x) \in C_0^\infty(\Omega)$, $\phi(0) = 0$ which can clearly not be true and hence u is not differentiable in the distributional sense.

To conclude this section, we present a result that will prove useful in later work.

Lemma A.11. *(see Driver [17])*
Consider a weakly differentiable function $u \in L^1_{loc}(\Omega)$ with $supp(u) \Subset \Omega$. Then there exists a function $v \in L^1_{loc}(\Omega)$ such that v is a weak derivative of u and $supp(v) \subset supp(u)$.

Proof. From the definition of the weak derivative, for each $\phi \in C_0^\infty(\Omega)$

$$-\int_\Omega D^{(1)}[u(x)]\ \phi(x)\mathrm{d}x = \int_\Omega u(x)\ \phi'(x)\mathrm{d}x$$
$$= \int_{supp(u)} u(x)\ \phi'(x)\mathrm{d}x \quad .$$

It therefore follows that,

$$\int_{\Omega\setminus supp(u)} u(x)\ \phi'(x)\mathrm{d}x = 0$$

and hence, for each $\phi \in C_0^\infty(\Omega \setminus supp(u))$

$$\int_\Omega D^{(1)}[u(x)]\ \phi(x)\mathrm{d}x = 0\ .$$

Thus

$$D^{(1)}[u] = 0 \ \text{ a.e on } \ \Omega \setminus supp(u)\ .$$

Due to the fact that the weak derivative is unique up to sets of measure zero, it follows that there exists a function v that is a weak derivative of u, with $supp(v) \subset supp(u)$. □

A.2 Sobolev Spaces

We begin this section by noting that since the distributional derivative is defined via integration, functions that differ only on sets of measure 0 will share distributional derivatives. Furthermore, the distributional derivative is not unique, in fact if a function v satisfies Definition A.6, then so does every function that differs from v at most on a set of measure 0. As with Lebesgue spaces, when referring to the distributional derivative of a function, we therefore mean the equivalence class of functions differing only on sets of measure 0. It therefore follows that the distributional derivative will preserve the nature of the elements within Lebesgue spaces and is hence suitable for use within the definition of a Sobolev space.

Definition A.12. Let $m \in \mathbb{N}$ and $1 \leq p < \infty$, then the $(m,p)-th$ order Sobolev norm is then defined by

$$\|u\|_{W^{m,p}(\Omega)} = \left(\sum_{i=0}^{m} \|D^{(i)}u\|^p_{L^p(\Omega)} \right)^{1/p} \quad .$$

Definition A.13. Let $m \in \mathbb{N}$, $1 \leq p < \infty$, then making use of the above Sobolev norm, we may define two different types of Sobolev spaces, namely

1. $W^{m,p}(\Omega)$ is defined to be the collection of all functions which the $(m,p)-th$ Sobolev norm is finite.
2. $W_0^{m,p}(\Omega)$ is defined to be the closure of $C_0^\infty(\Omega)$ in the $(m,p)-th$ Sobolev norm.

We note that in the case of a closed set $\bar{\Omega}$, the spaces $C_0^\infty(\bar{\Omega})$ and $L^1_{loc}(\bar{\Omega})$ loose their defining local characteristics and become $C^\infty(\bar{\Omega})$ and $L^1(\bar{\Omega})$ respectively. The use of these spaces within the definition of the distributional derivative is therefore non-nonsensical, as it alters the very nature of the definition. In order to maintain a meaningful definition, we will therefore define a Sobolev space over a closed set $\bar{\Omega}$ to be equivalent to that defined over their open counterparts Ω.

We will now proceed to present a number of key properties of Sobolev spaces. We note that while no proofs are provided in this section, we do direct the reader to appropriate texts where detailed proofs can be found and furthermore note that in general, proofs of the weighted versions of these results are provided within the main body of this dissertation.

We also note that since this dissertation focuses exclusively on functions in $\mathbb{R}$, results are presented in this context for convenience. Most texts however (including those to which we provide reference), present Sobolev theory in the more general setting of $\mathbb{R}^n$.

Theorem A.14. *(See Adams and Fournier [2] p.60-61)*
Let $m \in \mathbb{N}$, $1 \leq p < \infty$, then $W^{m,p}(\Omega)$ is a separable Banach Space. In particular, if $p = 2$, then $W^{m,2}(\Omega)$ is a separable Hilbert space, with inner product given by

$$(u, v)_{W^{m,p}(\Omega)} = \sum_{i=0}^{m} \int_{\Omega} D^{(i)}[u(x)] D^{(1)}[v(x)] \quad .$$

In reference to this fact, many authors write

$$W^{m,2}(\Omega) = H^m(\Omega) \quad .$$

Theorem A.15. *(See Adams and Fournier [2] p.67)*
For any bounded domain $\Omega \subset \mathbb{R}$, the space $C^\infty(\Omega)$ is dense in $W^{m,p}(\Omega)$.

We notice that this result cannot be extended to an unbounded Ω, a fact illustrated in the following example.

Example A.16. Consider the function $u(x) = 1$, defined on $\Omega = \mathbb{R}$. Clearly $u \in C^\infty(\mathbb{R})$, but the integral $\int_{\mathbb{R}} 1 \mathrm{d}x$ does not exist and hence $u \notin W^{m,p}(\mathbb{R})$ for any m and p.

