

ABSTRACT

Title of dissertation: Adaptive Finite Element Methods
For Variational Inequalities:
Theory And Applications In Finance

Chen-Song Zhang
Doctor of Philosophy, 2007

Dissertation directed by: Professor Ricardo H. Nochetto
Department of Mathematics

We consider variational inequalities (VIs) in a bounded open domain $\Omega \subset \mathbb{R}^d$ with a piecewise smooth obstacle constraint. To solve VIs, we formulate a fully-discrete adaptive algorithm by using the backward Euler method for time discretization and the continuous piecewise linear finite element method for space discretization. The outline of this thesis is the following.

Firstly, we introduce the elliptic and parabolic variational inequalities in Hilbert spaces and briefly review general existence and uniqueness results (Chapter 1). Then we focus on a simple but important example of VI, namely the obstacle problem (Chapter 2). One interesting application of the obstacle problem is the American-type option pricing problem in finance. We review the classical model as well as some recent advances in option pricing (Chapter 3). These models result in VIs with integro-differential operators.

Secondly, we introduce two classical numerical methods in scientific computing: the finite element method for elliptic partial differential equations (PDEs) and the

Euler method for ordinary differential equations (ODEs). Then we combine these two methods to formulate a fully-discrete numerical scheme for VIs (Chapter 4). With mild regularity assumptions, we prove optimal a priori convergence rate with respect to regularity of the solution for the proposed numerical method (Chapter 5).

Thirdly, we derive an a posteriori error estimator and show its reliability and efficiency. The error estimator is localized in the sense that the size of the elliptic residual is only relevant in the approximate noncontact region, and the approximability of the obstacle is only relevant in the approximate contact region (Chapter 6). Based on this new a posteriori error estimator, we design a time-space adaptive algorithm and multigrid solvers for the resulting discrete problems (Chapter 7).

In the end, numerical results for $d = 1, 2$ show that the error estimator decays with the same rate as the actual error when the space meshsize and the time step tend to zero. Also, the error indicators capture the correct local behavior of the errors in both the contact and noncontact regions (Chapter 8).

Adaptive Finite Element Methods
for Variational Inequalities:
Theory and Applications in Finance

by

Chen-Song Zhang

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2007

Advisory Committee:

Professor Nochetto, Ricardo (Chair/Advisor)
Professor von Petersdorff, Tobias
Professor Fu, Michael
Professor Machedon, Matei
Professor Madan, Dilip

© Copyright by
Chen-Song Zhang
2007

DEDICATION

To my parents, Jintang and Xiaoping.

ACKNOWLEDGMENTS

I owe my gratitude to all the people who have made this thesis possible and because of whom my graduate experience has been one that I will cherish forever.

I would like to thank my advisor Ricardo H. Nochetto for providing the vision and support for this work throughout my years at Maryland. His support includes not only mathematical guides to my research, but also providing adequate funds for conducting my research. I am grateful for his kindness, support, and encouragement, as well as his ability to challenge me to strive to be a better mathematician.

I would also like to thank my collaborators and friends. In particular, I am grateful to Tobias von Petersdorff for his insightful discussion and collaboration. I am also thankful for postdoctoral fellows and visitors, J. Manuel Cascon, Long Chen, and Kyoung-Sook Moon who helped me in various stages of this study. I am grateful to Professor Christopher Schwab for valuable discussion. I thank Alfred Schmidt and Kunibert G. Siebert. They share their adaptive finite element package, ALBERTA, with the community and gave me great help when I started to learn this package.

I would like to thank Prof. Michael Fu, Prof. Matei Machedon, and Prof. Dilip Madan for serving on my dissertation committee and sparing their invaluable time to review my thesis and provide feedback.

Finally, I have been continuously blessed by friends who have made my time in

graduate school memorable: Gunay Dogan, Khamron Mekchay, Xia Wang, Guanhua Lu and Qing Xia. Without these friendships and moral support, this work would not have been possible.

TABLE OF CONTENTS

1	Variational Inequalities: Existence and Regularity	4
1.1	Abstract Setting	4
1.2	Elliptic Variational Inequalities (EVIs)	6
1.2.1	Variational Formulation	6
1.2.2	Existence and Uniqueness	7
1.3	Parabolic Variational Inequalities (PVI)	9
1.3.1	Weak Formulation	9
1.3.2	Existence and Uniqueness	11
2	Obstacle Problems	12
2.1	Function Spaces	12
2.1.1	Hölder Spaces	13
2.1.2	Sobolev Spaces of Integer Order	13
2.1.3	Sobolev Spaces of Fractional Order	14
2.1.4	Norm Equivalence	16
2.2	Angle-Bounded Operators	17
2.2.1	Angle Condition	18
2.2.2	Coercivity Property	19
2.3	Obstacle Problems	20
2.3.1	Elliptic Obstacle Problems	20
2.3.2	Equivalent Formulations	21
2.3.3	Parabolic Obstacle Problems	24
2.3.4	Lagrange Multiplier	26
3	Option Pricing – An Application in Finance	27
3.1	Option Contract	27
3.2	Black-Scholes Model	28
3.2.1	A Simple Example: American Put Option	29
3.2.2	Black-Scholes Inequality	31
3.3	Beyond Black-Scholes Model	33
3.3.1	Lévy Processes	33
3.3.2	Lévy-Khintchine Formula	35
3.4	Option Pricing as a Variational Inequality	37

4	Numerical Methods for Obstacle Problems	41
4.1	Finite Element Methods	41
4.2	Euler Method for ODEs	46
4.3	Numerical Methods for Parabolic VI	48
4.3.1	Continuous Problem	48
4.3.2	Semi-discrete Problem	49
4.3.3	Fully-discrete Problem	51
5	A Priori Error Estimation	54
5.1	A Priori Error Estimation for EVIs	54
5.1.1	Abstract Error Estimation	55
5.1.2	Application to Stationary Obstacle Problems	56
5.2	A Priori Error Estimation for PVI's	57
5.2.1	Introduction	57
5.2.2	Estimation of Space Error	58
5.2.3	Positivity Preserving Operators	61
5.2.4	Optimal Convergence Rate	63
6	A Posteriori Error Estimation	65
6.1	Introduction	67
6.2	Stationary Problems	68
6.2.1	Lagrange Multiplier	69
6.2.2	Abstract Error Bounds	69
6.3	Approximation of Lagrange Multipliers	71
6.3.1	Discrete Contact and Noncontact Sets	72
6.3.2	Discrete Lagrange Multiplier	73
6.4	Residual-type Error Estimation	75
6.4.1	Upper Bound	75
6.4.2	Lower Bound	83
6.5	Time-dependent Problems	93
6.5.1	Lagrange Multiplier and Galerkin Functional	94
6.5.2	Abstract Error Bounds	95
6.5.3	Localized Error Estimators	97
6.6	General Obstacle	99
6.6.1	A Magic Bullet?	99
6.6.2	Obstacle Consistency Error	100
6.7	Mesh Changes and Coarsening Error	105
6.7.1	Transfer Operator	105
6.7.2	Residual and Galerkin Functional for Mesh Changes	107
6.7.3	Coarsening Error Estimate	108
6.7.4	Final A Posteriori Upper Bound	110

7	Adaptive and Multilevel Algorithms	114
7.1	Introduction	114
7.1.1	Adaptive Algorithm for Static Problem	115
7.1.2	Adaptive Algorithm for Evolution Problems	115
7.1.3	Convergence and Optimality	117
7.2	Estimate	118
7.3	Mark	120
7.3.1	Maximum Strategy	120
7.3.2	Equidistribution Strategy	121
7.3.3	Dörfler's Marking Strategy	121
7.4	Refine/Coarsen	121
7.4.1	Newest Vertex Bisection in 2d	122
7.4.2	Coarsening Algorithm	124
7.4.3	Compatible Bisection	124
7.4.4	Bisection Grids Revisited	125
7.5	Solve	127
7.5.1	Subspace Correction Methods for Obstacle Problems	128
7.5.2	Convergence Rate of SSC-CDM Methods	130
7.5.3	SSC-CDM on Adaptive Grids	135
8	Numerical Experiments	142
8.1	Asymptotic convergence rates (Part I: Differential Problems)	143
8.1.1	1d Tent Obstacle: Case $\chi_h = \chi$	143
8.1.2	1d Tent Obstacle: Case $\chi_h \neq \chi$	148
8.1.3	1d American Option	149
8.1.4	1d American Option with Moving Obstacle	151
8.1.5	2d Oscillating Moving Circle	152
8.2	Asymptotic convergence rates (Part II: Integral Problems)	155
8.2.1	Elliptic Equations	155
8.2.2	Elliptic Variational Inequalities	159
8.2.3	Parabolic Variational Inequalities	161
8.3	Adaptivity	163
8.3.1	1d American Option	163
8.3.2	1d Tent Obstacle	166
8.3.3	2d Tent Obstacle	168
8.4	Convergence of Discrete Solver	168
8.4.1	Smooth Constraint	168
8.4.2	Inactive Constraint	171
8.4.3	Kink Constraint	172
8.4.4	Singular Constraint	172
8.4.5	Unstable Constraint	174
8.5	Conclusions	174

LIST OF TABLES

8.1	1d tent obstacle problem (conforming case): asymptotic convergence rates	145
8.2	Asymptotic convergence rates: space error estimators	145
8.3	1d tent obstacle problem: approximation of the free boundary	147
8.4	1d tent obstacle problem (non-conforming case on uniform meshes): asymptotic convergence rates	148
8.5	1d tent obstacle problem (non-conforming case on graded meshes): asymptotic convergence rates	149
8.6	1d American put option problem: asymptotic convergence rates using uniform time partition	150
8.7	1d American put option problem: asymptotic convergence rates using graded time partition	151
8.8	1d American put option problem with moving obstacle: asymptotic convergence rates	152
8.9	2d oscillating moving circle problem: asymptotic convergence rates . .	153
8.10	2d oscillating moving circle problem: approximation of the free boundary	154
8.11	Elliptic integral equation: asymptotic convergence rates	156
8.12	Elliptic integral equation: asymptotic convergence rates of the oscillation term	157
8.13	Elliptic integral equation: asymptotic convergence rates for different Y	158
8.14	Elliptic integro-differential equation: asymptotic convergence rates . .	159
8.15	SSC-CDM convergence rate: inactive constraint.	173

8.16 SSC-CDM convergence rate: kink constraint	173
8.17 SSC-CDM convergence rate: singular constraint	173
8.18 SSC-CDM convergence rate: unstable constraint.	175

LIST OF FIGURES

3.1	American put options: payoff and exercise boundary	31
3.2	Foreign exchange rate: US dollars per Euro.	34
3.3	Foreign exchange rate: Yen per US dollars.	34
4.1	Conforming partition of Ω	43
6.1	Local Patch	72
6.2	Finite element function and its derivatives	80
6.3	Admissible region of p	81
6.4	Singular residual	88
6.5	Localization effect	100
6.6	Obstacle consistency error	102
6.7	Transfer operator	106
7.1	Flowchart of adaptive algorithm for static problems	116
7.2	Regular refinement	122
7.3	Newest vertex bisection	123
7.4	Compatible bisection	125
7.5	Patches generated by compatible bisections	126
7.6	Decomposition of a bisection grid	126
7.7	Decomposition of u	139

8.1	Solution of 1d tent obstacle problem	144
8.2	Asymptotic convergence rates of error and error estimator	146
8.3	1d tent obstacle problem: localization behavior of the space error estimator	147
8.4	2d oscillating moving circle problem: localization behavior of the space error estimator	154
8.5	Elliptic integral equation: asymptotic convergence rates and local behavior of the error estimator	158
8.6	Elliptic integro-differential equation: asymptotic convergence rates and local behavior of the error estimator	160
8.7	Elliptic variational inequality: asymptotic convergence rates and local behavior of the error estimator on uniform meshes	161
8.8	Elliptic variational inequality: asymptotic convergence rates and local behavior of the error estimator on graded meshes	162
8.9	Parabolic variational inequality: local behavior of space error estimator	162
8.10	Parabolic variational inequality: asymptotic convergence rates of time and space error estimators	163
8.11	Jump of the free boundary point	164
8.12	Time step sizes using different transfer operators	165
8.13	1d American option pricing problem: comparison of asymptotic convergence rates	166
8.14	1d tent obstacle: comparison of asymptotic convergence rates	167
8.15	2d tent obstacle: adaptive mesh and solution	169
8.16	2d tent obstacle: comparison of asymptotic convergence rates	170
8.17	Convergence rate of multilevel solver SSC-CDM on a graded mesh with $h_{\min} = 8.839 \times 10^{-2}$	171
8.18	Convergence rate of multilevel solver SSC-CDM on a graded mesh with $h_{\min} = 6.250 \times 10^{-2}$	172
8.19	Convergence rate of multilevel solver SSC-CDM on a graded mesh with $h_{\min} = 4.419 \times 10^{-2}$	174

8.20	Convergence rate of multilevel solver SSC-CDM on a graded mesh with $h_{\min} = 3.125 \times 10^{-2}$	175
8.21	The dependence of reduction rate on minimal meshsize	176

Introduction

The subject of variational inequalities has its origin in the calculus of variations associated with the minimization of infinite-dimensional energy functionals. Variational inequalities (VIs) arise from a wide range of application areas, like mechanics, control theory, and finance (see, for example, [120, 73, 142]). After decades of development, this subject has become very rich on both theory and numerics, especially for the special class of obstacle problems. For a general discussion on the existence and regularity, we refer the interested readers to [31, 84]. For numerical methods, we refer to [74] for a quick review.

One of the applications of variational inequalities in finance is valuation of American options. Options are derivative contracts where the future payoffs to the buyer and seller of the contract are determined by the price of another security, such as a common stock. Since the option gives the buyer a right and the writer an obligation, the buyer pays the option premium to the writer. Models of option pricing were very simple and incomplete until 1973 when F. Black and M. Scholes published the Black-Scholes pricing model [23]. For an American-style option in the Black-Scholes model, the arbitrage-free price of the option follows a parabolic variational inequality with a diffusion operator. Some more advanced models, like the CGMY model [40] gives integro-differential variational inequalities.

To solve parabolic variational inequalities numerically, we use the implicit Euler method for time-discretization and the finite element method (FEM) for space-discretization with adaptive mesh refinement techniques. Adaptive mesh refinement is an important tool to deal with multiscale phenomena and to reduce the size of the linear systems that arise from the finite element method. Generally speaking,

an adaptive FEM for static problems consists of iterations in the form

SOLVE \rightarrow ESTIMATE \rightarrow MARK \rightarrow REFINE/COARSEN.

The procedure ESTIMATE determines the error distribution within the domain. Since we cannot compute the exact error of the solution, finding a reliable and efficient a posteriori error estimator, which are locally computable, is critical aspect of adaptivity. This is the main subject of this thesis for a class of variational inequalities governed by integro-differential operators.

The outline and main contributions of the thesis are as follows.

- In Chapter 1, we introduce the general formulation of variational inequalities in Hilbert spaces and review the existence and uniqueness of the solutions of variational inequalities. In Chapter 2, we recall the regularity results of the obstacle problem, which is a special case of variational inequalities.
- In Chapter 3, we introduce option pricing models very briefly. The classical Black-Scholes model and more advanced models based on Lévy processes lead to parabolic variational inequalities with an integro-differential operator \mathcal{A} . For solving the class of parabolic variational inequalities proposed in Chapter 3, we use the implicit Euler method for time discretization and the finite element method for space discretization to formulate a fully-discrete numerical scheme in Chapter 4.
- In Chapter 5, we review the approximation properties of finite element methods for elliptic variational inequalities and generalize them to time-dependent problems. On the basis of recent advances in error estimation for the implicit Euler scheme, we prove optimal convergence rates of our numerical scheme in both space and time with respect to the regularity of solutions. This does not only generalize but also improve the a priori error estimation for the Laplacian by Johnson [82], in that convergence rate is now optimal.
- In Chapter 6, we consider the a posteriori error analysis for parabolic variational inequalities upon extending an idea of Fierro and Veiser for differential

operators [72]. First, we analyze elliptic variational inequalities to give a localized a posteriori error estimator for the error in energy norm for integro-differential operators. And then we generalize the analysis to time-dependent problems using the fact that the underlying operator \mathcal{A} is strongly sectorial. Finally, we also address the error due to lack of obstacle conformity as well as the effect of mesh changes in some detail. The discussion in this chapter is mainly based on two research papers: the one on a posteriori error estimation for parabolic variational inequalities with a general second order differential operator by Moon et al [104]; the other one on a posteriori error estimation for variational inequalities with an integro-differential operator by Nochetto et al [115].

- In Chapter 7, we design a fully-adaptive algorithm for solving parabolic variational inequalities. In particular, for solving discrete variational inequality problems, we generalize the multilevel constraint decomposition method by Tai [130] to graded meshes obtained by the bisection method. We prove that the convergence rate is globally linear and the reduction rate depends on the minimal meshsize mildly in 1d and 2d; the dependence is logarithmic.
- In Chapter 8, we perform several numerical experiments in 1d and 2d. The experiments confirm our theoretical expectations and show advantages of adaptive algorithms for American option pricing for both differential and integral operators alike.

Chapter 1

Variational Inequalities: Existence and Regularity

As a starting point, we introduce the concept of general variational inequalities and review the existence and uniqueness theorem. The rest of the chapter is organized as follows. In Section 1.1, we introduce the definition the elliptic variational inequality problem. In Section 1.2, we review the Lions-Stampacchia existence theorem in Hilbert space. In Section 1.3, we first introduce the notion of parabolic variational inequality problem as well as its weak formulation; then we review existence and uniqueness results.

1.1 Abstract Setting

We begin with notation and basic concepts. The symbol $a \lesssim b$ means $a \leq Cb$ with a generic constant C (not necessarily the same in any two places) and $a \approx b$ abbreviates $a \lesssim b \lesssim a$.

Let \mathcal{H} be a real Hilbert space associated with inner product (\cdot, \cdot) , which in turn induces a norm on \mathcal{H} via

$$\|w\|_{\mathcal{H}} := (w, w)^{\frac{1}{2}} \quad \forall w \in \mathcal{H}. \quad (1.1)$$

When there is no confusion, we use $\|\cdot\|$ to denote the \mathcal{H} -norm.

Denote the dual space of \mathcal{H} by \mathcal{H}^* , i.e.

$$\mathcal{H}^* := \left\{ \phi : \mathcal{H} \rightarrow \mathbb{R} \mid \sup_{w \in \mathcal{H}} \frac{|\phi(w)|}{\|w\|_{\mathcal{H}}} < \infty \right\}. \quad (1.2)$$

It is well-known that \mathcal{H}^* is isomorphic to \mathcal{H} due to the fact that \mathcal{H} is Hilbert and the Riesz representation theorem. Therefore, one can identify \mathcal{H}^* with \mathcal{H} later.

Now let $\mathcal{V} \subset \mathcal{H}$ be a dense real Hilbert subspace of \mathcal{H} with an inner product $(\cdot, \cdot)_{\mathcal{V}}$ and its associated norm $\|\cdot\|_{\mathcal{V}}$ such that the identity or canonical embedding is *continuous*, i.e.

$$\|v\|_{\mathcal{H}} \lesssim \|v\|_{\mathcal{V}} \quad \forall v \in \mathcal{V}. \quad (1.3)$$

For any functional $w \in \mathcal{H}$, the map $\mathcal{V} \ni v \rightarrow (w, v)$ belongs to the dual space \mathcal{H}^* because, by using the Cauchy-Schwarz inequality and the continuity of the canonical embedding, we have

$$|(w, v)| \leq \|w\|_{\mathcal{H}} \|v\|_{\mathcal{H}} \lesssim \|w\|_{\mathcal{H}} \|v\|_{\mathcal{V}}.$$

The dual space of \mathcal{V} can be defined in the same way as (1.2) and denoted by \mathcal{V}^* . Hence we can embed \mathcal{H} into \mathcal{V}^* and identify \mathcal{H} with its dual \mathcal{H}^* to obtain

$$\mathcal{V} \hookrightarrow \mathcal{H} \simeq \mathcal{H}^* \hookrightarrow \mathcal{V}^*. \quad (1.4)$$

We use $\langle \cdot, \cdot \rangle$ to denote the duality pair between \mathcal{V}^* and \mathcal{V} .