Under certain regularity conditions on the boundary of Ω, we may however extend the above result to ensure that the approximating smooth functions are bounded on the closure of Ω. To this end, we note the following definition due to Adams and Fournier [2].

Definition A.17. A domain $\Omega \subset \mathbb{R}$ is said to satisfy the "segment condition" if for each point $x \in \partial\Omega$ there exists a neighbourhood U_x of x and a number $y_x \neq 0$ such that given $z \in \overline{\Omega} \cap U_x$, we have that for every $0 < \delta < 1$, $z + \delta y_x \in \Omega$.

This condition essentially ensures that Ω does not contain any points that are separated from all other points in the set by some non-zero distance.

Theorem A.18. *(See Adams and Fournier [2] p.68)*
If the domain Ω satisfies the "segment condition" above, then $C_0^\infty(\mathbb{R})$ is dense in $W^{m,p}(\Omega)$ for any $m \in \mathbb{N}$, $1 \leq p < \infty$.

We note that since the restriction to Ω of any function in $C_0^\infty(\mathbb{R})$ clearly lies in $C^\infty(\overline{\Omega})$, this result is equivalent to the density of $C^\infty(\overline{\Omega})$ in $W^{m,p}(\Omega)$.

Furthermore, as a particular case of Theorem A.18 we have.

Theorem A.19. *(See Adams and Fournier [2] p.70)*
For all $m \in \mathbb{N}$ and $1 \leq p < \infty$,

$$W^{m,p}(\mathbb{R}) = W_0^{m,p}(\mathbb{R}) \quad .$$

This result may lead one to ask whether there are other domains Ω for which $W^{m,p}(\Omega) = W_0^{m,p}(\Omega)$? Adams and Fournier [2] provide a detailed answer in Chapter 3, however for our purposes it suffices to know that it cannot happen if the set $\mathbb{R} \setminus \Omega$ has non-zero measure. This is of particular interest as we shall demonstrate that the same is not true for certain weighted Sobolev spaces.

In many applications it is very useful to know under which conditions certain Sobolev spaces may be embedded into other spaces, to this end we will now present the well known Sobolev embedding theorems. We begin by recalling the definition of what it means for one space to be embedded in another.

Definition A.20. Let X, Y be normed spaces, then we say that X is continuously embedded into Y and write

$$X \hookrightarrow Y \quad ,$$

if X is a subspace of Y and there exists a constant $C > 0$ such that for each $u \in X$

$$\|u\|_Y \leq C\|u\|_X \quad .$$

We now recall the following definition.

Definition A.21. (Kreyszig [26] p.405)
Let X, Y be normed spaces, then a linear operator $T : X \to Y$ is said to be compact, if for every bounded subset M of X we have that $\overline{T(M)}$ is compact.

Definition A.22. Let X, Y be normed spaces, we then say that X is compactly embedded into Y and write

$$X \subset\subset Y \quad ,$$

if X is continuously embedded into Y and this embedding is a compact operator.

Similarly to some of the above results, the Sobolev embedding theorems require the boundary of Ω to satisfy certain regularity conditions. These conditions, when viewed within the context of $\mathbb{R}^n$, prove however to be quite complex and thus we direct the reader to the work of Adams and Fournier [2] (page 82 for the "cone condition" and page 83 for the "strong local Lipschitz condition") for a full treatment of these conditions. For the purposes of this dissertation, it will suffice to know that an interval subset of $\mathbb{R}$ satisfies both of these conditions.

Theorem A.23. *(See Adams and Fournier [2] p.168-172)*
Let $\Omega \subset \mathbb{R}$ be a bounded domain that satisfies the strong local Lipschitz condition, then for $2 \leq p < \infty$ and $m \in \mathbb{N}$, the following compact embedding holds

$$W^{m,p}(\Omega) \subset\subset C(\overline{\Omega}) \quad .$$

Theorem A.24. *(See Adams and Fournier [2] p.168-172)*
Let $\Omega \subset \mathbb{R}$ be a bounded domain that satisfies the cone condition, then for $1 \leq p < \infty$, $m, k \in \mathbb{N}$ and $k > m$, the following compact embedding holds

$$W^{k,p}(\Omega) \subset\subset W^{m,p}(\Omega) \quad .$$

A.3 Mollifiers

The proofs of many of the results in the previous section require the construction of sequences of smooth functions that approximate Sobolev functions in various norms. The construction of such sequences is achieved through a process termed mollification, in which the Sobolev function is convolved with a function from a special class of smooth functions termed mollifiers.

A large portion of the main body of this dissertation is dedicated to the derivation of weighted analogs of results in the previous section. Mollifiers will therefore play an important role in this regard and thus we will now provide a brief introduction to the topic.

Definition A.25. (Adams and Fournier [2])
For each $\epsilon > 0$ let J_ϵ be a non-negative function such that

1. $J_\epsilon \in C_0^\infty(-\epsilon, \epsilon)$
2. $\int_{\mathbb{R}} J_\epsilon(x)\mathrm{d}x = 1$

The function J_ϵ is called a mollifier and the convolution (defined for all u for which the below integral exists)

$$J_\epsilon * u(x) = \int_{\mathbb{R}} J_\epsilon\ (x-y)\ u(y)\mathrm{d}y\ ,$$

the mollification of u.