Assume that a bilinear form $a(\cdot, \cdot) : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}$ is *continuous* and *coercive* in the sense that there exists positive constants $C^* \geq C_* > 0$ such that, for any $v, w \in \mathcal{V}$,

$$|a(v, w)| \leq C^* \|v\|_{\mathcal{V}} \|w\|_{\mathcal{V}} \quad \text{continuous} \quad (1.5)$$

$$a(v, v) \geq C_* \|v\|_{\mathcal{V}}^2 \quad \text{coercive.} \quad (1.6)$$

For future reference, we denote $\mathcal{A} : \mathcal{V} \rightarrow \mathcal{V}^*$ to be the linear operator associated with the bilinear form a , i.e.

$$\langle \mathcal{A}v, w \rangle := a(v, w) \quad \forall v, w \in \mathcal{V}.$$

Using (1.5) and (1.6), we obtain that, for any $v \in \mathcal{V}$,

$$\|\mathcal{A}v\|_{\mathcal{V}^*} \leq C^* \|v\|_{\mathcal{V}} \quad \text{continuous} \quad (1.7)$$

$$\langle \mathcal{A}v, v \rangle \geq C_* \|v\|_{\mathcal{V}}^2 \quad \text{coercive.} \quad (1.8)$$

The natural norm associated with the bilinear form $a(\cdot, \cdot)$ (or operator \mathcal{A}) is usually called the *energy norm* which is denoted by $\|v\| := a(v, v)^{\frac{1}{2}}$. From (1.5) and (1.6), we immediately notice that $\|\cdot\|$ is equivalent to $\|\cdot\|_{\mathcal{V}}$:

$$C_* \|v\|_{\mathcal{V}}^2 \leq \|v\|^2 \leq C^* \|v\|_{\mathcal{V}}^2 \quad v \in \mathcal{V}. \quad (1.9)$$

1.2 Elliptic Variational Inequalities (EVIs)

Now we are ready to introduce the elliptic variational inequality problem in its weak form and discuss general existence results.

1.2.1 Variational Formulation

A general elliptic variational inequality (in weak form) can be written as follows:

Problem 1.1 (Abstract Elliptic Variational Inequality) *Let $\mathcal{K} \subseteq \mathcal{V}$ be a non-empty closed convex set and $f \in \mathcal{V}^*$. Find a $u \in \mathcal{K}$ such that*

$$\langle \mathcal{A}u, u - v \rangle \leq \langle f, u - v \rangle \quad \forall v \in \mathcal{K}. \quad (1.10)$$

Remark 1.2 (Relation with Variational Equations) It is easy to see that, if $\mathcal{K} = \mathcal{V}$, then the variational inequality (1.10) reduces to a variational equation

$$\langle \mathcal{A}u, v \rangle = \langle f, v \rangle \quad \forall v \in \mathcal{V}. \quad (1.11)$$

If we assume further that the bilinear form $a(\cdot, \cdot)$ is symmetric or \mathcal{A} is self-adjoint, then the Problem 1.1 can be rewritten as the following energy minimization problem:

Problem 1.3 (Convex Energy Minimization) *Let $\mathcal{K} \subseteq \mathcal{V}$ be a non-empty closed convex set and $f \in \mathcal{V}^*$, find $u \in \mathcal{K}$ such that*

$$\min_{v \in \mathcal{K}} \mathcal{J}(v) := \frac{1}{2}a(v, v) - \langle f, v \rangle. \quad (1.12)$$

Remark 1.4 (Equivalence) It is easy to see that (1.10) is the first order necessary condition for the constrained convex minimization problem, Problem 1.3. Furthermore, it is also sufficient because the objective function is strictly convex from (1.6). Hence the equivalence follows by elementary optimization theory. For details, see [33].

1.2.2 Existence and Uniqueness

The general existence theory has been developed by Lions and Stampacchia [93]. The proof of the existence can be reduced to an application of Banach fixed point theorem following the constructive approach in Rodrigues [120, Theorem 3.1].

Theorem 1.5 (Existence and Uniqueness) *Let $a(\cdot, \cdot) : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}$ be a continuous and coercive bilinear form on \mathcal{V} and \mathcal{K} be a non-empty closed convex subset of \mathcal{V} . Then, for any $f \in \mathcal{V}^*$, there exists a unique solution $u \in \mathcal{K}$ of the variational inequality (1.10).*

Proof of Existence. For any fixed $u \in \mathcal{V}$, the mapping $v \mapsto a(u, v)$ is in the dual space \mathcal{V}^* . We can find $\mathcal{B}u \in \mathcal{V}$ such that $a(u, v) = (\mathcal{B}u, v)_{\mathcal{V}}$. In the same spirit, we can find a representation of $f \in \mathcal{V}^*$, denoted by $b \in \mathcal{V}$. Rewrite the EVI (1.10) as: find

$$u \in \mathcal{K} : \quad (\mathcal{B}u, v - u)_{\mathcal{V}} \leq (b, v - u)_{\mathcal{V}} \quad \forall v \in \mathcal{K}.$$

For any constant $\beta > 0$, the above inequality is equivalent to

$$u \in \mathcal{K} : \quad (\beta b - \beta \mathcal{B}u + u - u, v - u)_{\mathcal{V}} \leq 0 \quad \forall v \in \mathcal{K}. \quad (1.13)$$

Let $P_{\mathcal{K}}(\cdot) : \mathcal{V} \rightarrow \mathcal{K}$ be the *projection* operator onto \mathcal{K} with respect to $(\cdot, \cdot)_{\mathcal{V}}$, i.e. for any $w \in \mathcal{V}$

$$(w - P_{\mathcal{K}}(w), v - P_{\mathcal{K}}(w))_{\mathcal{V}} \leq 0 \quad \forall v \in \mathcal{K}. \quad (1.14)$$

It then follows that (1.13) is equivalent to the nonlinear equation

$$u = P_{\mathcal{K}}(\beta b - \beta \mathcal{B}u + u). \quad (1.15)$$

Define, for any $v \in \mathcal{K}$, $G_{\beta}(v) := P_{\mathcal{K}}(\beta b - \beta \mathcal{B}v + v)$. Since a projection onto a closed convex set is Lipschitz with constant 1, we have

$$\|G_{\beta}(v_1) - G_{\beta}(v_2)\|_{\mathcal{V}} \leq \|(v_1 - v_2) - \beta \mathcal{B}(v_1 - v_2)\|_{\mathcal{V}} \quad \forall v_1, v_2 \in \mathcal{K}.$$

Hence, by using (1.5) and (1.6),

$$\begin{aligned} \|G_{\beta}(v_1) - G_{\beta}(v_2)\|_{\mathcal{V}}^2 &\leq \|v_1 - v_2\|_{\mathcal{V}}^2 - 2\beta(\mathcal{B}(v_1 - v_2), v_1 - v_2)_{\mathcal{V}} + \beta^2\|\mathcal{B}(v_1 - v_2)\|_{\mathcal{V}}^2 \\ &\leq (1 - 2\beta C_* + \beta^2 C^{*2})\|v_1 - v_2\|_{\mathcal{V}}^2. \end{aligned}$$

By choosing $\beta \in (0, 2C_*/C^{*2})$, we can make G_β a contraction and the existence of solution follows by the well-known Banach fixed point theorem. \blacksquare

Remark 1.6 (Lax-Milgram Theorem) This general existence and uniqueness theory is the so-called Lions-Stampacchia theorem [93]. In the case $\mathcal{K} = \mathcal{V}$ (no active constraint), this theorem reduces to the well-known Lax-Milgram theorem.

Remark 1.7 (Representation of \mathcal{B}) Let $I : \mathcal{V} \rightarrow \mathcal{V}^*$ be the canonical embedding operator characterized by

$$\langle Iu, v \rangle = (u, v)_{\mathcal{V}} \quad \forall u, v \in \mathcal{V}.$$

It follows directly that

$$(\mathcal{B}u, v)_{\mathcal{V}} = \langle I\mathcal{B}u, v \rangle \quad \forall u, v \in \mathcal{V}.$$

Since $a(u, v) = \langle \mathcal{A}u, v \rangle$, we then have $\mathcal{B} = I^{-1}\mathcal{A}$ formally.

Remark 1.8 (A Numerical Method) The constructive proof of Theorem 1.5 above suggests the following iterative method for the approximation of u : taking an initial guess $u^{(0)} \in \mathcal{K}$, obtain a sequence of approximate solutions $\{u^{(i)}\} \subset \mathcal{K}$ by

$$u^{(i)} = P_{\mathcal{K}}(u^{(i-1)} + \beta(b - \mathcal{B}u^{(i-1)})) \quad i = 1, 2, \dots$$

If $\beta \in (0, 2C_*/C^{*2})$, then this iterative method converges uniformly. To maximize the convergence speed, we can choose $\beta_* = C_*/C^{*2}$.

Remark 1.9 (A Different Approach) There is a different proof due to Stampacchia [84] which proves the existence result for symmetric problems via the well-known minimization principle and uses a continuation argument to handle the nonsymmetric part.

Remark 1.10 (Lipschitz Continuity and Uniqueness) Suppose $f_1, f_2 \in \mathcal{V}^*$ and $u_1, u_2 \in \mathcal{K}$ be the corresponding solutions of Problem 1.1, respectively. Taking

$v = u_2(u_1)$ in the variational inequality for $u_1(u_2)$ and adding the corresponding inequalities together, we obtain

$$a(u_1 - u_2, u_1 - u_2) \leq \langle f_1 - f_2, u_1 - u_2 \rangle.$$

By coercivity of $a(\cdot, \cdot)$ and Cauchy-Schwarz inequality,

$$C_* \|u_1 - u_2\|_{\mathcal{V}}^2 \leq \langle f_1 - f_2, u_1 - u_2 \rangle \leq \|f_1 - f_2\|_{\mathcal{V}^*} \|u_1 - u_2\|_{\mathcal{V}}.$$

Hence we can see that the mapping $f \rightarrow u$ is Lipschitz with Lipschitz constant $1/C_*$, i.e.

$$\|u_1 - u_2\|_{\mathcal{V}} \leq \frac{1}{C_*} \|f_1 - f_2\|_{\mathcal{V}^*}. \quad (1.16)$$

Uniqueness of the solution follows directly from (1.16).

1.3 Parabolic Variational Inequalities (PVI)

In this section, we deal with time dependent problems. We consider a time interval $[0, T] \subset \mathbb{R}$. More general time spans can always be shifted to this one.

Let \mathcal{H} and \mathcal{V} be the Hilbert spaces we defined in section 1.2. For $1 \leq p \leq \infty$, we introduce the concept of Bochner spaces. We denote by $L^p(0, T; \mathcal{H})$ the space of Lebesgue measurable functions $u : [0, T] \rightarrow \mathcal{H}$, with bounded norm

$$\|u\|_{L^p(0, T; \mathcal{H})} := \begin{cases} \left(\int_0^T \|u(t)\|_{\mathcal{H}}^p dt \right)^{1/p} & \text{if } 1 \leq p < \infty \\ \text{ess sup}_{t \in (0, T)} \|u(t)\|_{\mathcal{H}} & \text{if } p = \infty. \end{cases} \quad (1.17)$$

Other spaces, like $L^p(0, T; \mathcal{V})$, can be defined analogously.

1.3.1 Weak Formulation

In this part, we assume the operator \mathcal{A} satisfies the assumptions in §1.1. Similar to the elliptic case we discussed in the previous section, we first give the standard variational form of abstract PVIs:

Problem 1.11 (Abstract Parabolic Variational Inequalities) *Let $\mathcal{K} \subseteq \mathcal{V}$ be a non-empty closed convex set, $u_0 \in \mathcal{H}$ be a given initial solution, $T \leq \infty$ be the end time and $f : (0, T) \rightarrow \mathcal{V}^*$. Find $u : (0, T) \rightarrow \mathcal{V}$ such that $u(t) \in \mathcal{K}(t)$ a.e. $t \in (0, T)$ satisfying $u(0) = u_0$ and*

$$\langle \partial_t u(t) + \mathcal{A}u(t) - f(t), u(t) - v \rangle \leq 0 \quad \forall v \in \mathcal{K}(t) \text{ a.e. } t \in (0, T), \quad (1.18)$$

where $\partial_t u$ is the partial derivative of u in time variable.

This formulation is a natural extension from the elliptic problem in §1.2. It is convenient to introduce a weaker form for the purpose of the existence and uniqueness discussion.

Remark 1.12 (Sum Space and Its Dual) We introduce the spaces of “sum” and of “intersection” type by

$$\mathfrak{S}(0, T) := L^1(0, T; \mathcal{H}) + L^2(0, T; \mathcal{V}^*), \quad (1.19)$$

$$\mathfrak{J}(0, T) := L^\infty(0, T; \mathcal{H}) \cap L^2(0, T; \mathcal{V}), \quad (1.20)$$

The “sum” space $\mathfrak{S}(0, T)$ and its dual space $\mathfrak{J}(0, T)$ will be useful when we discuss the concept of weak solutions of time-dependent problems.

Now, we assume that u is the solution of Problem 1.11 and v satisfy

$$u, \partial_t u \in \mathfrak{J}(0, T), \quad u(t) \in \mathcal{K} \quad \text{a.e. } t \in (0, T), \quad (1.21)$$

$$v, \partial_t v \in \mathfrak{J}(0, T), \quad v(t) \in \mathcal{K} \quad \text{a.e. } t \in (0, T). \quad (1.22)$$

The inequality (1.18) gives that

$$\frac{d}{dt} \left(\frac{1}{2} \|u - v\|_{\mathcal{H}}^2 \right) + \langle \partial_t v + \mathcal{A}u - f, u - v \rangle \leq 0. \quad (1.23)$$

Although \mathcal{K} is closed in \mathcal{V} , it is not necessarily closed in \mathcal{H} . We denote $\overline{\mathcal{K}}$ be the closure of \mathcal{K} respect to the norm $\|\cdot\|_{\mathcal{H}}$. Then we define the weak solution of Problem 1.11 as following:

Definition 1.13 (Weak Solution of PVI) Give $u_0 \in \overline{\mathcal{K}}$, $0 < T \leq \infty$, $f \in \mathfrak{S}(0, T)$ and \mathcal{K} is a nonempty, closed, and convex subset in \mathcal{V} . Find

$$u \in \mathfrak{J}(0, T), \quad u(t) \in \mathcal{K} \quad \text{a.e. } t \in (0, T),$$

such that

$$\Theta'(t) \leq 0 \quad \text{and} \quad \Theta(t) \leq \frac{1}{2} \|u_0 - v(0)\|_{\mathcal{H}}^2 \quad \text{a.e. } t \in (0, T),$$

for all v satisfying (1.22), where

$$\Theta(t) := \frac{1}{2} \|u(t) - v(t)\|_{\mathcal{H}}^2 + \int_0^t \langle \partial_s v(s) + \mathcal{A}u(s) - f(s), u(s) - v(s) \rangle ds. \quad (1.24)$$

We call u the weak solution of PVI (1.18).

1.3.2 Existence and Uniqueness

Existence and uniqueness of the weak solution of parabolic variational inequality holds under very general assumptions on \mathcal{A} , u_0 and f ; for example \mathcal{A} could be a nonlinear monotone operator. We refer readers to the monograph [31]. Regularity as well as approximation results of the weak solution of general PVI can be found in Baiocchi [11]. For obstacle problems (which is the main topic of the thesis and will be discussed in the following chapters), Ito and Kunisch [79] introduced a new approach using Lagrange multiplier technique and proved the existence of strong and weak solutions.

Here we simply review the classical existence result of weak solution.

Theorem 1.14 (Existence and Uniqueness of Weak Solution) For any initial solution $u_0 \in \mathcal{H}$ and data $f \in L^2(0, T; \mathcal{V}^*)$, there exists a unique weak solution of Problem 1.11 and $u \in C^0([0, T]; \mathcal{H})$.

Remark 1.15 (Uniqueness) If $u_0 \in \mathcal{H}$ but not in \mathcal{K} , we need to modify the definition of the weak solution by replacing u_0 with $P_{\mathcal{K}}(u_0)$. Otherwise there might be multiple weak solutions; see [11].

Chapter 2

Obstacle Problems

In the previous chapter, we introduced variational inequalities in a general setting. In this chapter, we focus on a particular class of variational inequalities, namely obstacle problems. Obstacle problems were one of the main motivations of the theory of variational inequalities and have many important applications in various areas. One particular application in finance will be discussed in Chapter 3.

It is well known that the solution of an elliptic boundary value problem has certain degree of smoothness depending on the smoothness of the data and the boundary of its physical domain (see, for example, [62]). However, in general, the solution of an obstacle problem associate with a second order differential operator \mathcal{A} cannot be in C^2 even for smooth enough data. Lack of smoothness is one of the difficulties to handle this nonlinear problem.

In this chapter, we review some basic concepts and smoothness of solutions of obstacle problems. This chapter is organized as follows. First, we review the definition of Sobolev spaces of general order and angle-bounded operators in §2.1 and §2.2, respectively. Then we define the static as well as evolution obstacle problems in §2.3.

2.1 Function Spaces

Before we can discuss any concrete obstacle examples, we need to first recall the theory of Hölder spaces and Sobolev spaces [1]. Here we assume that $\Omega \subset$

$\mathbb{R}^d (d = 1, 2, 3)$ be open and bounded with boundary $\Gamma = \partial\Omega := \overline{\Omega} \setminus \Omega$. We denote the set of natural numbers by \mathbb{N} .

2.1.1 Hölder Spaces

Consider functions $v : \Omega \rightarrow \mathbb{R}$. $\beta = (\beta_1, \dots, \beta_d)^T \in \mathbb{N}^d$ be a multi-index of modulus $|\beta| = \sum_{i=1}^d \beta_i$. We denote by

$$D^\beta v := \frac{\partial^{|\beta|} v}{\partial x_1^{\beta_1} \dots \partial x_d^{\beta_d}} \quad (2.1)$$

the partial derivatives of v .

For any nonnegative integer $m \in \mathbb{N}$, we define by $C^m(\Omega)$ the linear space of continuous functions v on Ω whose partial derivatives $D^\beta v (|\beta| \leq m)$ is also continuous. Furthermore, it is a Banach space with the norm

$$\|v\|_{C^m(\Omega)} := \max_{0 \leq |\beta| \leq m} \sup_{x \in \Omega} |D^\beta v(x)|.$$

We define $C^{m,\alpha}(\Omega)$ for $0 < \alpha < 1$ to be the linear subspace of $C^m(\Omega)$ whose m -th order partial derivatives are Hölder continuous, i.e.

$$|D^\beta v(x) - D^\beta v(y)| \leq C_\beta |x - y|^\alpha \quad \forall x, y \in \Omega \text{ and } \beta \in \mathbb{N}^d \text{ with } |\beta| = m.$$

We then note that the Hölder- α space $C^{m,\alpha}(\Omega)$ is Banach with respect to the norm

$$\|v\|_{C^{m,\alpha}(\Omega)} := \|v\|_{C^m(\Omega)} + \max_{|\beta|=m} \sup_{x,y \in \overline{\Omega}} \frac{|D^\beta v(x) - D^\beta v(y)|}{|x - y|^\alpha}.$$

A very special case is that $C^\infty(\Omega)$ is the functions with continuous partial derivatives of any order. We denote by $C_c^\infty(\Omega)$ the subset of $C^\infty(\Omega)$ functions with compact support in Ω .