An example of such a function is given by

$$J_\epsilon(x) = \begin{cases} \dfrac{k}{\epsilon}\left(exp\left[\dfrac{-1}{(1+(x/\epsilon)^2)}\right]\right) & |x| < \epsilon \\ 0 & |x| \geq \epsilon\ . \end{cases}$$

The classical properties of mollifers are now given in the following well-known result (see Adams and Fournier [2] p.36).

Theorem A.26. *Let u be a real valued function with $supp(u) \subset \Omega$ then:*

1. *If $u \in L^1_{loc}(\Omega)$, then $J_\epsilon * u \in C^\infty(\Omega)$.*

2. *If $u \in L^1_{loc}(\Omega)$, then $J_\epsilon * u \in C_0^\infty(\Omega)$ provided that*

$$\epsilon < dist(supp(u), \partial\Omega) \ .$$

3. *If $u \in L^p(\Omega)$ for $1 \leq p < \infty$, then $J_\epsilon * u \in L^p(\Omega)$ and*

$$\lim_{\epsilon \to 0^+} \|J_\epsilon * u - u\|_{L^p, \Omega} = 0 \ .$$

4. *If $u \in C(\overline{\Omega})$, then $\lim_{\epsilon \to 0^+} J_\epsilon * u = u$ uniformly on Ω.*

To conclude this section, we will now demonstrate the application of mollifiers and prove a number of very useful properties of the distributional derivative.

Lemma A.27. *Consider a weakly differentiable function $u \in L^1_{loc}(\Omega)$ with $supp(u) \Subset \Omega$. It then follows that*

$$D^{(1)}[J_\epsilon * u(x)] = J_\epsilon * (D^{(1)}[u(x)]) \quad .$$

Proof. We begin by noting that since u is compactly supported within Ω,it must vanish within some neighbourhood of the boundary of Ω and hence Theorem A.26 implies that $J_\epsilon * u \in C_0^\infty(\Omega)$ and is hence classically differentiable. It therefore follows that

$$D^{(1)}[J_\epsilon * u(x)] = \frac{d}{dx}(J_\epsilon * u(x)) = \lim_{h \to 0^+} \left[\frac{1}{h} \int_\Omega u(y)\{J_\epsilon(x - y + h) - J_\epsilon(x - y)\} \mathrm{d}y \right] \quad .$$

The mean value theorem then implies that

$$\begin{aligned} \frac{1}{h} u(y)\{J_\epsilon(x - y + h) - J_\epsilon(x - y)\} &\leq \frac{1}{h}|u(y)|\,|J_\epsilon(x - y + h) - J_\epsilon(x - y)| \\ &= |u(y)| \left| \frac{J_\epsilon(x - y + h) - J_\epsilon(x - y)}{(x - y + h) - (x - y)} \right| \end{aligned}$$

$$= |u(y)||J'_\epsilon(c)|$$
$$\leq |u(y)| \sup_{c\in\mathbb{R}} |J'_\epsilon(c)|,$$

for some $c \in [x-y, x-y+h]$. The right hand side of this inequality is clearly a member of $L^1(\Omega)$ and hence, making use of the Lebesgue Dominated Convergence Theorem and the definition of the weak derivative it follows that

$$\begin{aligned} D^{(1)}[J_\epsilon * u(x)] &= \int_\Omega u(y) \lim_{h\to 0}\{\frac{J_\epsilon(x-y+h)-J_\epsilon(x-y)}{h}\}\mathrm{d}y \\ &= \int_\Omega \frac{\mathrm{d}}{\mathrm{d}x}[J_\epsilon(x-y)]u(y)dy = -\int_\Omega \frac{\mathrm{d}}{\mathrm{d}y}[J_\epsilon(x-y)]u(y)dy \\ &= \int_\Omega J_\epsilon(x-y)\ D^{(1)}[u(y)]dy \\ &= J_\epsilon * D^{(1)}[u(x)] \ \ . \end{aligned}$$

□

Lemma A.28. *(Driver [17])*
Suppose that u is weakly differentiable on the bounded domain $\Omega = (a,b)$, with $D^{(1)}[u] = 0$. It then follows that u is almost everywhere constant on Ω.