2.1.2 Sobolev Spaces of Integer Order

Let $L^p(\Omega)$ be the class of all measurable functions v defined on Ω with bounded norm

$$\|v\|_{L^p(\Omega)} := \begin{cases} \left(\int_{\Omega} |v|^p \right)^{1/p} & \text{if } 1 \leq p < \infty \\ \text{ess sup}_{x \in \Omega} |v(x)| & \text{if } p = \infty. \end{cases} \quad (2.2)$$

The space of functions of bounded variation (BV functions) will be useful for error analysis and can be defined as

$$\text{BV}(\Omega) = \left\{ v \in L^1(\Omega) \mid \sup_{\substack{\phi \in C_c^1(\Omega)^d \\ \|\phi\|_{L^\infty(\Omega)} < \infty}} \int_{\Omega} v \operatorname{div} \phi < \infty \right\}. \quad (2.3)$$

Sobolev spaces with integer order are normed spaces of functions with finite weak derivatives in L^p -norm. More precisely, for any nonnegative integer number $s \in \mathbb{N}$ and $1 \leq p \leq \infty$, the space $W^{s,p}(\Omega)$ is defined to be the subset of $L^p(\Omega)$ such that v and its weak derivatives up to order k have a finite $L^p(\Omega)$ -norm. With this definition, the Sobolev spaces admit a natural norm,

$$\|v\|_{W^{s,p}(\Omega)} := \left(\sum_{|\beta| \leq s} \|D^\beta v\|_{L^p(\Omega)}^p \right)^{\frac{1}{p}}.$$

We can identify $W^{0,p}(\Omega)$ with $L^p(\Omega)$. As a convention, we use $H^s(\Omega)$ to denote $W^{s,2}(\Omega)$.

It is well-known that functions in $H^1(\Omega)$ are not necessarily bounded or continuous. So we need to define the boundary value in trace sense. We denote $\tilde{H}^1(\Omega)$ to be the subspace of $H^1(\Omega)$ with zero boundary trace. It is equivalent to define $\tilde{H}^1(\Omega)$ as a completion of $C_c^\infty(\Omega)$ in the $H^1(\Omega)$ -norm (see for example [62, Theorem 2, Page 259]).

For negative integers $s < 0$ and $1 \leq p \leq \infty$, $W^{s,p}(\Omega)$ is the dual space of $\tilde{W}^{-s,q}(\Omega)$. Here q is the dual exponent of p , i.e. $\frac{1}{p} + \frac{1}{q} = 1$ and $q = \infty$ if $p = 1$; $q = 1$ if $p = \infty$.

2.1.3 Sobolev Spaces of Fractional Order

Up to now, the Sobolev spaces of integer order have been defined for any integer number. For noninteger s , there are several ways to define the fractional order norm, for example, by the growth of the Fourier coefficients, or by interpolation theory, or by double integrals.

We first give a definition of Sobolev spaces of noninteger order which also specify their norms. This is important for our later a posteriori error analysis since

different definitions of Sobolev norms behave differently as the domain becomes smaller, or as s approaches certain values (i.e., the equivalence constants are not uniform with respect to the size of the domain and s).

The space $\tilde{H}^s(\Omega)$ is defined for $s \in [0, 1]$ using interpolation [18] of

$$\tilde{H}^0(\Omega) := L^2(\Omega) \quad \text{and} \quad \tilde{H}^1(\Omega) := \{v \in H^1(\Omega) : v|_{\Gamma} = 0\}.$$

The space $H^{-s}(\Omega)$ is the dual space of $\tilde{H}^s(\Omega)$. Notice that this definition using interpolation is not restricted to the case when $p = 2$. $W^{s,p}$ for $0 < s < 1$ and $1 \leq p \leq \infty$ could be defined analogously.

From the definition, it is not difficult to show the following interpolation inequality:

Proposition 2.1 (Interpolation Inequality) *Let $s \in [0, 1]$. The Sobolev space interpolation inequality holds, i.e.*

$$\|v\|_{\tilde{H}^s(\Omega)} \leq \|v\|_{\tilde{H}^1(\Omega)}^s \|v\|_{L^2(\Omega)}^{1-s} \quad \forall v \in \tilde{H}^1(\Omega).$$

We now introduce the local version of the H^s -norm. Let ω be a sub-domain of Ω . We then define the spaces $H_{\Gamma}^s(\omega)$ for $s \in [0, 1]$ using interpolation of

$$L^2(\omega) \quad \text{and} \quad H_{\Gamma}^1(\omega) := \{v \in H^1(\omega) : v|_{\Gamma} = 0\}.$$

We will use $H_{\Gamma}^s(\omega)^*$ to denote the dual space of $H_{\Gamma}^s(\omega)$.

Remark 2.2 (Boundary Conditions) It is worth noticing that functions in $H_{\Gamma}^1(\omega)$ do not necessarily have zero boundary trace in the local domain ω . And it is clear that $H_{\Gamma}^s(\Omega) = \tilde{H}^s(\Omega)$ by their definitions.

Remark 2.3 (General Order) Although we only consider the case where $0 < s < 1$ here, it is clear that general s can be treated in a similar fashion via the interpolation of $\tilde{H}^{\lfloor s \rfloor}(\Omega)$ and $\tilde{H}^{\lfloor s \rfloor + 1}(\Omega)$, where $\lfloor s \rfloor$ is the maximum integer less than or equal to s .

Remark 2.4 (Zero Trace) The $\tilde{H}^s(\Omega)$ space is related to the standard fractional Sobolev space. In fact,

$$\tilde{H}^s(\Omega) = \begin{cases} H^s(\Omega) & 0 \leq s < \frac{1}{2} \\ H_0^s(\Omega) & \frac{1}{2} < s < 1. \end{cases}$$

When $s = \frac{1}{2}$, $\tilde{H}^{\frac{1}{2}}(\Omega)$ is a dense subspace of $H^{\frac{1}{2}}(\Omega)$ and is sometimes denoted by $H_{00}^{\frac{1}{2}}(\Omega)$. See, for example, [92].

2.1.4 Norm Equivalence

Now we will review the norm equivalence results for Sobolev spaces of noninteger order by Faermann [65]. Let $\{v_j\}_{j=1}^J$ be a set of functions with pairwise weakly disjoint support, i.e. the intersection of supports of any two functions v_i and v_j has zero measure. Due to the lack of orthogonality, the ordinary relation

$$\left\| \sum_{j=1}^J v_j \right\|_{H^1(\Omega)}^2 = \sum_{j=1}^J \|v_j\|_{H^1(\Omega)}^2$$

does not hold anymore for fractional-order norm $\|\cdot\|_{H^s(\Omega)}$ even if supports of v_j 's are pairwise disjoint.

In [68, Theorem 2.2], a weaker equivalence result has been proven.

Proposition 2.5 (Norm Equivalence) *There exists a constant $C_s > 0$ such that*

$$C_s^{-1} \sum_{j=1}^J \|v_j\|_{\tilde{H}^s(\Omega)}^2 \leq \left\| \sum_{j=1}^J v_j \right\|_{\tilde{H}^s(\Omega)}^2 \leq C_s \sum_{j=1}^J \|v_j\|_{\tilde{H}^s(\Omega)}^2, \quad (2.4)$$

for $\{v_j\}_{j=1}^J \subseteq \tilde{H}^s(\Omega)$ with weakly disjoint support.

Proof. We first prove the second inequality which is needed in our a posteriori error estimation in Chapter 6. We know that $\{v_j\}_{j=1}^J$ has pairwise weakly disjoint support (pairwise intersection has measure 0). Let Ω_j be the support of v_j for each $1 \leq j \leq J$. We can define an operator $T : \prod_{j=1}^J \tilde{H}^s(\Omega_j) \rightarrow \tilde{H}^s(\Omega)$ such that

$$T(v_1, \dots, v_J) = \sum_{j=1}^J v_j.$$

Obviously, for $s = 0, 1$, we have

$$\|T(v_1, \dots, v_J)\|_{\tilde{H}^s(\Omega)}^2 = \sum_{j=1}^J \|v_j\|_{\tilde{H}^s(\Omega)}^2.$$

Hence the interpolation argument gives the second inequality with $C_s = 1$. The reverse direction is much more involved and we refer readers to [65]. ■

Remark 2.6 (Different Definitions of Sobolev Spaces of Fractional Order)

Faermann defined in [65] the noninteger Sobolev norm by extending function v to \mathbb{R}^d and then using Fourier transform:

$$\|v(x)\|_{H^s(\mathbb{R}^d)} := \|(1 + |\xi|^2)^{s/2} \hat{v}(\xi)\|_{L^2(\mathbb{R}^d)},$$

where \hat{v} is the Fourier transformation of v . It is well known that this norm is equivalent to the \tilde{H}^s -norm by zero extension.

Remark 2.7 (Applications in A Posteriori Error Estimation) Based on this equivalence result, Faermann [68] gave a reliable and efficient (but unfortunately not computable) error estimator for boundary element methods for integral equations with s not an integer.

2.2 Angle-Bounded Operators

With the definitions of Sobolev spaces in the previous two subsections, from now on, we fix the general Hilbert triple consider in Chapter 1 to be the following particular setting (still quite general though):

$$(\mathcal{V}, \mathcal{H}, \mathcal{V}^*) := \left(\tilde{H}^s(\Omega), L^2(\Omega), H^{-s}(\Omega) \right),$$

for $0 \leq s \leq 1$. Furthermore, we will consider a class of operators \mathcal{A} , namely *angle-bounded* operators, in the following chapters (especially for applications in finance). This notion was introduced by Brézis and Browder [32] as a nonlinear generalization of sectorial operators, and more recently revisited by Caffarelli in the context of regularity theory [38].

2.2.1 Angle Condition

Let $\mathcal{A} : \tilde{H}^s(\Omega) \rightarrow H^{-s}(\Omega)$ be a linear monotone operator and we employ the energy norm (if it is a norm) as defined in Chapter 1

$$\|v\| := \langle \mathcal{A}v, v \rangle^{\frac{1}{2}} \quad \forall v \in \tilde{H}^s(\Omega),$$

induced by the operator \mathcal{A} , as well as its dual norm $\|\cdot\|_*$.

Definition 2.8 (Sectorial Operator) *A linear monotone operator \mathcal{A} is called sectorial if it satisfies the strong sector condition*

$$|\langle \mathcal{A}v, w \rangle|^2 \leq 4\gamma^2 \|v\|^2 \|w\|^2 \quad \forall v, w \in \tilde{H}^s(\Omega). \quad (2.5)$$

This is equivalent to the following inequality for the skew-symmetric part of \mathcal{A} [32, Prop. 11]:

$$|\langle \mathcal{A}v, w \rangle - \langle \mathcal{A}w, v \rangle| \leq 2\lambda \|v\| \|w\| \quad \forall v, w \in \tilde{H}^s(\Omega), \quad (2.6)$$

with a positive constant λ satisfying $\gamma^2 = (\lambda^2 + 1)/4$. We observe that (2.5) implies that \mathcal{A} is Lipschitz continuous and

$$\|\mathcal{A}v\|_* := \sup_{w \in \tilde{H}^s(\Omega)} \langle \mathcal{A}v, w \rangle / \|w\|$$

satisfies

$$\frac{1}{4\gamma^2} \|\mathcal{A}v\|_*^2 \leq \|v\|^2 \leq \|\mathcal{A}v\|_*^2 \quad \forall v \in \tilde{H}^s(\Omega).$$

Definition 2.9 (Angle-bounded) *Let \mathcal{H} be a Hilbert space, and let $\mathcal{D}(\mathfrak{F}) \subset \mathcal{H}$ be the domain of an operator $\mathfrak{F} : \mathcal{H} \rightarrow 2^{\mathcal{H}}$. Then \mathfrak{F} is said to be γ^2 -angle-bounded if there exists a positive constant γ such that*

$$\langle \mathfrak{F}(v) - \mathfrak{F}(w), w - z \rangle \leq \gamma^2 \langle \mathfrak{F}(v) - \mathfrak{F}(z), v - z \rangle \quad \forall v, w, z \in \mathcal{D}(\mathfrak{F}). \quad (2.7)$$

Lemma 2.10 (Equivalence) *The conditions (2.5) and (2.7) are equivalent for \mathcal{A} linear.*

Proof. We simply set $\tilde{v} = v - z$ and $\tilde{w} = w - z$ in (2.7) to get the equivalent formulation (we omit the tildes)

$$\langle \mathcal{A}v, w \rangle \leq \gamma^2 \langle \mathcal{A}v, v \rangle + \langle \mathcal{A}w, w \rangle \quad \forall v, w \in \mathcal{D}(\mathcal{A}). \quad (2.8)$$

Then replace v by αv with $\alpha \in \mathbb{R}$ and argue with the resulting quadratic inequality in α , i.e.

$$\alpha^2 \gamma^2 \langle \mathcal{A}v, v \rangle - \alpha \langle \mathcal{A}v, w \rangle + \langle \mathcal{A}w, w \rangle \geq 0$$

to realize that (2.5) and (2.8) are equivalent. \blacksquare

2.2.2 Coercivity Property

We conclude this section with the *coercivity* property [110, Lemma 4.3], which will be crucial in a posteriori error estimation later in Chapter 6.

Lemma 2.11 (Coercivity) *Let the linear sectorial operator \mathcal{A} satisfy the condition (2.7) (γ^2 -angle-bounded). Then we have*

$$\langle \mathcal{A}v - \mathcal{A}w, w - z \rangle \leq 2\gamma^2 \|v - z\|^2 - \frac{1}{4} \left(\|v - w\|^2 + \|z - w\|^2 \right) \quad \forall v, w, z \in \mathcal{K}. \quad (2.9)$$

Proof. In view of the Cauchy-Schwarz inequality, we get

$$\begin{aligned} \langle \mathcal{A}v - \mathcal{A}w, w - z \rangle &= \langle \mathcal{A}v - \mathcal{A}w, w - v \rangle + \langle \mathcal{A}v - \mathcal{A}w, v - z \rangle \\ &\leq -\|v - w\|^2 + 2\gamma \|v - w\| \|v - z\| \\ &\leq -\frac{1}{2} \|v - w\|^2 + 2\gamma^2 \|v - z\|^2. \end{aligned}$$

Similarly, we get

$$\begin{aligned} \langle \mathcal{A}v - \mathcal{A}w, w - z \rangle &= \langle \mathcal{A}z - \mathcal{A}w, w - z \rangle + \langle \mathcal{A}v - \mathcal{A}z, w - z \rangle \\ &\leq -\|z - w\|^2 + 2\gamma \|v - z\| \|w - z\| \\ &\leq -\frac{1}{2} \|z - w\|^2 + 2\gamma^2 \|v - z\|^2. \end{aligned}$$

Adding the last two inequalities gives (2.9). \blacksquare

2.3 Obstacle Problems

This presentation mainly follows Rodrigues [120] and Friedman [73]. Unfortunately, it is impossible to review all regularity results available in the literature. For regularity results for other types of variational inequalities, like the case of gradient constraint, the biharmonic obstacle problems, etc, we refer to the monograph [31].

Remark 2.12 (EVI and PVI) Since we shall focus on the variational inequalities with obstacle type constraints throughout this note, we will later refer to elliptic and parabolic obstacle problems as EVI and PVI, respectively, with a little abuse of notation.

2.3.1 Elliptic Obstacle Problems

Problem 2.13 (Elliptic Obstacle Problems) *Suppose in Problem 1.1, the convex set has the following structure*

$$\mathcal{K} := \{v \in \mathcal{V} \mid v \geq \chi\}, \quad (2.10)$$

where the function $\chi \in \mathcal{V}$ is the so-called obstacle. The corresponding VI problem

$$(VI) \quad \text{Find } u \in \mathcal{K} : \langle \mathcal{A}u - f, u - v \rangle \leq 0 \quad \forall v \in \mathcal{K}, \quad (2.11)$$

is called the elliptic obstacle problem.

Suppose $u \in \mathcal{K}$ is the solution of the obstacle problem, Problem 2.13, the set of points $\mathcal{C}(u) := \{x \in \Omega : u(x) = \chi(x)\}$ is called the *contact set* or *coincidence set*, and its complement $\mathcal{N}(u) = \Omega \setminus \mathcal{C}(u)$ the *noncontact set* or *non-coincidence set*. The boundary $\mathcal{F}(u)$ between the two sets is called the *free boundary* or *free interface*.

From now on, we use v_+ (v_-) to be the non-negative part of a function v ($-v$), i.e., $v_+ = \max\{v, 0\}$ and $v_- = -\min\{v, 0\}$.

We start by stating without proof a useful but relatively restricted regularity result [120]:

Proposition 2.14 (General Regularity Result) *Assume that*

$$(\chi - v)_+ \in \mathcal{V}, \quad \forall v \in \mathcal{V} \quad \text{and} \quad \|v_{\pm}\|_{\mathcal{H}} \leq \|v\|_{\mathcal{H}}, \quad \forall v \in \mathcal{H}. \quad (2.12)$$

If $f \in \mathcal{H}$ and $(\mathcal{A}\chi - f)_+ \in \mathcal{H}$, then the solution u of the obstacle problem

$$u \in \mathcal{K} : \quad \langle \mathcal{A}u - f, u - v \rangle \leq 0 \quad \forall v \in \mathcal{K}. \quad (2.13)$$

satisfies the estimate

$$\|\mathcal{A}u\|_{\mathcal{H}} \leq \|f\|_{\mathcal{H}} + \|(\mathcal{A}\chi - f)_+\|_{\mathcal{H}}.$$

Remark 2.15 (Dirichlet Obstacle Problem) The simplest example of \mathcal{A} is the Laplace operator, $-\Delta$. In this case, we take the Hilbert triple to be

$$\mathcal{V} = \tilde{H}^1(\Omega), \quad \mathcal{V}^* = H^{-1}(\Omega), \quad \text{and} \quad \mathcal{H} = L^2(\Omega) = \mathcal{H}^*.$$

The bilinear form $a(\cdot, \cdot) = (\nabla \cdot, \nabla \cdot)$ is an inner product which induces the energy norm for the Laplace equations. A direct application of Proposition 2.14 to the Dirichlet obstacle problem gives $H^2(\Omega)$ -regularity of the solution assuming $f \in L^2(\Omega)$, $\chi \in H^2(\Omega)$, $\chi \leq 0$ on $\partial\Omega$ and Ω being convex or $\partial\Omega \in C^{1,1}$ (see Brézis and Stampacchia [34]). It has been shown that the solution u of a Dirichlet Obstacle problem can never be better than $C^{1,1}(\Omega)$ regardless how smooth the obstacle χ and data f are (see Caffarelli [39]).

2.3.2 Equivalent Formulations

There are several different ways to formulate the variational inequality problem. We now discuss some of its equivalent formulations briefly.

Complementarity Problems

The most frequently used form is linear complementarity problem (LCP): find a solution $u \in \mathcal{V}$ such that

$$(\text{LCP}) \quad \begin{cases} \mathcal{A}u - f \geq 0 \\ u - \chi \geq 0 \\ \langle \mathcal{A}u - f, u - \chi \rangle = 0. \end{cases} \quad (2.14)$$

The last equation is the so-called *complementarity condition*. This is actually equivalent to (2.11) if $\chi \in \mathcal{V}$.

Proof of Equivalence. If u is a solution of LCP (2.14), then for any $v \in \mathcal{V}$ and $v \geq \chi$ we have

$$\langle \mathcal{A}u - f, u - v \rangle = \langle \mathcal{A}u - f, \chi - v \rangle \leq 0,$$

in view of the complementarity condition and the sign condition of $\mathcal{A}u - f$.

On the other hand, if u is solution of VI (2.11), it is trivial to see that u satisfies the first two conditions of LCP. The complementarity condition is obtained by taking $v = u + (u - \chi)$ and $v = \chi$. ■

Nonlinear Equation

Motivated by the proof of existence theorem 1.5, we can formulate the VI (2.11) as a nonlinear projection equation

$$(NE) \quad u = P_{\mathcal{K}}(u + (b - \mathcal{B}u)), \quad (2.15)$$

where $P_{\mathcal{K}}(\cdot) : \mathcal{V} \rightarrow \mathcal{K}$ is the projection operator defined as (1.14).

Proof of Equivalence. First the VI problem can be written equivalently as

$$(\mathcal{B}u - b, u - v)_{\mathcal{V}} \leq 0 \quad \forall v \in \mathcal{K}. \quad (2.16)$$

Define $e := u - P_{\mathcal{K}}(u + (b - \mathcal{B}u))$. If u is solution of VI (2.11), by taking $v = u - (b - \mathcal{B}u)$ and $v = u$ in the definition of projection (1.14), we get that

$$(e - (b - \mathcal{B}u), e)_{\mathcal{V}} \leq 0.$$

This, in turn, gives the sign condition

$$(b - \mathcal{B}u, e)_{\mathcal{V}} \geq \|e\|_{\mathcal{V}}^2 \geq 0.$$

By taking $v = P_{\mathcal{K}}(u - (b - \mathcal{B}u))$ in (2.16), we get $(b - \mathcal{B}u, e)_{\mathcal{V}} \leq 0$. Hence $\|e\|_{\mathcal{V}} = 0$.

The converse direction can be derived directly from (1.14) by taking $w = u - (b - \mathcal{B}u)$. ■

Variational Inclusion Problem

The VI (2.11) can also be viewed as an inclusion problem. If we write the VI problem as a variational inequality of second-type¹:

$$(VI2) \quad \langle \mathcal{A}u - f, u - v \rangle + \mathcal{I}_{\mathcal{K}}(u) - \mathcal{I}_{\mathcal{K}}(v) \leq 0 \quad \forall v \in \mathcal{V}. \quad (2.17)$$

Here $\mathcal{I}_{\mathcal{K}}$ is the indicator function of the convex set \mathcal{K} and it is convex lower semi-continuous:

$$\mathcal{I}_{\mathcal{K}}(v) := \begin{cases} 0 & \text{if } v \in \mathcal{K} \\ \infty & \text{if } v \notin \mathcal{K} \end{cases}$$

When \mathcal{A} is symmetric, it is clear that this problem is equivalent to a convex minimization problem

$$\min_{v \in \mathcal{V}} \frac{1}{2}a(v, v) - \langle f, v \rangle + \mathcal{I}_{\mathcal{K}}(v).$$

A more general formulation is given by Brezis and Stampacchia [34]. VI (2.17) can be written as a variational inclusion problem (IP):

$$(IP) \quad \mathcal{A}u + \partial\mathcal{I}_{\mathcal{K}}(u) \ni f. \quad (2.18)$$

Notice that the convex function $\mathcal{I}_{\mathcal{K}} : \mathbb{R} \rightarrow \mathbb{R}$ might not be differential in usual sense. We use the more general subdifferential mapping $\partial\mathcal{I}_{\mathcal{K}}$, which is a multivalued map such that, for any value $c \in \partial\mathcal{I}_{\mathcal{K}}(x)$

$$\mathcal{I}_{\mathcal{K}}(y) - \mathcal{I}_{\mathcal{K}}(x) \geq c(y - x) \quad \forall y \in \mathbb{R}.$$

Remark 2.16 (Lagrange Multiplier) If \mathcal{K} is the convex set defined in (2.21), we let $\mathfrak{F} : \mathcal{K} \rightarrow H^{-s}(\Omega)$ be the multivalued operator associated with the variational inequality in \mathcal{K} , i.e.

$$v_* \in \mathfrak{F}(v) \Leftrightarrow a(v, v - w) \leq \langle v_*, v - w \rangle \quad \forall w \in \mathcal{K}. \quad (2.19)$$

For details, see § 2.3.2. If we further define the multivalued operator $\lambda(v) := \mathfrak{F}(v) - \mathcal{A}v$ with $\mathcal{D}(\lambda) = \mathcal{K}$, we see that $\lambda(v) \leq 0$ in Ω and $\lambda(v) = 0$ in $\mathcal{N} = \{v > \chi\}$ (simply

¹The variational inequality in the form (2.11) is usually called variational inequality of first-type.

argue with $w = v + \varphi$). It turns out that λ is the subdifferential $\partial\mathcal{I}_{\mathcal{K}}$. Such a λ can be viewed as a Lagrange multiplier (see Definition 2.22 in §2.3.4) of the constraint $v \geq \chi$.

The following lemma provides an important insight for a posteriori error estimation which will be discussed in Chapter 6.

Lemma 2.17 (\mathfrak{F} is Angle-Bounded) *If \mathcal{A} is γ^2 -angle-bounded (see Definition 2.9), then the nonlinear operator $\mathfrak{F} = \mathcal{A} + \lambda$ is γ_0^2 -angle-bounded with constant $\gamma_0 = \max(1, \gamma)$. Moreover, \mathfrak{F} satisfies for all $v, w, z \in \mathcal{K}$*

$$\begin{aligned} \langle \mathfrak{F}(v) - \mathfrak{F}(w), w - z \rangle &\leq \gamma^2 \langle \mathcal{A}v - \mathcal{A}z, v - z \rangle + \langle \lambda(v), v - z \rangle \\ &\leq \gamma^2 \langle \mathcal{A}v - \mathcal{A}z, v - z \rangle + \langle \lambda(v) - \lambda(z), v - z \rangle. \end{aligned} \quad (2.20)$$

Proof. Since $\mathfrak{F}(v) = \mathcal{A}v + \lambda(v)$, in view of Lemmas 2.10 and (2.5) we only need to deal with $\lambda(v)$. We resort to the fact that $\lambda(v) = \partial\mathcal{I}_{\mathcal{K}}(v)$, which translates into the property

$$\langle \lambda(v), w - v \rangle \leq 0 \quad \forall v, w \in \mathcal{K}.$$

In fact, if $v > \chi$ then $\lambda(v) = 0$ whereas if $v = \chi \leq w$ then $\lambda(v) \leq 0$. Consequently

$$\begin{aligned} \langle \lambda(v) - \lambda(w), w - z \rangle &= \langle \lambda(v), v - z \rangle + \langle \lambda(v), w - v \rangle + \langle \lambda(w), z - w \rangle \\ &\leq \langle \lambda(v), v - z \rangle \leq \langle \lambda(v) - \lambda(z), v - z \rangle, \end{aligned}$$

whence we deduce (2.20)

$$\begin{aligned} \langle \mathfrak{F}(v) - \mathfrak{F}(w), w - z \rangle &\leq \gamma^2 \langle \mathcal{A}v - \mathcal{A}w, w - z \rangle + \langle \lambda(v) - \lambda(z), v - z \rangle \\ &\leq \gamma_0^2 \langle \mathfrak{F}(v) - \mathfrak{F}(z), v - z \rangle. \end{aligned}$$

The last inequality implies that \mathfrak{F} is γ_0^2 -angle-bounded, as asserted. ■

2.3.3 Parabolic Obstacle Problems

The parabolic obstacle problems can be defined in an analogous way,

Problem 2.18 (Parabolic Obstacle Problems) *Suppose that, in (1.18), the convex set has the following structure*

$$\mathcal{K} := \{v \in \mathcal{V} \mid v \geq \chi(t) \quad \text{a.e. } t \in (0, T)\}. \quad (2.21)$$

Then the corresponding variational inequality problem, Problem 1.11, is called the parabolic obstacle problem.

Remark 2.19 (Equivalent Formulations) Similar to the elliptic problem (Problem 2.13) discussed in §2.3.2, we can write the parabolic problem (Problem 2.18) as equivalent LCP, NE, IP formulations also.

For $\mathcal{V} = \tilde{H}^1(\Omega)$ and a second order elliptic operator $\mathcal{A} : \tilde{H}^1(\Omega) \rightarrow H^{-1}(\Omega)$ satisfying (1.5) and (1.6), the following classical regularity result is well-known (see [30, Section 2.4]).

Lemma 2.20 (Regularity) *Suppose the obstacle $\chi(t) \in H^2(\Omega)$ a.e. $t \in (0, T)$ and $\chi(t) < 0$ on the boundary $(0, T) \times \Gamma$. If*

$$f \in C([0, T]; L^2(\Omega)), \quad \frac{\partial f}{\partial t} \in L^1(0, T; L^2(\Omega)), \quad \text{and} \quad u_0 \in H^2(\Omega) \cap \mathcal{K},$$

then the problem 2.18 has a unique solution u satisfying

$$u \in L^\infty(0, T; H^2(\Omega)), \quad \frac{\partial u}{\partial t} \in L^\infty(0, T; L^2(\Omega)) \cap L^2(0, T; H^1(\Omega)).$$

Remark 2.21 (Singularity in Time Horizon) For parabolic problems without constraint ($\mathcal{K} = \mathcal{V}$), the smoothness of u in time is directly related to the smoothness of f in time under compatibility assumptions of f and u_0 on Γ . In fact,

$$f \in H^s(0, T; \mathcal{V}^*) \implies u \in H^s(0, T; \mathcal{V}) \cap H^{s+1}(0, T; \mathcal{V}^*).$$

On the contrary, for obstacle problems, no matter how smooth u_0 and f are, the time derivative $\partial_t u$ could be discontinuous.

2.3.4 Lagrange Multiplier

We now look at a very important quantity for constrained energy minimization, namely the Lagrange multiplier. In Chapter 6, we shall employ it for a posteriori error estimation.

Definition 2.22 (Lagrange Multiplier) *We denote the residual of u by*

$$\mathcal{V}^* \ni \lambda(u) := \begin{cases} f - \mathcal{A}u & \text{for elliptic problems} \\ f - \partial_t u - \mathcal{A}u & \text{for parabolic problems;} \end{cases} \quad (2.22)$$

$\lambda(u)$ is often referred to as the Lagrange multiplier.

It is clear that $\lambda = 0$ for problems without obstacle constraint (linear equations). For problems with constraint, this quantity encodes information about the contact region. It may be regarded as a reaction in elasticity applications.

To be able to understand the properties of λ better, we first look at the elliptic obstacle problem. It is easy to see, from the definition of λ as well as the variational inequalities (1.10), that

$$\begin{cases} \lambda \leq 0 \\ \lambda = f - \mathcal{A}\chi & \text{in } \mathcal{C}(u) \\ \lambda = 0 & \text{in } \mathcal{N}(u) \end{cases} \quad (2.23)$$

These important characteristics of λ tells us:

- When the constraint is not active ($\mathcal{N}(u)$ or $u > \chi$), λ vanishes as in the linear equations.
- When the constraint is active ($\mathcal{C}(u)$ or $u = \chi$), $\lambda < 0$ is nonzero; furthermore, the magnitude of λ measures the interaction between the solution and the obstacle.

Remark 2.23 (First-order Optimal Condition) The condition (2.23) can be viewed as an extension of first-order optimal condition for constrained minimization problems. For stationary problems, when \mathcal{A} is symmetric, continuous, and coercive, (2.23) is equivalent to the well-known Karush-Kuhn-Tucker (KKT) condition [89] for constrained minimization.

Chapter 3

Option Pricing – An Application in Finance

The evaluation of the price of an option contract is of considerable importance in finance [78]. It is well-known that there is no general closed-form analytical solution for the price of American-style options. To solve this problem, people usually resort to numerical methods, whose improvement is still an active field of research. The American-style option pricing problem based on the classical Black-Scholes model can be written as a variational inequality for a differential operator. This reformulation is crucial to construct a successful numerical treatment of the problem, as suggested by Wilmott, Dewynne, and Howison [142]. However, in some more advanced models (like the CGMY model [40]), the problem is more complicated and involves a pseudo-differential operator.

3.1 Option Contract

An *option* is a contract between the writer and the holder that gives the right, but not the obligation, to the holder to buy or sell a risky asset at a prespecified fixed price within a specified period [142, Chapter 1]. The underlying risky asset could be stocks, stock indices, futures, currencies, commodities, or even weather.

An option contract is a form of derivative instrument, which can be traded on exchanges or over the counter. A *call* (*put*) option allows its holder to buy (sell) the underlying asset at the *strike price* K . Option holders can only exercise their *European-style* options at the expiration or maturity date, T ; in contrast, *American-*

style options can be exercised at any time before they expire.

Purchasing options offers you the ability to position yourself accordingly with your market expectations so as to both profit and protect yourselves with limited risk. The decision as to what type of options to buy depends on whether your outlook for the respective security is positive (bullish) or negative (bearish). If your outlook is positive, buying a call option with lower strike price creates the opportunity to share in the upside potential of a stock without having to risk more than a fraction of its market value. Conversely, if you anticipate downward movement, buying a put option with high strike price will enable you to protect your investment against downside risk without limiting profit potential.

The option premium is the price at which the option contract trades. In return, the writer of the call option is obligated to deliver the underlying security to an option buyer if the call is exercised or buy the underlying security if the put is exercised. The writer keeps the premium whether or not the option is exercised. Then it is natural to ask what is a fair price of an option.

Because options are derivatives, they can be combined with the underlying security to create a risk neutral portfolio (zero risk, zero cost, zero return). Implementing this in practice may be difficult because of “stale” stock prices, large bid/ask spreads, market closures. If stock market prices do not follow a random walk (due, for example, to insider trading) this delta neutral strategy or other model-based strategies may encounter further difficulties. Even for veteran traders using very sophisticated models, option trading is not an easy game to play. Hence, the *option pricing problem* is an important and fundamental financial problem. A good estimation of an option’s theoretical price contributed to the explosion of trading in options.

3.2 Black-Scholes Model

Models of option pricing were very simple and incomplete until 1973 when Black and Scholes [23] published the Black-Scholes pricing model. Their model

gives theoretical values for European put and call options on non-dividend paying stocks.

3.2.1 A Simple Example: American Put Option

To introduce this classical model, we take the pricing problem of an American put option on a non-dividend paying stock as a model problem. In the classical Black-Scholes model, we assume that the price $S(t)$ of the underlying risky asset (e.g., a stock) is described by geometric Brownian motion

$$\frac{dS}{S} = rdt + \sigma dW \quad (3.1)$$

with volatility $\sigma > 0$ and interest rate $r > 0$. When no confusion arises, we will assume the random variables all have dependence in time t and drop the argument t .

Remark 3.1 (Wiener Process) A Brownian motion (name from physics) is often called the Wiener process. A Wiener process W_t is characterized by the following three facts:

- $W_0 = 0$;
- W_t is almost surely continuous;
- The increments $W_{t+\Delta t} - W_t$ satisfies independent normal distribution with mean value 0 and variance Δt for any $t, \Delta t \geq 0$.

The Wiener process is the simplest continuous Lévy process which will be discussed in the next Section.

An *American put option* with strike price K and expiration date T gives the holder the right to sell one asset at *any time* t before the expiration date at price K . At any time t when the option is exercised, its value is given by $P(S(t))$ with the payoff function

$$P(S) = (K - S)_+ = \max\{K - S, 0\}.$$

We want to solve the following problem: If at time t we have an asset priced at $S(t)$,

- What is the fair price $V(S, t)$ of the option?
- When is the optimal time to exercise the option?

Let $S(t)$ denote the underlying stock price and $V(S, t)$ be the American put option price at time t . It is well-known that the price of an American option satisfies the Black-Scholes equation:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S} - rV = 0 \quad \forall S > S_f(t) \text{ and } t \in [0, T], \quad (3.2)$$

where σ is the volatility of the underlying stock, r is the interest rate, and $S_f(t)$ denotes the exercise boundary at time t . We know that the price of an American option is never less than the pay-off function $P(S)$ because of the non-arbitrage assumption¹; therefore

$$V(S, t) = P(S) \quad \forall 0 \leq S \leq S_f(t) \text{ and } t \in [0, T]. \quad (3.3)$$

The final and boundary conditions are given by

$$\begin{cases} V(S, T) = P(S), \quad S \geq 0, \\ V(S_f(t), t) = P(S_f(t)), \quad \frac{\partial V}{\partial S}(S_f(t), t) = -1, \quad 0 < t \leq T, \\ \lim_{S \rightarrow \infty} V(S, t) = 0, \quad 0 \leq t \leq T. \end{cases} \quad (3.4)$$

In this way, we write the price of an American put option as the solution of a free boundary problem (3.2)–(3.4). In Figure 3.1, we see that, for an American put option, when the underlying stock price is greater than the exercise boundary, we should hold the put option; otherwise, early exercise could avoid possible loss. Although this formulation is mathematically beautiful, a major difficulty under this setting is that one needs to solve for V along with the unknown exercise boundary² S_f .

¹It simply means no one can make immediate risk-free profit

²For American option holders, they need to decide whether and when to exercise an option. This leads to an optimal exercise policy problem.

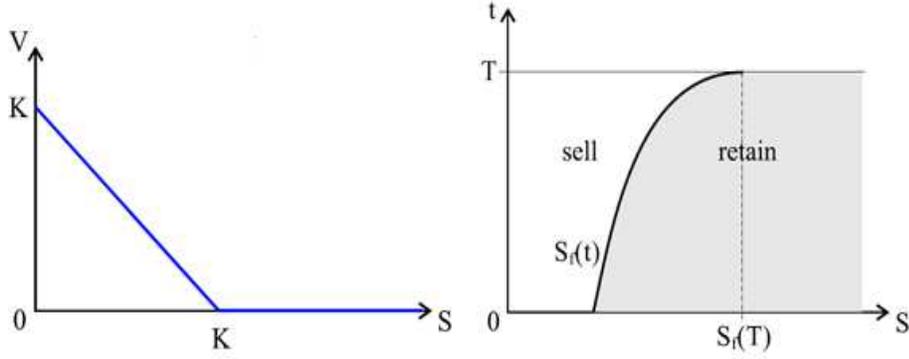


Figure 3.1: Price of American Put Option. Left: pay-off function P ; Right: exercise boundary S_f .

3.2.2 Black-Scholes Inequality

The idea is to reformulate the problem such that the free boundary does not show up explicitly and the degeneracy at the origin is avoid [80, 81]. If we use the time to maturity $\tilde{t} = T - t$ and log price $x = \log S$ as independent variables, then the function

$$u(x, \tilde{t}) := V(e^x, T - \tilde{t})$$

satisfies the following linear complementarity problem LCP (we will write t instead of \tilde{t} for time to maturity from now on):

Problem 3.2 (Black-Scholes Inequality) Find $u(x, t)$ such that

$$\frac{\partial u}{\partial t} - \frac{\sigma^2}{2} \frac{\partial^2 u}{\partial x^2} + \left(\frac{\sigma^2}{2} - r \right) \frac{\partial u}{\partial x} + ru \geq 0 \quad \text{for } x \in \mathbb{R} \text{ and } 0 \leq t \leq T, \quad (3.5)$$

with the obstacle condition

$$u(x, t) \geq \chi(x) \quad \text{for } x \in \mathbb{R} \text{ and } 0 \leq t \leq T \quad (3.6)$$

and the initial condition

$$u(x, 0) = u_0(x) \quad \text{for } x \in \mathbb{R}, \quad (3.7)$$

where $u_0(x) = \chi(x) = P(e^x)$ is the payoff function in the log of the asset price. Moreover, for each point $(x, t) \in \mathbb{R} \times [0, T]$, the complementarity condition has to be satisfied, i.e., there holds equality in at least one of (3.5) and (3.6).

We have shown in §2.3.2 that LCP's can be also written as variational inequalities. So it is clear that Problem 3.2 is a special example of parabolic variational inequalities.

Remark 3.3 (Localization of Domain) To solve problems like Problem 3.2, which is formulated on an infinite domain, we usually truncate the infinite domain to get a finite domain $[-L, L]$ (this procedure is usually called localization). It introduces truncation error which decreases exponentially fast as L increases. On the other hand, the localization also removes the degeneracy (when $S = 0$) artificially. To get around this, there is a different approach which avoids using the log-price has been proposed by [5].

Remark 3.4 (Solving the B-S Problems) Generally speaking, there are two basic ways to solve option pricing problems: analytical methods and numerical methods. Black and Scholes [23] derived explicit pricing formulas for European call and put options on stocks which do not pay dividends. For American options, the Black-Scholes model results in a variational inequality. One can not find explicit closed-form solutions to the American option pricing problem in general. When the formulas for the exact solutions are too difficult to be practically used, we resort to numerical methods, such as lattice methods, simulation-based methods, PDE-based methods, etc. We refer to the book by Wilmott, Dewynne, and Howison [142], the recent review by Broadie and Detemple [37], and the references therein for a review and comparison of many numerical strategies for pricing American options.

Remark 3.5 (Perpetual Options) A perpetual option is an option with no maturity date. Of course, only American-style perpetual options make sense then. For pricing perpetual options in the B-S model, we only need to modify Problem 3.2 by removing the time-derivative term to obtain a steady state variational inequality.