Proof. Given $\delta > 0$, we define $\Omega_\delta = [a+\delta, b-\delta]$ and then set

$$\hat{u}(x) = \begin{cases} u(x) & \text{if} \quad x \in \Omega_\delta \\ 0 & \text{otherwise} \end{cases} .$$

Clearly $supp(\hat{u}) \Subset \Omega_\delta$ and $D^{(1)}[\hat{u}] = D^{(1)}[u]$ almost everywhere on Ω_δ. Furthermore, if $0 < \epsilon < \delta$, it follows from Theorem 1.14 that $J_\epsilon * \hat{u} \in C_0^\infty(\Omega)$, while Lemma A.27 implies that

$$\begin{aligned} D^{(1)}[J_\epsilon * \hat{u}] &= J_\epsilon * D^{(1)}[\hat{u}] \\ &= 0 \ \ a.e. \text{ on } \Omega_\delta \ . \end{aligned}$$

It therefore follows that there exists a $C \in \mathbb{R}$ such that

$$J_\epsilon * \hat{u} = C \ \ a.e. \text{ on } \Omega_\delta \ .$$

Recalling that, by Theorem 1.14, $J_\epsilon * \hat{u}$ converges to $\hat{u}$ in $L^1(\Omega_\delta)$, it follows that u is constant almost everywhere on Ω_δ. The result now follows by noting that this is true for every $\delta > 0$ and that

$$\Omega = \bigcup_{\delta>0} \Omega_\delta \ .$$

□

Lemma A.29. *(Driver [17])*
If u is weakly differentiable on the bounded domain $\Omega = (a,b)$, then u has a version that is absolutely continuous on every closed interval $[c,d] \subset (a,b)$.

Proof. Let u be weakly differentiable on $\Omega = (a,b)$ and recall that from the definition of weak differentiability, for every $[c,d] \subset (a,b)$, we have that

$$u \in L^1[c,d] \ \text{ and } \ D^{(1)}[u] \in L^1[c,d] \ .$$

Now, for $x \in [c,d]$, define

$$w(x) = \int_c^x D^{(1)}[u(y)]\mathrm{d}y \ .$$

Clearly w is classically differentiable, with derivative $D^{(1)}[u(x)]$ and hence has a weak derivative that agrees with the classical derivative almost everywhere. That is,

$$D^{(1)}[u(x) - w(x)] = 0 \ \ a.e. \text{ on } [c,d] \ .$$

It therefore follows from Lemma A.28 that $u(x) - w(x)$ is almost everywhere constant on $[c,d]$ and thus, for some $C \in \mathbb{R}$

$$u(x) = w(x) + C \ \ a.e. \text{ on } [c,d] \ .$$

There thus exists a version $\tilde{u}$ of u such that

$$\tilde{u}(x) = w(x) + C \ \text{ for every } x \in [c,d].$$

Setting $x = c$, it follows that

$$\begin{aligned}\tilde{u}(c) &= \int_c^c D^{(1)}[u(y)]\mathrm{d}y + C \\ &= C \ ,\end{aligned}$$

and hence, for $x \in [c, d]$, it follows that

$$\tilde{u}(x) = \tilde{u}(c) + \int_c^x D^{(1)}[u(y)]\mathrm{d}y \quad .$$

Thus, as required,$\tilde{u}$ is absolutely continuous on $[c, d]$. □

Lemma A.30. *Let (v_n) be a sequence of functions that converges uniformly on Ω to some function $v \in L^1(\Omega)$. Now, given a function $u \in L^1(\Omega)$ such that $uv \in L^1(\Omega)$, it follows that*

$$\lim_{n\to\infty} \int_\Omega u(x)v_n(x)\mathrm{d}x = \int_\Omega u(x)v(x)\mathrm{d}x$$

Proof. Since v_n converges to v uniformly on Ω, there exists $N \in \mathbb{N}$ such that for $x \in \Omega$ and $n \geq N$,

$$|v_n(x)| - |v(x)| \leq |v_n(x) - v(x)| < 1 \quad ,$$

or

$$|v_n(x)| < 1 + |v(x)| \quad .$$

We therefore have that for each $n \geq N$,

$$\begin{aligned}|uv_n| &< u(1 + |v|) \\ &\leq u + |uv| \quad .\end{aligned}$$

Noting that the right hand side of this inequality is clearly a member of $L^1(\Omega)$, the result follows due to the Lebesgue dominated convergence theorem. □

Lemma A.31. *(Wong [39])*
Let u and v be weakly differentiable functions on Ω such that both uv and $D^{(1)}[u]v + uD^{[1]}(v)$ are members of $L^1_{loc}(\Omega)$. It then follows that uv is also weakly differentiable on Ω with

$$D^{(1)}[uv] = D^{(1)}[u]v + uD^{(1)}[v] \quad .$$

Proof. We begin by fixing $\phi \in C_0^\infty(\Omega)$ and then note that since $supp(\phi) = \Omega_\phi \Subset \Omega$, the definition of local integrability and Lemma A.29 imply that

$$u, v \in C(\Omega_\phi) \ \text{ and } \ \mathrm{D}^{(1)}[\mathrm{v}] \in \mathrm{L}^1(\Omega_\phi) \quad .$$

Given $\epsilon > 0$, we now set

$$v_\epsilon = \ J_\epsilon * v$$

and note that due to Lemmas 1.14 and A.27 ,

1. $$v_\epsilon \in C^\infty(\Omega^\phi)$$

2. $$v_\epsilon \to v \ \text{ uniformly on } \Omega$$

3. $$v'_\epsilon = J_\epsilon * D^{(1)}[v]$$

4. $$v'_\epsilon \to D^{(1)}[v] \text{ in } L^1(\Omega^\phi) \quad .$$

Finally, we note that since u, v, ϕ, ϕ' are all bounded on Ω_ϕ, their products with integrable functions remain integrable and due to property (4) above

$$uv'_\epsilon\phi \to uD^{(1)}[v] \text{ in } L^1(\Omega_\phi) \quad . \tag{A.3}$$