3.3 Beyond Black-Scholes Model

In the classical Black-Scholes (B-S) model, the underlying risky assets are assumed to be geometric Brownian motions. In practice, all the parameters (strike price, expiration date, interest rate, etc) can be observed except the volatility. This implies an one-to-one relation between the value of an option contract and the volatility. However it is observed in “real” world that it is necessary to use different volatility for different strike price or maturity to fit the Black-Scholes formula with quoted prices of European options. This phenomenon is called *volatility skew* or *volatility smile* depending on the shape of the volatility curve. Because of the existence of the volatility smile, traders usually need to use a matrix of implied volatilities [141] to adjust prices.

3.3.1 Lévy Processes

Many advanced models beyond the classical B-S models have been proposed to overcome this difficulty. We only mention one of the approaches, which enriches the stochastic dynamics of the underlying risk asset by allowing jumps (see [4] and the reference therein for a quick review). These models can be treated in a general framework using Lévy processes. In real life, it is observed that the price of a risky asset could have sudden jumps. For example, in Figure 3.2 and 3.3, it shows the exchange rate of US dollars to Euro from the beginning of century till now. We can see jumps if we examine the picture carefully.

Starting from the seminal work by Merton [102], many models were developed along this direction in the last two decades. The variance Gamma model by Madan and Seneta [95] was the first model which used a particular Lévy process to model the asset dynamics. It was extended to option pricing later by Madan et al. [94]. All these models as well as the classical B-S model can be considered in the framework of Lévy processes [91]. In this section, we shall first review some basic concepts of Lévy processes.

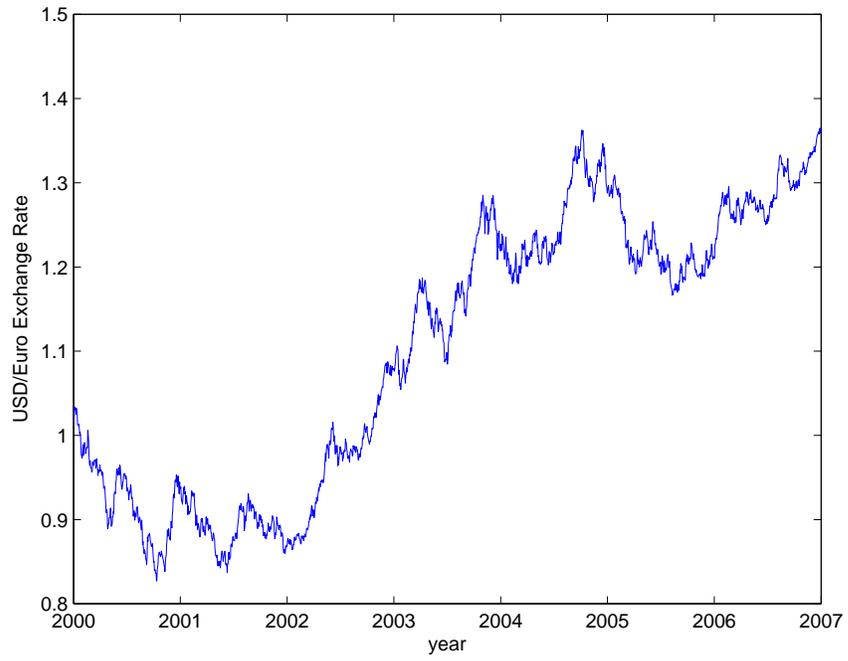


Figure 3.2: Foreign exchange rate: US dollars per Euro.

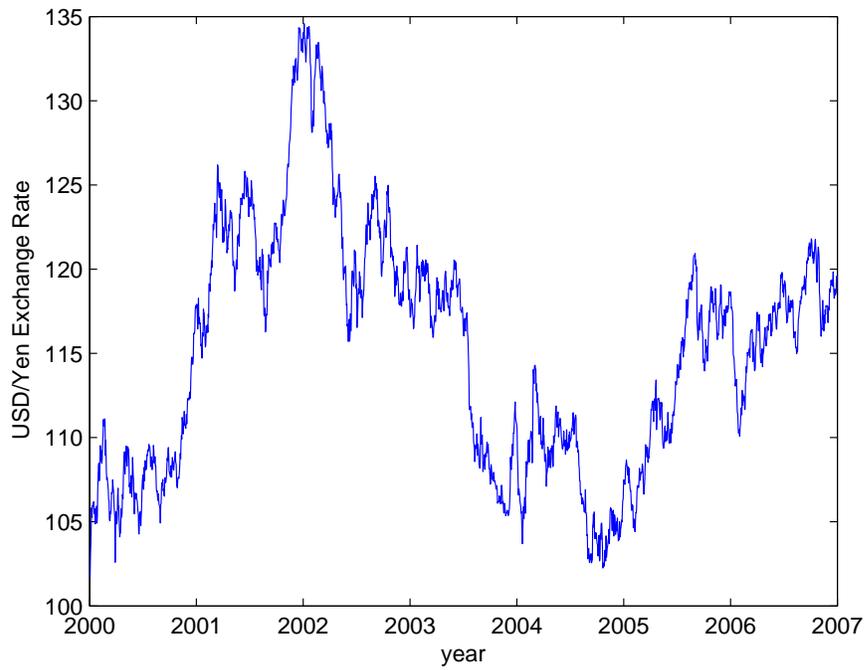


Figure 3.3: Foreign exchange rate: Yen per US dollars.

Definition 3.6 (Lévy Process) *A stochastic process, X_t ($0 < t < \infty$ and $X_0 = 0$), is a Lévy process if and only if it has independent and stationary increments.*

Remark 3.7 (Independent and Stationary Increments) By the definition, for any Lévy process X_t , the random variable $X_{t+\Delta t} - X_t$ has same but independent distribution as the $X_{t'+\Delta t} - X_{t'}$ with $0 < t, t', \Delta t < \infty$. It is then clear that the Wiener process introduced in Remark 3.1 is a particular example of Lévy processes.

Example 3.8 (Poisson Process) In addition to a Wiener process, another simple example of Lévy processes is a Poisson process. The Poisson process $N_t (t \geq 0)$ represents the number of events since time $t = 0$ and increment $N_{t+\Delta t} - N_t$ satisfies a Poisson distribution for any t and $\Delta t \geq 0$. Merton [102] used Poisson processes to model the occurrence of jumps in real market

$$\frac{dS}{S} = rdt + \sigma dW + \eta dN.$$

It is often called the jump-diffusion model.

3.3.2 Lévy-Khintchine Formula

The characteristic function of a Lévy process can be represented using the following Lévy-Khintchine formula (detailed discussion can be found in the monograph by Sato [121]).

Proposition 3.9 (Lévy-Khintchine Formula) *Let X_t be a Lévy process. Then we have the following representation of the characteristic function of X_t*

$$\ln \mathbf{E}[e^{i\theta X_t}] = i\alpha t\theta - \frac{1}{2}\sigma^2 t\theta^2 + t \int_{\mathbb{R}} (e^{i\theta x} - 1 - i\theta x \mathbf{1}_{|x|<1}) \nu(dx).$$

where $\alpha \geq 0$, $\sigma \in \mathbb{R}$, and $\mathbf{1}_{|x|<1}$ is the characteristic function and a measure ν on $\mathbb{R} \setminus \{0\}$ satisfying

$$\int_{\mathbb{R}} \min\{1, x^2\} \nu(dx) < \infty.$$

Remark 3.10 (Lévy-Khintchine Triplet) From the proposition above, a Lévy process is a combination of a drift component, a Brownian motion component and

a jump component. These three components can be determined by the Lévy-Khintchine triplet (α, σ^2, ν) .

- The first parameter α is called the drift term which determines the development of the process X_t on the average.
- The second parameter σ^2 defines the variance of the Gaussian part of X_t .
- The last parameter ν (the so-called Lévy measure) is responsible for the behavior of jumps. It is usually assumed that $\nu(dx) = k(x)dx$ with $k(\cdot)$ being the Lévy density of X_t . Intuitively speaking, the Lévy measure describes the expected number of jumps of a certain height in a unit time interval.

Remark 3.11 (Regularization) We notice that the Lévy density might not be integrable near the origin. Regularization is necessary to make the integral in the Lévy-Khintchine formula integrable. The function $1 + i\theta x \mathbf{1}_{|x| < 1}$ is used for regularization (to guarantee integrability around zero) here.

Remark 3.12 (CGMY Model) The CGMY model [40] is a generalization of the variance Gamma model [95]. Here we just give the Lévy density of the CGMY model without getting into details. The density function can be written as

$$k_{CGMY}(x) := \begin{cases} C \frac{\exp(-G|x|)}{|x|^{1+Y}} & \text{if } x < 0 \\ C \frac{\exp(-M|x|)}{|x|^{1+Y}} & \text{if } x > 0, \end{cases} \quad (3.8)$$

where constants $C > 0$, $G, M \geq 0$, and $Y < 2$. Here C is a measure of the overall level of activity; G and M control the rate of exponential decay of the Lévy density (they are usually different due to different reasons causing up and down movement of the price of risk assets); Y is used to model the fine structure of the stochastic process.

Remark 3.13 (Relation with Fractional Laplacian) It is well known that the Fourier transform of the Laplace operator can be written as

$$(-\Delta u)^\wedge(\xi) = |\xi|^2 \hat{u}(\xi).$$

In this manner, we can define square root of the Laplace operator to be

$$((-\Delta)^{\frac{1}{2}}u)^\wedge(\xi) := |\xi|\hat{u}(\xi).$$

More generally, we can define [55], for all $s \in \mathbb{R}_+$, that

$$((-\Delta)^s u)^\wedge(\xi) := |\xi|^{2s}\hat{u}(\xi). \quad (3.9)$$

This is related to the so-called fractional integral operator. In fact, we can compute the fractional Laplacian $(-\Delta)^s$ using a singular integral

$$(-\Delta)^s u(x) = C_{d,s} \cdot \mathbf{PV} \int_{\mathbb{R}^d} \frac{u(x) - u(y)}{|x - y|^{d+2s}} dy, \quad (3.10)$$

This integral operator is then related to the CGMY model ($G = M = 0$, $Y = s$ for $d = 1$).

Using similar techniques as in the case of Black-Scholes, it has been shown (see [118]) that value of options written on an underlying geometric Lévy process can be formulated as integro-differential equations (European-style) or variational inequalities (American-style) [4]. In the following section, we will give a general formulation of a class of integro-differential variational inequalities which can cover the important cases of European and American option pricing problems with Lévy asset.

3.4 Option Pricing as a Variational Inequality

In this section, we shall specify a class of problems which will be treated numerically in the following chapters. We shall introduce fully-discrete numerical methods to solve the problem in Chapter 4; we analyze the a priori as well as a posteriori errors of the numerical methods in Chapter 5 and 6; finally we shall propose adaptive algorithms to improve efficiency in Chapter 7.

Assume the linear operator $\mathcal{A} : \mathcal{V} \rightarrow \mathcal{V}^*$ to be continuous and coercive and $a(\cdot, \cdot)$ to be its associated bilinear form. To cover the interesting applications men-

tioned in the previous two sections, we consider the following evolution integro-differential variational inequalities: find $u(t) \in \mathcal{K}(t)$ such that

$$\langle \partial_t u(t) + \mathcal{A}u(t) - f(t), u(t) - v \rangle \leq 0 \quad \forall v \in \mathcal{K}(t) \text{ a.e. } t \in (0, T), \quad (3.11)$$

where the convex set

$$\mathcal{K}(t) := \{v \in \mathcal{V} \mid v \geq \chi(t)\}.$$

Here f, u, v are obviously also functions of x , which we omit for convenience.

Now we shall introduce a general variational inequality problem which can be used for American option pricing problems on assets whose prices are modelled by a Lévy process. Let Ω be an open and bounded polygonal domain in \mathbb{R}^d and $\mathcal{Q} := \Omega \times (0, T)$. For a real constant $Y < 2$, we define a continuous pseudo-differential operator $\mathcal{A}_I: \tilde{H}^{Y/2}(\Omega) \rightarrow H^{-Y/2}(\Omega)$

$$\mathcal{A}_I u(x) := \int_{\Omega} k(x-y)u(y) dy \quad \forall u \in \tilde{H}^{Y/2}(\Omega), \quad (3.12)$$

where $k(x)$ is a given kernel function. We assume that, in the definition (3.12), the kernel function $k(x) \in C^\infty(\mathbb{R} \setminus \{0\})$, and that the condition

$$|\partial_x^m k(x)| \lesssim |x|^{-d-Y-m} \quad (3.13)$$

near $x = 0$.

Remark 3.14 (More General Pseudo-differential Operators) For financial applications considered in this thesis, the pseudo-differential operator \mathcal{A}_I (3.12) is general enough to cover most important models, like Lévy jump-diffusion models and the CGMY model. However, the theory, which will be developed in the following chapters, can be extended to more general classes of operators. For example, we can allow operators which are not transition invariant, i.e. $\mathcal{A}_I u(x) = \int_{\Omega} k(x, y)u(y)dy$, also. In differential operator case, operator \mathcal{A} with coefficients depends on x are considered in [104].

Remark 3.15 (Singular Kernel) Since we could and would like to (to allow jumps) have singular kernel as discussed in previous section, we need to give the integral operator in (3.12) a proper meaning. Taking the kernel function as in CGMY model as an example, i.e. $k(x) = \frac{e^{-C|x|}}{|x|^{1+Y}}$, we usually consider the following cases:

1. $\int_{\mathbb{R}} k(x)dx < \infty$ or $Y < 0$: In this case, the integral is not singular and the corresponding underlying asset has finite activity and finite variation.
2. $\int_{\mathbb{R}} xk(x)dx < \infty$ or $0 \leq Y < 1$: In this case, the integral need to be regularized by $\int_{\mathbb{R}} k(x-y)(u(y) - u(x))dy$. This corresponds to the case when the underlying asset has infinite activity but finite variation.
3. $\int_{\mathbb{R}} x^2k(x)dx < \infty$ or $1 \leq Y < 2$: In this case, the kernel function is more singular; the underlying asset could have infinite activity and infinite variation. We could regularize the integral by $\int_{\mathbb{R}} k(x-y)(u(y) - u(x) - (e^{y-x} - 1)u'(x))dy$ for example.

Let $\rho \in (0, 2]$ be a positive constant. We define $\mathcal{V} := \tilde{H}^{\rho/2}(\Omega)$. We consider the following class of linear operators.

Definition 3.16 (Operator \mathcal{A}) Define $\mathcal{A}: \tilde{H}^{\rho/2}(\Omega) \rightarrow H^{-\rho/2}(\Omega)$ in the following three class where coefficients $c_2 \in \mathbb{R}^{d \times d}$, $0 \leq c_I \in \mathbb{R}$, $c_1 \in \mathbb{R}^d$, $c_0 \in \mathbb{R}$ are constants:

- Case I ($\rho = 2$): In this case $Y < 2$

$$\mathcal{A}u := -\nabla \cdot (c_2 \nabla u) + c_I \mathcal{A}_I u + c_1 \cdot \nabla u + c_0 u,$$

where $c_2 \in \mathbb{R}^{d \times d}$ is a positive definite matrix.

- Case II ($1 \leq \rho < 2$): In this case $Y = \rho$ and

$$\mathcal{A}u := c_I \mathcal{A}_I u + c_1 \cdot \nabla u + c_0 u,$$

where \mathcal{A}_I satisfies the Gårding inequality:

$$\langle \mathcal{A}_I v, v \rangle \geq \kappa_\rho \|v\|_{\tilde{H}^{\rho/2}}^2 - \kappa_\sigma \|v\|_{\tilde{H}^\sigma(\Omega)}^2 \quad (3.14)$$

with $\kappa_\rho > 0$ and $\sigma < \rho/2$.

- Case III ($0 < \rho < 1$): In this case $Y = \rho$ and

$$\mathcal{A}u := c_I \mathcal{A}_I u + c_0 u,$$

where \mathcal{A}_I satisfies the Gårding inequality (3.14).

From now on, we define $s = \rho/2$ and the operator $\mathcal{A} : \tilde{H}^s(\Omega) \rightarrow H^{-s}(\Omega)$. We note that $0 < s \leq 1$ depends on the specific application.

Remark 3.17 (Financial Meaning) For a Levy process, c_2 corresponds to the covariance matrix of a Brownian motion; the integral operator \mathcal{A}_I corresponds to a jump process; the term with c_1 is necessary to achieve the Martingale condition.

Remark 3.18 (Continuity and Coercivity) In all these three cases, we can see that (1.7) always holds and (1.8) is satisfied if c_0 is sufficiently large. Hence the existence and uniqueness of the solution can be proved by the general theory introduced in Chapter 1 and 2. Furthermore, the energy norm associated with \mathcal{A} , $\|\cdot\|$, is equivalent to the $\tilde{H}^s(\Omega)$ -norm.

Remark 3.19 (Strong Sector Condition) From continuity and coercivity of \mathcal{A} , it is then clear that the operator \mathcal{A} satisfies the strong sector condition (2.5), i.e.

$$|\langle \mathcal{A}v, w \rangle| \leq \|\mathcal{A}v\|_{\mathcal{V}^*} \cdot \|w\|_{\mathcal{V}} \leq 2\gamma \|v\| \cdot \|w\| \quad v, w \in \mathcal{V},$$

where $\gamma = \frac{C^*}{2C_*}$ and $\mathcal{V} = \tilde{H}^s(\Omega)$. Hence, by Lemma 2.10, \mathcal{A} is an angle-bounded operator which satisfies the coercivity condition (2.9).

Remark 3.20 (Smooth Pasting) Regularity results for obstacle problems with fractional power of Laplacian are discussed by Silvestre [128]. He proved that the solution u is in $C^{1,s}(\Omega)$ for time-independent obstacle problem with $\mathcal{A} = (-\Delta)^s$ ($0 < s \leq 1$). For more general problems, it has been shown by Boyarchenko and Levendorskii [24] that (for perpetual American options) the smooth pasting property (C^1 solution) may fail in the pure jump cases (for example, $c_2 = 0$, $c_I = 1$, $c_1 \neq 0$, and $Y < 1$). So in general, we can not assume that smooth pasting holds for our numerical treatments.

Chapter 4

Numerical Methods for Obstacle Problems

Many numerical schemes have been developed and analyzed for variational inequalities in the past three decades. The standard techniques for both static and evolution variational inequalities can be found in the book by Glowinski [75, 74]. For option pricing problem, several numerical algorithms [3, 76, 77, 98, 99, 100] have been proposed recently.

In this chapter we discuss numerical methods for the obstacle problems of general form discussed in §2. We review the finite element method for space-discretization and the general θ -scheme for time-discretization. And then we discuss a fully-discrete numerical scheme for PVI to prepare ourselves for later chapters on error analysis and adaptive methods. The rest of this chapter is organized as follows. In section 4.1, we review basic concepts of continuous Galerkin method and finite element approximation. Then we introduce the general θ -scheme which is commonly used for evolution equations in section 4.2. Finally, we give a fully-discrete numerical scheme for parabolic variational inequalities in 4.3.

4.1 Finite Element Methods

The finite element method (FEM) has a long history in practical use and is widely applied to lots of problems in physics and engineering. It has been proved to be very successful in many areas, like structural mechanics. After forty years extensive development, the subject of standard finite element method has become

a well-understood and successful area in scientific computation.

Remark 4.1 (Why FEM?) The reason we choose to use the finite element method instead of the finite difference method is due to the following features of the finite element method:

- The most attractive feature of the FEM is its ability to handle complex geometries, boundaries, and operators with relative ease. Since we are going to handle differential and integral operators in a uniform framework, it is much easier to use the finite element method.
- The finite difference method approximates the differential equation whereas the finite element method approximates the underlying function space. It is more natural to enforce the obstacle constraint in the finite element approximation.
- The finite element method provides a mathematically sound framework for deriving a priori and a posteriori error estimates along with adaptive algorithms.

For elliptic partial differential equations, the *Galerkin method* exploit the weak formulation and replaces the underlying function space by an appropriate finite dimensional subspace. And FEM is a Galerkin method that uses piecewise polynomial spaces for approximate test and trial function spaces. The readers are referred to [50, 83, 25, 29] for more detailed discussion on construction and error analysis of the standard finite element method. This idea can be naturally extended to elliptic variational inequalities.