Making use of the above facts, the result is derived as follows,

$$\int_{\Omega} u(x)v(x)\phi'(x)\mathrm{d}x = \int_{\Omega_\phi} u(x)v(x)\phi'(x)\mathrm{d}x \quad ,$$

making use of Lemma A.30,

$$= \lim_{\epsilon\to 0^+} \int_{\Omega_\phi} u(x)v_\epsilon(x)\phi'(x)\mathrm{d}x \quad ,$$

applying the classical product rule,

$$= \lim_{\epsilon\to 0^+} \int_{\Omega_\phi} u(x)[v_\epsilon(x)\phi(x)]'\mathrm{d}x - \lim_{\epsilon\to 0^+} \int_{\Omega_\phi} u(x)v_\epsilon'(x)\phi(x)\mathrm{d}x \quad ,$$

noting that $v_\epsilon\phi \in C_0^\infty(\Omega)$, the definition of the weak derivative implies,

$$\begin{aligned} &= - \lim_{\epsilon\to 0^+} \int_{\Omega_\phi} D^{(1)}[u(x)]v_\epsilon(x)\phi(x)\mathrm{d}x \\ &\quad - \lim_{\epsilon\to 0^+} \int_{\Omega_\phi} u(x)v_\epsilon'(x)\phi(x)\mathrm{d}x \quad . \end{aligned}$$

Thus, applying Theorem A.30 and (A.3),

$$\begin{aligned} &- \int_{\Omega_\phi} D^{(1)}[u(x)]v(x)\phi(x)\mathrm{d}x - \int_{\Omega_\phi} u(x)D^{(1)}(v(x))\phi(x)\mathrm{d}x \\ &= - \int_{\Omega_\phi} \left[D^{(1)}[u(x)]v(x) - u(x)D^{(1)}[v(x)]\right]\phi(x)\mathrm{d}x \quad . \end{aligned}$$

□

Appendix B

Options and the Black-Scholes Equation

In this appendix we will provide a brief introduction to financial concepts used within the body of the dissertation and will provide a brief introduction to the theory of options and option pricing, with an emphasis on the Black-Scholes option pricing model. For a deeper treatment of these topics, we direct the reader to the well known works of Hull [24] or Bjork [6].

B.1 Options

An option is a contract that affords its holder the opportunity to buy or sell a given financial asset for a predetermined price at some time in the future. We call the asset that is to be bought or sold the underlying (whose value at some time $t > 0$. is denoted by $x(t)$) and the predetermined price and future time, the strike price (denoted by $K > 0$) and expiration date (denoted by $T > 0$) respectively. Options that allow the holder to buy the underlying are termed call options, while options to sell the underlying are referred to as put options. When the holder of an option makes use of their right to buy or sell the underlying asset, we say that they have exercised the option. Options may be classified as being of either European or American type based on when the exercise may occur. European options allow the holder to exercise the option only at the expiry date, whereas American options allow the holder more freedom in that they may be exercised at any time prior to, or at the expiration date. There are many choices for the financial asset on which an option is written, with some of the more common examples being stocks, foreign currencies and futures, however for the purposes of this dissertation we will focus on the case of stock options of European type.

We now recall that options offer their holders a right to exercise, but do not oblige them to do so. It therefore follows that the holder will only exercise the option if it is profitable to do so. Consider for example a European call option. At expiry, the holder has the opportunity to pay the strike price K and receive a share of the underlying. Clearly this is only advantageous if the current value of the underlying $x(T)$ exceeds the strike price K, in which case the holder receives a profit equal to $x(T) - K$. If the holder does not exercise the option, it expires with a value of 0. It therefore follows that at expiry, a European call option will have the value

$$[x(T) - K]^+ = \max\{x(T) - K, 0\} \quad ,$$

which we will term the payoff of the option.
A similar argument holds for the case of a European put option; at expiry the holder has the right to sell a share of the underlying for the strike price K. Once again, the holder of the option will only make this sale if it is advantageous to do so and hence the payoff of the option is given by

$$[K - x(T)]^+ = \max\{K - x(T), 0\} \quad .$$

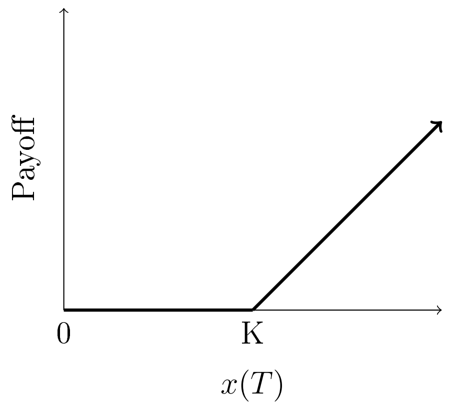

Figure B.1: The payoff function $[x(T) - K]^+$ of a call option.