To explain the main idea, we first introduce the finite element method for the following elliptic variational inequality as an example:

Problem 4.2 (Elliptic Variational Inequality) *Let $0 < s \leq 1$ and the elliptic operator $\mathcal{A} : \tilde{H}^s(\Omega) \rightarrow H^{-s}(\Omega)$. Given data $f \in L^2(\Omega)$ and a closed convex set \mathcal{K} , find $u \in \mathcal{K}$ such that*

$$\langle \mathcal{A}u, u - v \rangle \leq \langle f, u - v \rangle \quad \forall v \in \mathcal{K} := \{v \in \tilde{H}^s(\Omega) \mid v \geq \chi\}. \quad (4.1)$$

For the definition of Sobolev spaces, see §2.1.

Step 1. Domain Partition

We first partition the domain into small subdomains. Let Ω be an open polygonal domain. We then discretize the polygonal domain Ω into simplexes $\tau \in \mathcal{T}$. Let $h_\tau = |\tau|^{\frac{1}{d}}$ be the diameter of $\tau \in \mathcal{T}$ and $h(x)$ be the local meshsize, that is the piecewise constant function with $h|_\tau := h_\tau$ for all $\tau \in \mathcal{T}$. The collection \mathcal{T} of elements (triangles or tetrahedrons) is called a mesh or triangulation. Throughout this work, we will only consider *conforming* meshes, i.e. the intersection of any two elements in \mathcal{T} is either an edge(2d)/face(3d), vertex, or empty (see Figure 4.1 for an example). We denote by $\mathcal{P}_h(\mathcal{T})$ the set of all nodes in the mesh \mathcal{T} . Here we use the subscript h to describe the discrete nature and this does not imply the underlying meshes are quasi-uniform with meshsize h . Given a node $z \in \mathcal{P}_h(\mathcal{T})$, we define the local meshsize to be $h_z := \max\{h_\tau : \tau \in \mathcal{T} \text{ and } z \in \tau\}$. Let $h_{\min} := \min_{z \in \mathcal{P}_h(\mathcal{T})} h_z$ to be the minimum meshsize of \mathcal{T} .

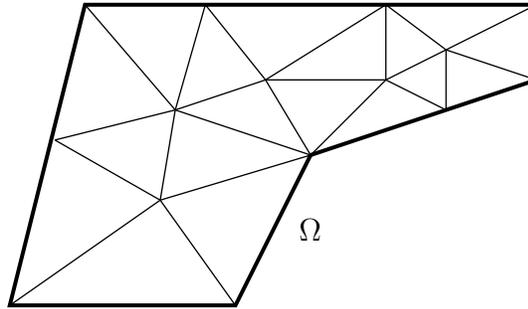


Figure 4.1: A Conforming Partition of Ω

Step 2. Finite-dimensional Approximation

Let $\mathbb{V}(\mathcal{T}) \subset \tilde{H}^s(\Omega)$ be the space of continuous piecewise polynomial finite element functions over the mesh \mathcal{T} which vanishes on the boundary $\Gamma := \bar{\Omega} \setminus \Omega$, i.e.

$$\mathbb{V}(\mathcal{T}) := \{v \in C(\bar{\Omega}) : v|_\tau \text{ is a polynomial for all } \tau \in \mathcal{T}, v = 0 \text{ on } \Gamma\}. \quad (4.2)$$

We then use a finite-dimensional set

$$\mathbb{K} := \{u_h \in \mathbb{V}(\mathcal{T}) \mid u_h \geq \chi_h\} \quad \text{closed and convex}$$

to approximate the feasible set \mathbb{K} , where χ_h is an approximation of the obstacle χ . We notice that there are many ways to approximate the continuous test function space. Different choices will result in different numerical methods. In this note, we shall focus on the simplest case – linear finite element method on triangles or tetrahedrons, i.e. $v|_\tau$ is a linear polynomial on each $\tau \in \mathcal{T}$. The weak form of the finite element approximation reads

$$\text{Find } u_h \in \mathbb{K} : \quad a(u_h, u_h - v_h) \leq \langle f, u_h - v_h \rangle \quad \forall v_h \in \mathbb{K}. \quad (4.3)$$

Step 3. Solving the Finite-dimensional Problem

Let polynomials $\{\psi_i\}_{i=1}^I$ be a basis of the I -dimensional space $\mathbb{V}(\mathcal{T})$. Let $\mathbf{A} := (a(\psi_i, \psi_j))_{i,j=1}^I$ be the resulting stiffness matrix of (4.1). If $\vec{\mathbf{U}} = (\mathbf{U}_i)_{i=1}^I$, $\vec{\mathbf{X}} = (\mathbf{X}_i)_{i=1}^I \in \mathbb{R}^I$ are the vectors of coefficients of u_h and χ_h , namely $u_h = \sum_{i=1}^I \mathbf{U}_i \psi_i$ and $\chi_h = \sum_{i=1}^I \mathbf{X}_i \psi_i$, and $\vec{\mathbf{F}} = (\mathbf{F}_i)_{i=1}^I := (\langle f, \psi_i \rangle)_{i=1}^I$, then $\vec{\mathbf{U}}$ satisfies the finite-dimensional variational inequality:

$$\text{Find } \vec{\mathbf{U}} \geq \vec{\mathbf{X}} : \quad (\mathbf{A}\vec{\mathbf{U}} - \vec{\mathbf{F}})^T (\vec{\mathbf{U}} - \vec{\mathbf{V}}) \leq 0 \quad \forall \vec{\mathbf{V}} \geq \vec{\mathbf{X}}.$$

Upon solving this finite-dimensional problem, we obtain a discrete approximation u_h of Problem (4.2). It is clear that this discrete problem admits a unique solution (see, for example, [74]). There are various ways to solve this finite-dimensional variational inequality. For the moment, we assume that there is a magic black box which can give us the solution of this problem. Once this discrete VI problem is solved, we get an approximation of the exact solution.

Remark 4.3 (Approximation of χ) There are several ways to approximate the convex set \mathcal{K} . For example, we can take $\chi_h = \chi$ (conforming, i.e. $\mathbb{K} \subset \mathcal{K}$, but not practical) or take χ_h to be the Lagrange interpolant of χ for continuous χ (might not be conforming).

Error Estimations

For standard finite element approximation of elliptic equations, the most important property is an orthogonality property (i.e. the so-called *Galerkin orthogonality*).

nality)

$$a(u - u_h, v_h) = 0 \quad \forall v_h \in \mathbb{V}. \quad (4.4)$$

This is a simple observation of the weak formulations of the exact and discrete solutions:

$$\begin{cases} a(u, v) = \langle f, v \rangle & \forall v \in \mathcal{V} \\ a(u_h, v_h) = \langle f, v_h \rangle & \forall v_h \in \mathbb{V}. \end{cases}$$

Taking $v = v_h$ in the first equation and simply subtracting the two equations gives the Galerkin orthogonality (4.4).

A Priori Error Analysis. Using the definition of the energy norm, the Galerkin orthogonality (4.4) and the strong sector condition (2.5), we have, for any $v_h \in \mathbb{V}$, that

$$\|u - u_h\|^2 = a(u - u_h, u - u_h) = a(u - u_h, u - v_h) \lesssim \|u - u_h\| \|u - v_h\|.$$

Hence, we obtain the quasi-optimality of the finite element approximation

$$\|u - u_h\| \lesssim \inf_{v_h \in \mathbb{V}} \|u - v_h\|. \quad (4.5)$$

This means u_h is almost the best approximation of u in the subspace \mathbb{V} . We shall discuss this in Chapter 5 in detail.

A Posteriori Error Analysis. A posteriori error estimation relies on the following error equation (or residual equation). It is straightforward that

$$a(u - u_h, v) = a(u, v) - a(u_h, v) = \langle f, v \rangle - a(u_h, v) = \langle f - \mathcal{A}u_h, v \rangle.$$

Hence, by continuity and coercivity of \mathcal{A} and the Cauchy-Schwarz inequality, we obtain

$$\|f - \mathcal{A}u_h\|_* \lesssim \|u - u_h\| \lesssim \|f - \mathcal{A}u_h\|_*. \quad (4.6)$$

Notice that, on the right-hand side, we only have the data f and the discrete solution u_h . This upper bound does not depend on the unknown solution u . Of course, to make the upper bound useful in adaptive algorithms, we need it to be local and computable. This will be addressed later in Chapter 6 and 7.

Some Comments

We now summarize this short introduction of finite element methods with a few comments.

Remark 4.4 (Quasi-uniform Meshes) Let \mathcal{T} be a mesh over the domain Ω . We say that \mathcal{T} is *quasi-uniform* if there exists a constant h_0 independent of τ such that

$$h_0 \lesssim h_\tau \lesssim h_0 \quad \forall \tau \in \mathcal{T}.$$

Remark 4.5 (Shape-regularity) Let $\{\mathcal{T}_j\}$ be a family of conforming meshes over the domain Ω . We refer $\{\mathcal{T}_j\}$ as a *shape-regular* family if there exists a generic constant C independent of j such that

$$\frac{\text{diam}(\tau)}{h_\tau} \leq C \quad \forall \tau \in \bigcup_j \mathcal{T}_j,$$

where $\text{diam}(\tau)$ is the diameter of the smallest ball containing τ . Notice that the shape-regular family allows meshes that may be very highly locally refined (containing elements of very different sizes). This condition is equivalent to the *maximum angle condition* which is crucial for standard finite element analysis [7].

Remark 4.6 (Higher-order Finite Element Spaces) The test function space $\mathbb{V}(\mathcal{T})$ does not necessarily have to be a piecewise linear polynomial space. It could contain high-order polynomials to achieve better approximability. By choosing different trial function spaces and different convex sets \mathbb{K} , one can construct different finite element methods.

4.2 Euler Method for ODEs

Before we can introduce a fully-discrete numerical method, we review a simple time discretization scheme, the Euler method, for the Cauchy problem (initial value problem): find $u : [0, T] \rightarrow \mathbb{R}$ satisfying

$$\begin{cases} u'(t) + \mathfrak{F}(t, u(t)) = 0 & \forall t \in (0, T) \\ u(0) = u_0 \end{cases} \quad (4.7)$$

where $\mathfrak{F}(t, \cdot), t \in [0, T]$ be a family of continuous and coercive operators from \mathcal{V} to \mathcal{V}' .

We partition the time domain $[0, T]$ into N subintervals, i.e. $0 = t_0 < t_1 < \dots < t_N = T$ and let $k_n := t_n - t_{n-1}$ be the time step-size. We denote the approximate solution at each time step t_n by U^n , for $n = 0, \dots, N$.

For any sequence $\{W^n\}_{n=1}^N$, we define the piecewise constant interpolant \overline{W} and the piecewise linear interpolant W as

$$\overline{W}(t) := W^n, \quad W(t) := l(t)W^{n-1} + (1 - l(t))W^n \quad \forall t \in (t_{n-1}, t_n], \quad (4.8)$$

for $1 \leq n \leq N$, where the linear function $l(t)$ is defined by

$$l(t) := \frac{t_n - t}{k_n} \quad \forall t \in (t_{n-1}, t_n]. \quad (4.9)$$

We also denote by $\{\delta W^n\}_{n=1}^N$ the discrete derivative of the sequence $\{W^n\}_{n=1}^N$

$$\delta W^n := \frac{W^n - W^{n-1}}{k_n} \quad \forall 1 \leq n \leq N. \quad (4.10)$$

Since W is piecewise linear in time, we denote $\delta_t W$ to be the left derivative of W in time. From this definition, it is easy to see that

$$\delta_t W(t) = \delta W^n \quad \forall t \in (t_{n-1}, t_n]. \quad (4.11)$$

For a function w continuous in time, we let $W^n(\cdot) := w(t_n, \cdot)$ be its semi-discrete approximation. Hence, by the convention above, \overline{W} is the piecewise constant approximation of w and W is the piecewise linear interpolation (in time) of w .

Now we are ready to formulate the θ -scheme: given an initial guess U^0 of u_0 , solve the following discrete problem

$$\delta U^n + \theta \mathfrak{F}(t_n, U^n) + (1 - \theta) \mathfrak{F}(t_{n-1}, U^{n-1}) = 0, \quad (4.12)$$

for $n = 1, \dots, N$ and $0 \leq \theta \leq 1$. For different θ , we get different finite difference schemes:

- Forward Euler Method : $\theta = 0$ (explicit scheme)

- Backward Euler's Method : $\theta = 1$ (implicit scheme)
- Trapezoidal Method: $\theta = \frac{1}{2}$ (Crank-Nicolson Method).

The convergence, stability and consistency results for these methods are standard (see, for example, [6, Chapter 5]).

4.3 Numerical Methods for Parabolic VI

With the two basic building blocks introduced in §4.1 and §4.2, we can now introduce a class of fully-discrete numerical methods for the parabolic obstacle problem (2.18). We first recall the continuous problem and then give a fully-discrete numerical scheme to solve it.

4.3.1 Continuous Problem

To simplify the representation, we assume that Ω be an open bounded polygonal domain in \mathbb{R}^d with boundary Γ and $\mathcal{Q} := \Omega \times (0, T)$ be the parabolic cylinder. Consider an obstacle $\chi \in H^1(\mathcal{Q})$ such that $\chi \leq 0$ on $\Gamma \times (0, T)$ and nonempty convex sets

$$\mathcal{K}(t) := \{v \in \tilde{H}^s(\Omega) : v \geq \chi(t)\} \quad \text{a.e. } t \in [0, T]. \quad (4.13)$$

We consider the linear operator $\mathcal{A} : \tilde{H}^s(\Omega) \rightarrow H^{-s}(\Omega)$ for $0 < s \leq 1$ given in Definition 3.16. The operator \mathcal{A} gives rise to the continuous and coercive bilinear form $a(\cdot, \cdot) : [\tilde{H}^s(\Omega)]^2 \rightarrow \mathbb{R}$ defined by

$$a(v, w) := \langle \mathcal{A}v, w \rangle \quad \forall v, w \in \tilde{H}^s(\Omega).$$

For the moment, we further assume that $\chi \in C(0, T; H^1(\Omega) \cap C(\Omega))$. We can use linear Lagrange interpolation χ_h^n to approximate $\chi(t_n)$. Instead of using the interpolation to define the approximate obstacle, we can also employ an operator based on averaging. This will be discussed in Chapter 5. Hence this restriction will be removed later.

Problem 4.7 Given data $f \in L^1(0, T; L^2(\Omega))$ and initial condition $u_0 \in \mathcal{K}$, find $u \in L^2(0, T; \mathcal{K}) \cap H^1(0, T; H^{-s}(\Omega))$ such that

$$\langle \partial_t u(t) + \mathcal{A}u(t), u(t) - v \rangle \leq \langle f(t), u(t) - v \rangle \quad \forall v \in \mathcal{K}(t) \text{ a.e. } t \in (0, T). \quad (4.14)$$

4.3.2 Semi-discrete Problem

We can apply the backward Euler method to parabolic variational inequality (4.14) to get a semi-discrete numerical scheme:

Method 4.8 (Backward Euler Method) Given the initial guess $U^0 = u_0$ and

$$F^n := \frac{1}{k_n} \int_{t_{n-1}}^{t_n} f(t) dt, \quad (4.15)$$

find an approximate solution $U^n \in \mathcal{K}$ for $1 \leq n \leq N$ such that

$$\langle \delta U^n, U^n - v \rangle + a(U^n, U^n - v) \leq \langle F^n, U^n - v \rangle \quad \forall v \in \mathcal{K}. \quad (4.16)$$

Remark 4.9 (Implicit Scheme) The backward Euler method, Method 4.8, is fully implicit. At each time step n , we need to solve an elliptic variational inequality

$$\langle U^n, U^n - v \rangle + k_n a(U^n, U^n - v) \leq \langle U^{n-1} + k_n F^n, U^n - v \rangle \quad \forall v \in \mathcal{K}.$$

This problem has a unique solution from Theorem 1.5. We can apply the finite element method discussed in §4.1 to solve it at each time step once the initial guess U^0 is given.

We now recall some convergence results of the semi-discrete solution of Method 4.8. These results will be useful when we discuss the a priori error estimate for fully-discrete problems in Chapter 5. The following lemma is first proved by Biaocchi [11, Theorem 2.1] and then generalized and improved by Savaré [122, Theorem 4] and gives the regularity of the semi-discrete solution as well as its first time derivative.

Lemma 4.10 (Regularity of Semi-discrete Solution) For any initial guess $U^0 \in \mathcal{V}'$, the temporal semi-discrete problem (4.16) admits a unique solution $\{U^n\}$ and

$U^n \in \mathcal{K}$, for $1 \leq n \leq N$. If $U^0 = u_0 \in \mathcal{K}$ and $f \in \mathfrak{S}(0, T)$, then we have the piecewise linear (in time) function

$$U \in \mathfrak{J}(0, T).$$

Furthermore, if $f \in BV(0, T; \mathcal{H})$, we have that

$$\partial_t U \in \mathfrak{J}(0, T)$$

and there exists a constant C depends on f and u_0 such that

$$\|u - U\|_{\mathfrak{J}(0, T)} \leq Ck.$$

Remark 4.11 (Comments on Regularity) As discussed in Remark 2.21, we can not expect $\partial_t u$ to be continuous even if data is sufficiently smooth. Under this consideration, $\partial_t U \in \mathfrak{J}(0, T)$ is *almost* the maximal regularity one can ask; maximal regularity of u is explored in [122]. Using Proposition 2.14, we observe that $\mathcal{A}U$ is in $L^\infty(0, T; \mathcal{H})$ because $f \in BV(0, T; \mathcal{H})$ and $\partial_t U \in L^\infty(0, T; \mathcal{H})$.

Next we recall the following convergence rate for backward Euler method in [110], which is optimal respect to the time stepping method and the regularity of the solution. In this work, Nochetto et al. exploit the angle-bounded condition without assuming further regularity of the solution to prove the optimal convergence rate via a novel *a posteriori* error estimator. This result is consistent with Lemma 4.10.

Lemma 4.12 (Error Estimation for Semi-discrete Solution) *Let the operator \mathcal{A} be γ -angle-bounded. If*

$$U^0 = u_0 \in \{v \in \mathcal{K} \mid \mathcal{A}v \in \mathcal{H}\} \text{ and } f \in BV(0, T; \mathcal{H}),$$

then we have the error

$$\max \left\{ \max_{0 \leq t \leq T} \|u - U\|, \left(\int_0^T \| \|u - U\|^2 dt \right)^{\frac{1}{2}}, \left(\int_0^T \| \|u - \bar{U}\|^2 dt \right)^{\frac{1}{2}} \right\} \leq Ck,$$

where the constant C depends on γ , u_0 , and f only.

Proof. The result is a direct consequence of [110, Corollary 4.10]. ■

4.3.3 Fully-discrete Problem

We can solve Problem 4.7 numerically by a θ -scheme for time-discretization and a conforming finite element method for space-discretization. Apparently, there are many possible combinations in this class. We will focus on one of the simplest combinations: backward Euler and linear finite element method. In the next two chapters, we shall consider the error committed by this particular fully-discrete numerical scheme.

Discretization

For the numerical treatment of Problem 4.7, we discretize the spatial domain Ω into simplexes $\tau \in \mathcal{T}$, and partition the time domain $[0, T]$ into N subintervals, i.e. $0 = t_0 < t_1 < \dots < t_N = T$ and let $k_n := t_n - t_{n-1}$.

Let $\mathbb{V}(\mathcal{T})$ be the usual conforming piecewise linear finite element subspace of $\tilde{H}^s(\Omega)$ over the mesh \mathcal{T} . For the moment, we assume that the finite element space does not change in time. We shall consider the case of mesh changes in time in Chapter 7.

Consider the corresponding discrete convex set at time $t = t_n$

$$\mathbb{K}^n := \{v \in \mathbb{V}(\mathcal{T}) : v \geq \chi_h^n\} \quad (4.17)$$

where the sequence $\chi_h^n \in \mathbb{V}(\mathcal{T})$ is a piecewise linear approximation of the obstacle $\chi(t_n)$ for $0 \leq n \leq N$. For example, when the obstacle χ is continuous, we could take χ_h^n to be the piecewise linear Lagrange interpolant of $\chi(t_n)$. For convenience, we denote the set of space-time piecewise linear functions which satisfies the discrete constraints all the time as

$$\mathbb{K} := \{V \mid V(t_n) \in \mathbb{K}^n \text{ and } V(t) \text{ linear in } [t_{n-1}, t_n], n = 1, \dots, N\}. \quad (4.18)$$

Given an initial guess $U_h^0 \in \mathbb{K}^0$, we define feasible set

$$\tilde{\mathbb{K}} := \{V \in \mathbb{K} \mid V(t_0) = U_h^0\}.$$

Numerical Scheme

Now we formulate the following fully discrete numerical approximation of Problem 4.7 by using linear finite elements in space and backward Euler method in time:

Method 4.13 (*Fully-discrete Method*) Given the approximation $F^n \in L^2(\Omega)$ of f at time t_n for $1 \leq n \leq N$, and initial guess $U_h^0 \in \mathbb{K}^0$, find an approximate solution $U_h^n \in \mathbb{K}^n$ for $1 \leq n \leq N$ such that

$$\frac{1}{k_n} \langle U_h^n - U_h^{n-1}, U_h^n - v_h \rangle + a(U_h^n, U_h^n - v_h) \leq \langle F^n, U_h^n - v_h \rangle \quad \forall v_h \in \mathbb{K}^n. \quad (4.19)$$

Remark 4.14 (Existence and Uniqueness of Solution) Based on the general existence theory for elliptic problems developed in Chapter 1, we know that the inequality (4.19) has a unique solution for any $1 \leq n \leq N$.

Discrete Problem

The discrete problem (4.19) admits a unique solution [74]. Moreover, let $\{\psi_{z_i}\}_{i=1}^I$ be the set of nodal basis functions, and let

$$\mathbf{A} := \left(\langle \psi_i, \psi_j \rangle + k_n a(\psi_i, \psi_j) \right)_{i,j=1}^I$$

be the resulting matrix of (4.19). If $\vec{\mathbf{U}} = (\mathbf{U}_i)$, $\vec{\mathbf{X}} = (\mathbf{X}_i) \in \mathbb{R}^I$ are the vector of nodal values of U_h^n and χ_h^n , namely

$$U_h^n = \sum_{i=1}^I \mathbf{U}_i \psi_{z_i} \quad \text{and} \quad \chi_h^n = \sum_{i=1}^I \mathbf{X}_i \psi_{z_i},$$

and the right-hand side

$$\vec{\mathbf{F}} = (\mathbf{F}_i) := \left(\langle U_h^n + k_n F^n, \psi_i \rangle \right)_{i=1}^I,$$

then $\vec{\mathbf{U}}$ satisfies the variational inequality:

$$\text{Find } \vec{\mathbf{U}} \geq \vec{\mathbf{X}} : \quad (\mathbf{A}\vec{\mathbf{U}} - \vec{\mathbf{F}})^T (\vec{\mathbf{U}} - \vec{\mathbf{V}}) \leq 0 \quad \forall \vec{\mathbf{V}} \geq \vec{\mathbf{X}}. \quad (4.20)$$

In (4.20), it is trivial to see $\vec{\mathbf{U}} \geq \vec{\mathbf{X}}$. Taking $\vec{\mathbf{V}} = \vec{\mathbf{U}} + \vec{\mathbf{W}}$ for any $\vec{\mathbf{W}} \geq 0$, we obtain that $\mathbf{A}\vec{\mathbf{U}} - \vec{\mathbf{F}} \geq 0$. Furthermore, by taking $\vec{\mathbf{V}} = \vec{\mathbf{X}}$ and $\vec{\mathbf{U}} + (\vec{\mathbf{U}} - \vec{\mathbf{X}})$,

respectively, we can see $(\mathbf{A}\vec{\mathbf{U}} - \vec{\mathbf{F}})^T(\vec{\mathbf{U}} - \vec{\mathbf{X}}) = 0$. In this way, we obtain a discrete linear complementarity problem (LCP) as in §2.3.2:

$$\mathbf{A}\vec{\mathbf{U}} \geq \vec{\mathbf{F}}, \quad \vec{\mathbf{U}} \geq \vec{\mathbf{X}}, \quad (\mathbf{A}\vec{\mathbf{U}} - \vec{\mathbf{F}})^T(\vec{\mathbf{U}} - \vec{\mathbf{X}}) = 0; \quad (4.21)$$

We shall discuss how to solve this finite dimensional variational inequality in great detail in § 7.5.

Chapter 5

A Priori Error Estimation

In this chapter, we consider *a priori* error estimation for the numerical methods proposed in Chapter 4 for both stationary and evolutionary variational inequalities. Here we shall assume both the time horizon $[0, T]$ and the polygonal space domain Ω are partitioned uniformly. The main purpose is to derive discretization error in terms of time step-size k and space meshsize h .

5.1 A Priori Error Estimation for EVIs

Before we look at the parabolic variational inequality (4.14), we first review the convergence results of linear finite element method for the elliptic variational inequality (4.1). This discussion motivates the optimal convergence rate proof for parabolic problems in §5.2. The general discussion on a priori error estimations of finite element methods for linear elliptic PDEs can be found, for example, in [29].

The first a priori error estimation for elliptic variational inequality was given by Falk [69] for symmetric bilinear form $a(\cdot, \cdot)$ (but the proof works for non-symmetric problems also) in the abstract setting discussed in chapter 1. In [69], Falk proved optimal convergence rate for linear elements for problems with homogenous boundary data. Later, the result was extended to the nonhomogenous case, quadratic elements and mixed finite elements by Brezzi, Hager, and Raviart [35, 36].

5.1.1 Abstract Error Estimation

We now present an optimal approximation result which is a modification of [69, Theorem 1].

Theorem 5.1 (Optimal Approximation Property) *Suppose \mathcal{A} is continuous and elliptic as in (1.7) and (1.8). Let u and u_h be the solutions of (4.1) and (4.3), respectively. If $f - \mathcal{A}u \in \mathcal{V}^*$, then*

$$\begin{aligned} \|u - u_h\|_{\mathcal{V}}^2 &\lesssim \inf_{v_h \in \mathbb{K}} \left\{ \|u - v_h\|_{\mathcal{V}}^2 + \|f - \mathcal{A}u\|_{\mathcal{V}^*} \|u - v_h\|_{\mathcal{V}} \right\} \\ &\quad + \|f - \mathcal{A}u\|_{\mathcal{V}^*} \inf_{v \in \mathcal{K}} \|u_h - v\|_{\mathcal{V}}. \end{aligned} \quad (5.1)$$

Furthermore, if $f - \mathcal{A}u \in \mathcal{H}$, then

$$\begin{aligned} \|u - u_h\|_{\mathcal{V}}^2 &\lesssim \inf_{v_h \in \mathbb{K}} \left\{ \|u - v_h\|_{\mathcal{V}}^2 + \|f - \mathcal{A}u\|_{\mathcal{H}} \|u - v_h\|_{\mathcal{H}} \right\} \\ &\quad + \|f - \mathcal{A}u\|_{\mathcal{H}} \inf_{v \in \mathcal{K}} \|u_h - v\|_{\mathcal{H}}. \end{aligned} \quad (5.2)$$

Before we prove the results above, it is worth mentioning the following comments for better understanding of the theorem.

Remark 5.2 (Approximation Error) The first term of the two inequalities above, (5.1) and (5.2), is the approximation error due to replacing the infinite-dimensional test function spaces by the finite element subspace. If the solution satisfies $f - \mathcal{A}u = 0$ in distribution sense, the above theorem reduces to the standard quasi-optimality (4.5) of finite element methods for linear elliptic boundary value problems.

Remark 5.3 (Non-conformity Error) The second term of (5.1) and (5.2) measures non-conformity of the approximate constraint set \mathbb{K} . If $\mathbb{K} \subset \mathcal{K}$ (conforming), this term vanishes and we only have the first approximability term; otherwise it tells how “different” the sets \mathbb{K} and \mathcal{K} are.

Proof of Theorem 5.1. Recall the continuous and discrete variational inequalities, (4.1) and (4.3):

$$\begin{aligned} a(u, u - v) &\leq \langle f, u - v \rangle \quad \forall v \in \mathcal{K} \\ a(u_h, u_h - v_h) &\leq \langle f, u_h - v_h \rangle \quad \forall v_h \in \mathbb{K}. \end{aligned}$$

By adding the last two inequalities, it is easy to see that

$$a(u, u) + a(u_h, u_h) \leq \langle f, u - v \rangle + \langle f, u_h - v_h \rangle + a(u, v) + a(u_h, v_h).$$

Subtracting $a(u, u_h) + a(u_h, u)$ from both sides of the inequality above, we obtain

$$a(u - u_h, u - u_h) \leq \langle f, u - v \rangle + \langle f, u_h - v_h \rangle + a(u, v - u_h) + a(u_h, v_h - u).$$

Since

$$a(u_h, v_h - u) = a(u - u_h, u - v_h) - a(u, u - v_h),$$

we regroup terms on the right-hand side to get

$$\begin{aligned} a(u - u_h, u - u_h) &\leq \left(\langle f, u - v_h \rangle - a(u, u - v_h) \right) + \left(\langle f, u_h - v \rangle - a(u, u_h - v) \right) \\ &\quad + a(u - u_h, u - v_h) \\ &= \langle f - \mathcal{A}u, u - v_h \rangle + \langle f - \mathcal{A}u, u_h - v \rangle + a(u - u_h, u - v_h) \end{aligned} \tag{5.3}$$

By coercivity of the bilinear form, the left-hand side of the above inequality yields

$$a(u - u_h, u - u_h) \geq C_* \|u - u_h\|_{\mathcal{V}}^2;$$

On the other hand, by continuity,

$$a(u - u_h, u - v_h) \leq C^* \|u - u_h\|_{\mathcal{V}} \|u - v_h\|_{\mathcal{V}}.$$

Then the theorem follows immediately from the Cauchy-Schwarz inequality and the last two inequalities and (5.3). \blacksquare

5.1.2 Application to Stationary Obstacle Problems

Based on the previous general approximation theorem, we obtain the following optimal error approximation of the linear finite element method for the Dirichlet obstacle problem, i.e. $\mathcal{V} = \tilde{H}^1(\Omega) = H_0^1(\Omega)$, $\mathcal{V}^* = H^{-1}(\Omega)$, $\mathcal{H} = L^2(\Omega)$, which has been discussed in Remark 2.15. The proof hinges on the regularity result (see Remark 2.15) and classical interpolation theory. We leave the proof out (for details, we refer to [69]).

Corollary 5.4 (Dirichlet Obstacle Problem) *Let Ω be a bounded convex polygonal domain. Let $f \in L^2(\Omega)$ and $\chi \in H^2(\Omega)$ be admissible. If u and u_h are the solutions of (4.1) and (4.3), respectively, then there exists a constant C which depends only on Ω, f and χ such that*

$$\|u - u_h\|_{H^1(\Omega)} \leq Ch.$$

Remark 5.5 (Higher Order Finite Element Approximation) From the previous corollary, the energy error converges optimally for linear finite element method respect to the approximation space and regularity of u . On the other hand, for quadratic finite element method, it has been shown [35, Lemma 4.3] that the convergence rate is $\mathcal{O}(h^{3/2-\varepsilon})$ for smooth enough f, χ and Ω . This is due to the lack of regularity of solutions of obstacle problems discussed in Remark 2.15.

5.2 A Priori Error Estimation for PVIs

In this section, we shall consider parabolic variational inequalities, and derive optimal a priori error bound in $\mathfrak{J}(0, T)$ -norm. We can further assume that the conforming condition $\mathbb{K} \subseteq \mathcal{K}(0, T)$ is satisfied. In fact, we can use the transformation $w := u - \chi$ to transform the original PVI to a problem with a simple constraint $\mathcal{K}(t) := \{v \in \mathcal{V} \mid v \geq 0\}$. Hence, for simplicity, we assume $\chi = 0$ in this chapter. Let $\Omega \subset \mathbb{R}^d$ be an open and bounded polygonal domain and \mathcal{T} be a quasi-uniform mesh (with meshsize h) of Ω . Let $\mathbb{V}(\mathcal{T})$ be the \mathbb{P}^1 finite element space associated with \mathcal{T} . Furthermore, we use uniform time partition with time step-size k .

5.2.1 Introduction

A priori error of the semi-discrete problem (4.16) has been studied in [11, 109, 110]. Baiocchi [11] proved that for initial solution $u_0 \in H^2(\Omega)$ the error $u - U$ in the energy norm converges a priori with order of $\mathcal{O}(k)$ for \mathcal{A} being the Laplace operator. In [109, 110], Nochetto et al. proved optimal convergence rate of the

backward Euler solution for more general evolution problems with a very different approach, a special case of which has been given in Lemma 4.12.

There are a number of a priori error estimates available for the fully-discrete method (4.19) for parabolic variational inequality (4.14). Johnson assumes that $u_0 \in W_\infty^2(\Omega)$ and obtains (with some additional assumptions on the speed of propagation of the free boundary) an error estimate $\mathcal{O}((\log k^{-1})^{1/4}k^{3/4} + h)$ for the $L^2(0, T; H^1(\Omega))$ error for implicit Euler (time) and linear finite element (space) discretization. Vuik [138] generalized the error estimation for general θ -scheme; he used the same techniques as Johnson and obtained same suboptimal convergence rate in time. Berger and Falk [17] analyzed the convergence of truncation method (using linear finite element, explicit time scheme) for a class of parabolic variational inequalities and obtained the $L^2(0, T; H^1(\Omega))$ error can be bounded by $C_\epsilon(h + k^{1-\epsilon})$ when $k \lesssim h^2$. More recently, Fetter [71] obtained an almost optimal L^∞ error bound using an auxiliary parabolic variational inequalities assuming $u_{tt} \in L^2(0, T; L^2(\Omega))$.

To prove optimal convergence of the fully-discrete scheme (4.19), we will take full advantage of recent developments in the error analysis for time-discretization for evolution problems [109, 110]. We carry out the error estimation in two steps: first we look at the error between the temporal semi-discretization solution U and fully-discrete solution U_h and introduce a general estimation for the energy error; then we apply the known results for the error of semi-discrete solution U (Lemma 4.12) as well as the regularity result for semi-discrete solution U (Lemma 4.10) to prove Theorem 5.11.

5.2.2 Estimation of Space Error

Applying the standard energy method, we give a general estimation of the “space” error $\|U - U_h\|_{L^2(0, T; \mathcal{V})}$. Recall that, by convention, $\|\cdot\|$ denotes the \mathcal{H} -norm and $\|\!\|\!\|\cdot\|\!\|\!\|$ is the energy norm.

Lemma 5.6 (Abstract Error Estimation of Space Error) *Let U and U_h be the solutions of the temporal semi-discrete problem (4.16) and the fully-discrete prob-*

lem (4.19), respectively. If

$$E^2(0, T; \Omega) := \|(U - U_h)(T)\|^2 + \int_0^T \|\bar{U} - \bar{U}_h\|^2 dt,$$

and $\bar{F} - \delta_t U - \mathcal{A}\bar{U} \in L^\infty(0, T; \mathcal{H})$, then we have that

$$E^2(0, T; \Omega) \lesssim \inf_{V_h \in \tilde{\mathbb{K}}} \left(\|U(T) - V_h(T)\|^2 + \int_0^T \mathfrak{E}^2 dt \right),$$

where

$$\mathfrak{E}^2(t) := \|\bar{F} - \delta_t U - \mathcal{A}\bar{U}\| \cdot \|\bar{U} - \bar{V}_h\| + \|\bar{U} - \bar{V}_h\|^2 + \|\delta_t(U - V_h)\|^2$$

is a piecewise constant function.

Proof. For convenient of the presentation, we first define

$$L(t) := \langle \delta_t(U - U_h), \bar{U} - \bar{U}_h \rangle + a(\bar{U} - \bar{U}_h, \bar{U} - \bar{U}_h).$$

Here, as we defined in §4.2, we use the following notation

$$\delta_t V(t) := \frac{V^n - V^{n-1}}{k_n} \quad \text{if } t \in (t_{n-1}, t_n],$$

for any piecewise linear (in time) function V .

Integrating $L(t)$ in time and applying the triangle inequality, we get that

$$\begin{aligned} \int_0^T L(t) dt &= \sum_{n=1}^N \int_{t_{n-1}}^{t_n} \langle \delta_t(U - U_h), \bar{U} - \bar{U}_h \rangle dt + \int_0^T \|\bar{U} - \bar{U}_h\|^2 dt \\ &= \sum_{n=1}^N \langle (U^n - U_h^n) - (U^{n-1} - U_h^{n-1}), U^n - U_h^n \rangle + \int_0^T \|\bar{U} - \bar{U}_h\|^2 dt \\ &\geq \sum_{n=1}^N \left\{ \frac{1}{2} \|U^n - U_h^n\|^2 - \frac{1}{2} \|U^{n-1} - U_h^{n-1}\|^2 \right\} + \int_0^T \|\bar{U} - \bar{U}_h\|^2 dt. \end{aligned} \tag{5.4}$$

On the other hand, for any finite element function $v_h \in \mathbb{K}$, we always have $L(t) = \text{I} + \text{II} + \text{III}$ where

$$\begin{aligned} \text{I} &:= \langle \delta_t(U - U_h), \bar{U} - v_h \rangle, \\ \text{II} &:= a(\bar{U} - \bar{U}_h, \bar{U} - v_h), \\ \text{III} &:= \langle \delta_t(U - U_h), v_h - \bar{U}_h \rangle + a(\bar{U} - \bar{U}_h, v_h - \bar{U}_h). \end{aligned}$$

We keep the first part as it is and start to estimate the other two parts.