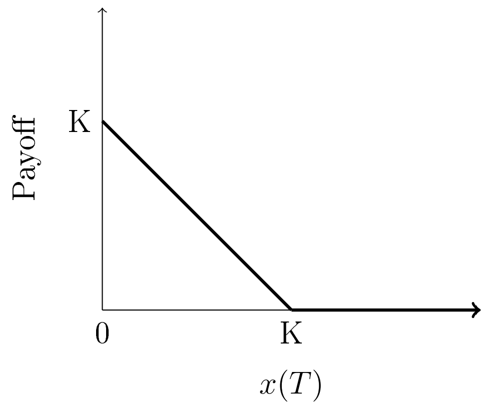

Figure B.2: The payoff function $[x(T) - K]^+$ of a put option.

The standard put and call options we have described thus far are termed vanilla options and are the most basic within the class. More advanced options, or so called exotic options, may be constructed by the addition of further conditions. Such conditions may for example alter the underlying, introduce some sort of path dependence or change the terms under which the option may be exercised. For examples of such exotics we direct the reader to Hull [24] (Chapter 24).

In this dissertation we will focus our attention on the class of path-dependent exotic options referred to as barrier options. These options extend the idea of vanilla puts and calls by placing a restriction on the values that the underlying may assume over the life of the option. These restrictions are classified as either knock-in or knock-out conditions. A knock-in condition asserts that the option will expire worthless unless the underlying asset attains a certain value (termed the knock-in barrier) over the course of the options life. A knock-out condition on the other hand causes the option to expire worthless if the underlying attains a certain value (termed the knock-out barrier) over the course of the options life. Barrier options may further be classified as up, down or double barrier options. If the barrier condition is triggered by the value of the underlying exceeding some value $U > 0$, the option is termed an up-type option, while a down-type barrier option is one in which the barrier condition is triggered by the value of the underlying falling below a certain value L. As with vanilla options, it is clear that the value of a barrier option is known at the expiry date. An up and out European call option for example has a payoff of the form

$$\begin{cases} [x(T) - K]^+ & \text{provided } \; x(t) < U \; \forall \; t \in [0, T] \\ 0 & \text{otherwise} \end{cases},$$

while the payoff of a down and in European put option is given by

$$\begin{cases} [K - x(T)]^+ & \text{provided } \; x(t) < L \text{ for some } t \in [0, T] \\ 0 & \text{otherwise} \end{cases}.$$

We call the option that combines the conditions of up and down type barriers, a double barrier option.

In the theory of option pricing, barriers are usually assumed to be monitored in continuous time. In other words, the knock-in and knock-out conditions are applied if the underlying reaches the barrier at any time during the options life. This is however an unrealistic assumption in practice as it is impossible to monitor the value of the underlying at every point in time. In fact, the majority of traded barrier options specify only a fixed number of times $0 \leq t_1 < t_2 < ... < t_N \leq T$ at which the value of the underlying is checked against the barrier and allow the underlying to assume any value during the interim time periods. As before, the payoff of such an option is known. For example, a double barrier knock-out call option with monitoring dates $0 \leq t_1 < t_2 < ... < t_N \leq T$ has a payoff of

$$\begin{cases} [x(T) - K]^+ & \text{provided that } \ L < x(t_n) < U \text{ for every } n = 1, 2, ...N \\ 0 & \text{otherwise} \end{cases} . \tag{B.1}$$

A special case of discretely monitored barrier options is the so called binary option, which may be viewed as a barrier option with a single monitoring time that coincides with the expiry and pays out 1 if the option is not knocked out and 0 otherwise.

B.2 Option Pricing and the Black Scholes Equation

As we can see from the above discussion, the value of an option is known at the time it is exercised/expires. Since this value cannot be negative (the holder of the option is under no obligation to exercise the option) and may be positive (the holder has a chance of making a profit), it is clear that the option must have a positive value at inception. We term this value the fair price of the option and it is the amount an investor must pay in order to enter into the contract. At the time the option is written, we cannot know the value of the underlying at the expiry date (or hence the payoff of the option) and thus the problem of calculating the price of an option is stochastic in nature.

In their seminal paper, Black and Scholes [7] show that under certain assumptions the value of European options satisfy a deterministic partial differential equation of parabolic type. This model (termed the Black-Scholes model) has become one of the cornerstones of quantitative finance and will form a basis for this dissertation.

Before we may proceed to introduce the workings of the model, we recall the following basic stochastic processes.

Definition B.1. (Bjork [6])
A stochastic process $(W_t)_{t\geq 0}$ is called a Wiener process (or standard Brownian motion), provided that it satisfies

1. $W_0 = 0$.
2. The paths $t \to W_t$ are almost surely continuous.
3. Given times $0 < s < t$, the random variable $W_t - W_s$ is normally distributed with mean 0 and variance $\sqrt{t-s}$.
4. The process (W_t) has independent increments. That is, given $0 < r < s < t$, $W_s - W_r$ and $W_t - W_s$ are independent.

Definition B.2. Let x_t be the solution to the stochastic differential equation

$$dx_t = \mu x_t dt + \sigma x_t dW_t \quad t > 0$$

$$x_0 = X_0 \quad .$$

Then x_t is called geometric Brownian motion with drift parameter μ and volatility σ.

Making use of these processes, the model of Black and Scholes begins with the following assumptions.

1. The risk-free interest rate r is known and constant.

2. The value of the underlying is geometric Brownian motion, i.e. for $t \geq 0$

$$dx_t = rxdt + \sigma x dW_t \tag{B.2}$$

   where W_t is a Wiener process and r and σ are the risk-free interest rate and volatility of the underlying respectively.

3. The underlying does not pay dividends (in the case of the underlying being a share of stock).

4. There are no transaction costs when buying or selling either the underlying or the option.

5. There are no restrictions to short selling.

6. Financial assets are divisible, in other words, an investor may purchase any fraction of an asset.

7. The market is free of arbitrage.

If these conditions are satisfied, then for all times $0 \leq t < T$, the value of a European option (denoted by $V(x,t)$) satisfies the Black-Scholes partial differential equation.

$$\frac{\partial V}{\partial t} + rx\frac{\partial V}{\partial x} + \frac{\sigma^2}{2}x^2\frac{\partial^2 V}{\partial x^2} - rV = 0 \tag{B.3}$$

With terminal condition

$$V(x,T) = f(x) \quad .$$

Under this framework, one may derive closed form solutions for the values of a number of common European options, including standard puts and calls (see Black and Scholes [7]), as well as continuously monitored barrier options (see Merton [30]). More complicated exotic options (including discretely monitored barrier options) do not in general have closed form valuation formulae and as such it is often necessary to resort to numerical valuation methodologies.

Numerically, options are most commonly valued by binomial (see Cox, Ross and Rubinstein [15]) and trinomial (see Boyle [9]) lattice methods. These methods rely on constructing a mesh of possible paths for the value of the underlying, by discretising the time horizon and then assuming that at each discrete point in time the value of the underlying may either make a fixed up or downward movement (or remain unchanged, in the case of the trinomial method). Once the mesh has been constructed, the payoff of the option is calculated at each of the terminal points in the mesh and then these values are discounted back through the tree to obtain a day one price for the option. These methods are particularly useful for the valuation of American style options, as they allow for the inclusion of early exercise features. Lattice methods are however generally unable to accurately value path dependant options (such as Asian or Lookback options) as the lattice is unable to distinguish the path that the underlying took to reach a certain node. Furthermore, it is well documented with in literature (see Boyle and Lau [10], as well as Ritchken [32]) that the application of lattice methods to barrier options may result in large errors and very slow convergence, unless great care is taken to ensure that the barrier is well positioned relative to the mesh.

Another common numerical approach to option valuation is Monte Carlo simulation(see Boyle, Broadie and Glasserman [8]). Monte Carlo simulation relies on the assumption that the underlying follows geometric Brownian motion, as is assumed under the Black-Scholes framework. Making use of this assumption, the method the simulates sample paths for the underlying and then calculates the value of the option under each simulation. These values are then averaged to obtain the price for the option. Since this methods simulates paths for the underlying individually, it is well suited to the valuation of path dependant options. Monte Carlo simulation does however not have the flexibility to consider options with early exercise possibilities and often requires a large number of simulations to obtain an accurate value .

Finally, since the Black-Scholes framework phrases the valuation problem in terms of a PDE, the well known PDE methods of finite difference and finite elements (see Seydel [34]) may also be applied within option pricing. These methods do however present a number of complications due to the nature of the Black-Scholes PDE. Neither method is well equipped to accept degenerate problems on unbounded domains and as such, many authors elect to transform the Black-Scholes PDE to the more common heat equation and truncate the spacial domain when applying either of these methods. Despite these complexities, the finite element method in particular allows a large degree of flexibility and as such is often applied to value more complex contracts that cannot be accurately valued under more standard methods.

Bibliography

[1] Yves Achdou and Olivier Pironneau. *Computation Methods for Option Pricing*. SIAM, 2005.

[2] Robert A. Adams and John J. F. Fournier. *Sobolev spaces*, volume 140 of *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, second edition, 2003.

[3] Dong-Hyun Ahn, Stephen Figlewski, and Bin Gao. Pricing discrete barrier options with an adaptive mesh model. *The Journal of Derivatives*, 6(4):33–43, 1999.

[4] Ivo Babuška. The finite element method for infinite domains. i. *Mathematics of Computation*, 26(117):1–11, 1972.

[5] P. Bettess. *Infinite Elements*. Penshaw Press, 1992.

[6] Tomas Björk. *Arbitrage Theory in Continuous Time*. Oxford University Press, 1998.

[7] Fischer Black and Myron Scholes. The pricing of options and corporate liabilities. *The journal of political economy*, pages 637–654, 1973.

[8] Phelim Boyle, Mark Broadie, and Paul Glasserman. Monte carlo methods for security pricing. *Journal of economic dynamics and control*, 21(8):1267–1321, 1997.

[9] Phelim P. Boyle. Option valuation using a three-jump process. *International Options Journal*, 3(7-12):1, 1986.

[10] Phelim P. Boyle and Sok Hoon Lau. Bumping up against the barrier with the binomial method. *The Journal of Derivatives*, 1(4):6–14, 1994.

[11] H. Brezis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Universitext (Berlin. Print). Springer, 2010.

[12] Mark Broadie, Paul Glasserman, and Steven Kou. A continuity correction for discrete barrier options. *Mathematical Finance*, 7(4):325–348, 1997.

[13] P.G. Ciarlet. *The Finite Element Method for Elliptic Problems*. Studies in Mathematics and its Applications. Elsevier Science, 1978.

[14] J.B. Conway. *A Course in Functional Analysis*. Graduate Texts in Mathematics. Springer, 1990.

[15] John C. Cox, Stephen A. Ross, and Mark Rubinstein. Option pricing: A simplified approach. *Journal of financial Economics*, 7(3):229–263, 1979.

[16] Gearoid De Barra. *Introduction to measure theory*. Van Nostrand Reinhold, 1974.

[17] Bruce K. Driver. *Analysis Tools with Applications*. Lecture Notes, 2003. URL: `http://www.math.ucsd.edu/~bdriver/240-01-02/Lecture_Notes/anal2p.pdf`.

[18] A. Ern and J.L. Guermond. *Theory and Practice of Finite Elements*. Number v. 159 in Applied Mathematical Sciences. Springer, 2004.

[19] Lawrence C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.

[20] P.A. Forsyth, K.R. Vetzal, and R. Zvan. A finite element approach to the pricing of discrete lookbacks with stochastic volatility. *Applied Mathematical Finance*, 6(2):87–106, 1999.

[21] Donald A. French. The finite element method for a degenerate elliptic equation. *SIAM J. Numer. Anal.*, 24(4):788–815, August 1987.

[22] Gianluca Fusai, David Abrahams, and Carlo Sgarra. An exact analytical solution for discrete barrier options. *Finance and Stochastics*, 10(1):1–26, 2006.

[23] Pierre Grisvard. *Elliptic Problems in Nonsmooth Domains*. Pitman Advanced Publishing Program, 1985.

[24] J. Hull. *Options, Futures and Other Derivatives*. Options, Futures and Other Derivatives. Pearson/Prentice Hall, 2009.

[25] Kapil D. Joshi. *Introduction to general topology*. New Age International, 1983.

[26] Erwin Kreyszig. *Introductory functional analysis with applications*. Wiley Classics Library. John Wiley & Sons, Inc., New York, 1989.

[27] Alois Kufner. *Weighted Sobolev spaces*, volume 31 of *Teubner-Texte zur Mathematik [Teubner Texts in Mathematics]*. BSB B. G. Teubner Verlagsgesellschaft, Leipzig, 1980. With German, French and Russian summaries.

[28] Alois Kufner and Bohumír Opic. How to define reasonably weighted Sobolev spaces. *Comment. Math. Univ. Carolin.*, 25(3):537–554, 1984.

[29] A. D. Lyashko and M.R. Timerbaev. Accuracy estimates for the schemes of the finite element method for second order degenerate elliptic equations. *Differentsialńye Uravmeniya (Russian)*, 29:1210–1215, 1993.

[30] Robert C. Merton. Theory of rational option pricing. *Bell Journal of Economics*, 4(1):141–183, 1973.

[31] John T. Oden and Junuthula N. Reddy. *An introduction to the mathematical theory of finite elements*. Wiley, 1976.

[32] Peter Ritchken. On pricing barrier options. *The Journal of Derivatives*, 3(2):19–28, 1995.

[33] Simona Sanfelici. Galerkin infinite element approximation for pricing barrier options and options with discontinuous payoff. *Decis. Econ. Finance*, 27(2):125–151, 2004.

[34] Rüdiger Seydel. *Tools for computational finance*, volume 4. Springer-Verlag, Berlin, 2002.

[35] M. R. Timerbaev. Embedding theorems for weighted Sobolev spaces. *Izv. Vyssh. Uchebn. Zaved. Mat.*, (9):56–60, 1991. URL: http://www.mathnet.ru/php/archive.phtml?wshow=paper&jrnid=ivm&paperid=5151&option_lang=eng.

[36] M.R. Timerbaev. Error estimates for n-dimensional spline interpolation in weighted norms. *Izvestiya Vysshikh Uchebnykh Zavedenii. Matematika*, (10):54–60, 1992. URL: http://www.mathnet.ru/php/archive.phtml?wshow=paper&jrnid=ivm&paperid=8461&option_lang=eng.

[37] Jürgen Topper. Finite element modeling of exotic options. In *Operations Research Proceedings*, pages 336–341. Springer, 2000.

[38] Gunter Winkler, Thomas Apel, and Uwe Wystup. Valuation of options in hestons stochastic volatility model using finite element methods. *Foreign Exchange Risk*, pages 283–303, 2001.

[39] Willie Wong. Lecture notes on sobolev spaces for cambridge centre for analysis. *Lecture Notes*. URL: https://documents.epfl.ch/users/w/ww/wwywong/www/lecturenotes/NotesOnSobolevSpaces.pdf.

[40] E. Zeidler. *Nonlinear Functional Analysis and Its Applications: II/ A: Linear Monotone Operators*. Springer New York, 1990.

[41] Zili Zhu and Nick Stokes. A finite element platform for pricing path dependent exotic options. Technical report, 1998.

[42] R. Zvan, P.A. Forsyth, and KR Vetzal. A general finite element approach for pde option pricing models. *Working Paper*.