By the Cauchy-Schwarz and Hölder's inequalities, we get that

$$\mathbb{II} \leq \frac{1}{2} \|\bar{U} - \bar{U}_h\|^2 + \frac{1}{2} \|\bar{U} - v_h\|^2.$$

For the third part, we first divide it into three parts and apply the semi-discrete as well as fully-discrete variational inequalities (4.16) and (4.19). Then we obtain that

$$\begin{aligned} \mathbb{III} &= \langle \delta_t U, v_h - \bar{U} \rangle + a(\bar{U}, v_h - \bar{U}) + \underbrace{\langle \delta_t U, \bar{U} - \bar{U}_h \rangle + a(\bar{U}, \bar{U} - \bar{U}_h)}_{\text{apply (4.16)}} \\ &\quad + \underbrace{\langle \delta_t U_h, \bar{U}_h - v_h \rangle + a(\bar{U}_h, \bar{U}_h - v_h)}_{\text{apply (4.19)}} \\ &\leq \langle \delta_t U + \mathcal{A}\bar{U} - \bar{F}, v_h - \bar{U} \rangle. \end{aligned}$$

Hence, in the above two inequalities, we take a piecewise constant function $v_h \in \mathbb{K}$ such that

$$v_h(t) = V_h^n \in \mathbb{K}^n \quad \text{for } t \in (t_{n-1}, t_n], \quad n = 1, \dots, N$$

and obtain that

$$\begin{aligned} \mathbb{L}(t) &\leq \langle \delta_t(U - U_h), \bar{U} - \bar{V}_h \rangle + \frac{1}{2} \|\bar{U} - \bar{U}_h\|^2 + \frac{1}{2} \|\bar{U} - \bar{V}_h\|^2 \\ &\quad + \|\bar{F} - \delta_t U - \mathcal{A}\bar{U}\| \cdot \|\bar{U} - \bar{V}_h\| \\ &\leq \langle \delta_t(U - U_h), \bar{U} - \bar{V}_h \rangle + \frac{1}{2} \|\bar{U} - \bar{U}_h\|^2 + \mathfrak{E}^2. \end{aligned} \quad (5.5)$$

Combining (5.4) with (5.5), we directly get

$$\begin{aligned} &\frac{1}{2} \|(U - U_h)(T)\|^2 + \frac{1}{2} \int_0^T \|\bar{U} - \bar{U}_h\|^2 dt \\ &\leq \frac{1}{2} \|U^0 - U_h^0\|^2 + \int_0^T \langle \delta_t(U - U_h), \bar{U} - \bar{V}_h \rangle dt + \int_0^T \mathfrak{E}^2 dt. \end{aligned} \quad (5.6)$$

Now we are left with the term $\int_0^T \langle \delta_t(U - U_h), \bar{U} - \bar{V}_h \rangle dt$. Using summation by parts, we get

$$\begin{aligned} \int_0^T \langle \delta_t(U - U_h), \bar{U} - \bar{V}_h \rangle dt &= \langle U^N - U_h^N, U^N - V_h^N \rangle - \langle U^0 - U_h^0, U^0 - V_h^0 \rangle \\ &\quad - \sum_{n=1}^N \int_{t_{n-1}}^{t_n} \langle U^n - U_h^n, \delta_t(U - V_h) \rangle dt. \end{aligned}$$

On the right-hand side, we take any $V_h \in \widetilde{\mathbb{K}}$ ($V_h^0 := V_h(t_0) = U_h^0$) to obtain by the Cauchy-Schwarz inequality

$$\begin{aligned} \int_0^T \langle \delta_t(U - U_h), \bar{U} - \bar{V}_h \rangle dt &\leq \frac{1}{4} \|(U - U_h)(T)\|^2 + \|(U - V_h)(T)\|^2 - \|U^0 - U_h^0\|^2 \\ &\quad + \sum_{n=1}^N \int_{t_{n-1}}^{t_n} \frac{\varepsilon}{4} \|U^n - U_h^n\|^2 + \frac{1}{\varepsilon} \|\delta_t(U - V_h)\|^2 dt \end{aligned}$$

Hence, by choosing an appropriate ε , it follows from the last inequality that

$$\begin{aligned} \int_0^T \langle \delta_t(U - U_h), \bar{U} - \bar{V}_h \rangle dt &\leq \frac{1}{4} \|(U - U_h)(T)\|^2 + \frac{1}{4} \int_0^T \|\bar{U} - \bar{U}_h\|^2 dt \\ &\quad - \|U^0 - U_h^0\|^2 + \|(U - V_h)(T)\|^2 + \int_0^T \mathfrak{E}^2 dt. \end{aligned} \quad (5.7)$$

Combining inequalities (5.6) and (5.7), we get the desired result. \blacksquare

Remark 5.7 (Comparison with Existing Analysis) Notice that, in the previous lemma, we only deal with piecewise constant and piecewise linear functions (in time). This gives us the advantage to get around a mixed term (in our notation) like

$$\left\langle \partial_t u(t_{n+1}) - \frac{u^{n+1} - u^n}{k_n}, U_h^{n+1} - u(t_{n+1}) \right\rangle$$

as analyzed in [82, 138], which is responsible for a suboptimal convergence rate as well as an additional requirement on the free boundary.

5.2.3 Positivity Preserving Operators

Positivity preserving operators are of particular interests for obstacle problems because we usual need the outcome of the approximation operator still satisfies the obstacle constraints. The piecewise linear interpolation operator preserves positivity and gives optimal approximation property; but unfortunately, it is well known that the interpolation operators are not stable in $H^1(\Omega)$ and can be only well-defined for continuous functions. The usual averaging approximation operators, like the Clément operator [51] or the Scott-Zhang operator [125], are stable but not positive. A positive operator which is stable and has optimal approximation properties on

polygonal domains has been constructed by Chen and Nchetto [49] and further analyzed in [116].

First we define the positive preserving operator given by Chen and Nchetto [49]. We denote the interior nodes of \mathcal{T} by $\{x_i\}_{i=1}^I$. Recall that $\{\psi_i\}_{i=1}^I$ are the canonical nodal basis functions of $\mathbb{V}(\mathcal{T})$, i.e. $\psi_i(x_j) = \delta_{ij}$ for $j = 1, \dots, I$. For each $1 \leq i \leq I$, let ω_i be the support of ψ_i , i.e.

$$\omega_i := \cup\{\tau \in \mathcal{T} \mid \text{supp}(\psi_i) \cap \tau \neq \emptyset\}.$$

For any $\tau \in \mathcal{T}$, we denote the union of elements surrounding τ by ω_τ :

$$\omega_\tau := \cup\{\tau' \in \mathcal{T} \mid \tau' \cap \tau \neq \emptyset\}.$$

Let B_i be the maximal ball centered at x_i and $B_i \subset \omega_i$. For any $v \in L^1(\Omega)$, we define the operator $\Pi_h : L^1(\Omega) \rightarrow \mathbb{V}(\mathcal{T})$ by

$$(\Pi_h v)(x) := \sum_{i=1}^I \left(\frac{1}{|B_i|} \int_{B_i} v \right) \psi_i(x). \quad (5.8)$$

From the definition above, it is clear that the operator Π_h preserves positivity, i.e.

$$\Pi_h v \geq 0 \quad \forall v \geq 0. \quad (5.9)$$

Furthermore, due to the symmetry of B_i with respect to x_i , we have

$$(\Pi_h v)(x_i) = v(x_i) \quad \forall v \in \mathbb{P}^1(B_i).$$

Next we review briefly the stability and optimal approximation results of Π_h ; for the proof, see [49, Section 3].

Lemma 5.8 (Stability) *For any $\tau \in \mathcal{T}$ and $1 \leq p \leq \infty$, the following estimates hold*

1. $\|\Pi_h v\|_{L^p(\tau)} \lesssim \|v\|_{L^p(\tau)} \quad \forall v \in L^p(\Omega);$
2. $\|\nabla \Pi_h v\|_{L^p(\tau)} \lesssim \|\nabla v\|_{L^p(\tau)} \quad \forall v \in W^{1,p}(\Omega).$

Lemma 5.9 (Optimal Approximation) For any $\tau \in \mathcal{T}$ and $1 \leq p \leq \infty$, we have the following estimation

$$\|v - \Pi_h v\|_{W^{j,p}(\tau)} \lesssim h_\tau^{m-j} \|D^m v\|_{L^p(\omega_\tau)} \quad \forall v \in W^{m,p}(\Omega) \cap \tilde{W}^{1,p}(\Omega),$$

where $j = 0, 1$ and $m = 1, 2$.

Remark 5.10 (General Order) Using the interpolation estimate (Proposition 2.1), this result can also be applied for any real number $0 \leq s \leq 1$ to obtain optimal approximation property

$$\|v - \Pi_h v\|_{W^{s,p}(\tau)} \lesssim h_\tau^{m-s} \|D^m v\|_{L^p(\omega_\tau)} \quad \forall v \in W^{2,p}(\Omega) \cap \tilde{W}^{1,p}(\Omega).$$

5.2.4 Optimal Convergence Rate

In this section, we shall present an optimal convergence result for the fully-discrete method (4.13) in $L^\infty(0, T; L^2(\Omega)) \cap L^2(0, T; H^1(\Omega))$ -norm.

Theorem 5.11 (A Priori Error Estimation for PVI) Let Ω be a convex polygonal domain. Let $\mathcal{A} = -\Delta$. Let

$$f \in BV(0, T; L^2(\Omega)) \quad \text{and} \quad u_0 \in H^2(\Omega) \cap \mathcal{K}.$$

Given an initial guess U_h^0 satisfying

$$U_h^0 \geq 0 \quad \text{and} \quad \|u_0 - U_h^0\| = \mathcal{O}(h),$$

we have the error estimate

$$\max_{1 \leq n \leq N} \|u(t_n) - U_h^n\|^2 + \int_0^T \|u - \bar{U}_h\|_{H^1(\Omega)}^2 \leq C(k^2 + h^2). \quad (5.10)$$

Proof of Theorem 5.11: Recall that, in our convention, $u^n = u(t_n)$ and \bar{u} is the piecewise linear (in time) function. Applying the triangle inequality, we obtain that

$$\int_0^{t_{n_0}} \| \|u - \bar{U}_h\| \|^2 dt \leq 2 \int_0^{t_{n_0}} \| \|u - \bar{U}\| \|^2 dt + 2 \int_0^{t_{n_0}} \| \| \bar{U} - \bar{U}_h \| \|^2 dt, \quad (5.11)$$

for any integer $1 \leq n_0 \leq N$.

For the first term (time error) on the right-hand side of (5.11), a consequence of Lemma 4.12

$$\int_0^{t_{n_0}} \left\| \|u - \bar{U}\| \right\|^2 dt \lesssim \mathcal{O}(k^2). \quad (5.12)$$

For the second term (space error) on the right-hand side of (5.11), we choose piecewise linear function V_h in the approximation property, Lemma 5.6, such that

$$V_h(0) = U_h^0, \quad \text{and} \quad V_h(t_n) = \Pi_h U(t_n), \quad n = 1, \dots, n_0,$$

where Π_h be the positive operator defined in §5.2.3. For any $0 \leq n \leq N$, since $U(t_n) \geq 0$, we have $\Pi_h U(t_n) \geq 0$. Hence $V_h \in \tilde{\mathbb{K}}$ is admissible. Consequently, the regularity results $U \in L^\infty(0, T; H^2(\Omega))$ (see Remark 2.15 and Lemma 2.20) and $\delta_t U \in L^\infty(0, T; L^2(\Omega)) \cap L^2(0, T; H^1(\Omega))$ (see Lemma 4.10) give the estimate

$$\|(U - U_h)(t_{n_0})\|^2 + \int_0^{t_{n_0}} \left\| \bar{U} - \bar{U}_h \right\|^2 dt \lesssim \mathcal{O}(h^2). \quad (5.13)$$

Plugging (5.12) and (5.13) into (5.11), we arrive at

$$\|(u - U_h)(t_{n_0})\|^2 + \int_0^{t_{n_0}} \left\| \|u - \bar{U}_h\| \right\|^2 dt \lesssim \mathcal{O}(k^2 + h^2).$$

Note that the last inequality is true for arbitrary positive integer $0 < n_0 \leq N$. We can pick n_0 such that $\|(u - U_h)(t_{n_0})\|^2$ is maximized. Hence the estimation (5.10) is established. ■

Remark 5.12 (More General Operator) The operator \mathcal{A} does not need to be $-\Delta$ in the previous theorem. The proof can be extended to general second order elliptic operator case.

Chapter 6

A Posteriori Error Estimation

Since the seminal work by Babuška and Rheinboldt [8], a considerable amount of effort has been made in developing reliable and efficient adaptive algorithms for boundary value problems over the last three decades. The main idea of adaptive algorithms is to generate a discretization of the time-space domain such that local error is equally distributed.

Since local error is not available in general, computable local error estimators play a major role in designing adaptive schemes. Compared with a priori error estimates discussed in the previous chapter, *a posteriori* error estimators possess the following important features:

- They are computable and depend only on discrete solutions and data, instead of the exact solutions.
- They are quantitative and so instrumental for adaptive mesh generation and error control.

Before we discuss the a posteriori error estimation for our particular problem, it is worth mentioning some of its general principles:

1. **Reliability.** We require the computable error estimator (denoted by \mathcal{E}) to be a global upper bound of the error in certain norm (denoted by E) up to a multiplicative constant, i.e. $E \leq C_1 \mathcal{E}$. This means the error estimator \mathcal{E} is *reliable* in the following sense: if the error estimator is small enough, then the real error will not be too big neither.

2. **Efficiency.** A reliable error estimator \mathcal{E} could over-estimate the error E . To guarantee over-estimation does not happen, we require \mathcal{E} to be efficient, i.e. \mathcal{E} is also a global lower bound of the error, i.e. $\mathcal{E} \leq C_2 E$.
3. **Estimation Quality.** The ratio C_1/C_2 provides important information of the quality of the error estimator. If this ratio is close to 1, then the error estimator is very close to the error.
4. **Local Error Estimation.** To derive an adaptive algorithm from a reliable and efficient a posteriori estimator \mathcal{E} , the global upper and lower bounds are not enough. We need information of local error to decide where more computational effort is needed. To achieve this, the estimator \mathcal{E} should be localizable, i.e. $\mathcal{E} = \sum_{\tau \in \mathcal{T}} \mathcal{E}(\tau)$, with each local indicator $\mathcal{E}(\tau)$ providing some information of the local error $E(\tau)$ on element τ . Mathematically, this can be expressed as local efficiency or a local lower bound of the form $\mathcal{E}(\tau) \lesssim E(\tau)$. This suggests that we have to reduce the local estimator $\mathcal{E}(\tau)$ to reduce the local error.

For classical theories and techniques of a posteriori error estimation of elliptic partial differential equations, we refer interested readers to the reviews by Verfürth [135] and Ainsworth and Oden [2].

Since reliable and efficient a posteriori error estimation is the key to develop efficient adaptive schemes, we shall explain this part carefully in this chapter. The main material of this chapter is based on [104, 115, 117]. The rest of the chapter is organized as follows. We first introduce the main idea of a posteriori error estimation for obstacle problems in §6.1. Then we consider the conforming case when the discrete obstacle $\chi_h = \chi$: we give a posteriori error estimators for elliptic variational inequalities in §§6.2, 6.3, and 6.4 and discuss how to deal with time-dependent problems and time discretization error in §6.5. We then extend our analysis for general obstacle χ for which numerical approximation of χ introduces additional obstacle consistency error in §6.6. Finally, we consider mesh changes as well as coarsening error in §6.7.

6.1 Introduction

For variational inequalities (VI), the *a posteriori* error analysis is very recent and rather intricate. One of the difficulties is that VI's lead to non-Lipschitz nonlinearities and the *linearization* techniques [135] used for nonlinear problems do not work any longer.

To gain some insight on the difficulties involved, we let $\mathfrak{F}(u) := \mathcal{A}u + \lambda(u)$ be the nonlinear operator discussed in §2.3.2, which consists of the linear operator \mathcal{A} and the nonlinear part λ that accounts for the unilateral constraint $u \geq \chi$. The Lagrange multiplier, λ , satisfies

$$\lambda(u) = \begin{cases} f - u_t - \mathcal{A}u \leq 0 & \text{in } \mathcal{C} = \{u = \chi\} \\ 0 & \text{in } \mathcal{N} = \{u > \chi\}; \end{cases} \quad (6.1)$$

hence $\lambda(u)$ restores the equality in (3.11), namely,

$$u_t + \mathcal{A}u + \lambda(u) = f. \quad (6.2)$$

A posteriori error estimates of residual type are obtained by plugging the discrete solution U into the PDE. Roughly speaking, we get the defect measure

$$\mathcal{G} = f - U_t - \mathcal{A}U - \lambda(U), \quad (6.3)$$

which is called *Galerkin functional* in this nonlinear context; the precise definition is given in §6.2 for elliptic VI and §6.5 for parabolic VI, respectively. This is a replacement for the usual residual in linear theory. To obtain sharp a posteriori error estimators, we must be able to provide a discrete multiplier $\lambda(U)$ with properties similar to (6.1).

In fact, the linear part r of \mathcal{G} , that is $r := f - U_t - \mathcal{A}U$, does not give correct information in the contact set \mathcal{C} , where the solution adheres to the obstacle regardless of the size of r . Notice that r is the usual residual for linear PDE. We point out that failure to recognize the importance of $\lambda(u)$ leads to a global upper bound of the error but not to a global lower bound [49]; overestimation is thus possible.

This issue was first addressed for elliptic variational inequalities by Veerer [134] and further improved by Fierro and Veerer [72] in $H^1(\Omega)$. Nochetto, Siebert,

and Veerer extended these estimates to $L^\infty(\Omega)$ and derived barrier set estimates [113, 114]. The *duality* approach, reported in [12], is not suitable in this setting because of the singular character of $\lambda(u)$.

A residual-type $L^2(0, T; H^1(\Omega))$ error estimator was proposed for parabolic variational inequalities by Moon et al [104]. If the variational inequality becomes an equality, the *energy* estimates in [104] reduce to those in [16, 119, 136]. More recently the estimator proposed in [104] was extended to variational integro-differential inequalities [115].

For problems with integro-differential operators, another difficulty arises, namely the non-local character of integral operators. On the other hand, in many practical problems, the integral operators are of pseudo-differential type and possess some pseudo-local properties. In particular, for the integral operator \mathcal{A}_I (3.12), we have

$$\text{sing supp } \mathcal{A}_I v \subset \text{sing supp } v,$$

for any $v \in C^\infty(\Omega)^*$ [133, Theorem II.2.1]. Here the singular support of a distribution v , denoted by $\text{sing supp } v$, is the complement of the open set on which v is smooth. Due to the pseudo-local properties, the adaptive algorithms work well in practice [147]. Adaptive finite and boundary element methods have been discussed for integral equations in several papers [139, 140, 43, 41, 42, 68, 66, 67].

6.2 Stationary Problems

To explain the main idea of our a posteriori error estimation, we first look at the elliptic variational inequality problem, Problem 2.13: find $u \in \mathcal{K}$ such that

$$\langle \mathcal{A}u - f, u - v \rangle \leq 0 \quad \forall v \in \mathcal{K} := \{v \mid v \geq \chi, v \in \mathcal{V}\}. \quad (6.4)$$

We use linear finite element method (see §4.1) to solve this problem numerically. Consider the discrete convex set corresponding to \mathcal{K}

$$\mathbb{K} := \{v \in \mathbb{V} : v \geq \chi_h\}, \quad (6.5)$$

where $\mathbb{V} \subset \tilde{H}^s(\Omega)$ is the continuous piecewise linear finite element space.

For the moment, we assume that the approximate obstacle is exactly equal to the real obstacle, i.e. $\chi_h = \chi$. This means our discrete feasible set is conforming, $\mathbb{K} \subset \mathcal{K}$. The more general case where \mathbb{K} is not a subset of \mathcal{K} will be discussed later. Now we formulate the following numerical approximation of the inequality (6.4) by using piecewise linear finite elements: find $u_h \in \mathbb{K}$ such that

$$\langle \mathcal{A}u_h - f, u_h - v \rangle \leq 0 \quad \forall v \in \mathbb{K}. \quad (6.6)$$

6.2.1 Lagrange Multiplier

As in the linear case, we define the residual to be

$$r_h = f - \mathcal{A}u_h. \quad (6.7)$$

Note that, for variational inequalities, the error equation $\mathcal{A}(u - u_h) = r_h$, which is the starting point for residual-type error estimations for linear elliptic PDEs (see §4.1), does not hold any more. Residual-type error estimators for elliptic variational inequalities have been given in [134, 113, 72, 111, 26].

The basic idea is to introduce an appropriate *computable* approximation λ_h of the Lagrange multiplier (see Definition 2.22)

$$\lambda := f - \mathcal{A}u \in H^{-s}(\Omega). \quad (6.8)$$

In Section 2.3.4, it has been shown that the Lagrange multiplier λ is *non-positive* and *vanish in the noncontact region* in the sense of distributions. Furthermore, it is clear that we have the following error equation

$$\mathcal{A}(u - u_h) = r_h - \lambda, \quad (6.9)$$

which corresponds to the error equation for linear equations.

6.2.2 Abstract Error Bounds

For the moment, we assume that we have obtained a computable approximate Lagrange multiplier $\lambda_h \leq 0$ and focus on how to get upper and lower bounds of

the error. Notice that the error bounds developed here are independent of the particular choices of the discrete Lagrange multiplier λ_h . We will discuss how to define a practical λ_h in Section 6.3.

From (6.9), it is easy to see that

$$\|u - u_h\|^2 = \langle \mathcal{A}(u - u_h), u - u_h \rangle = \langle r_h - \lambda, u - u_h \rangle. \quad (6.10)$$

Adding and subtracting $\lambda_h \leq 0$, by the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned} \|u - u_h\|^2 &= \langle r_h - \lambda_h + \lambda_h - \lambda, u - u_h \rangle \\ &\leq \frac{1}{2} \|r_h - \lambda_h\|_*^2 + \frac{1}{2} \|u - u_h\|^2 - \langle \lambda - \lambda_h, u - u_h \rangle. \end{aligned}$$

On the other hand, by definition (6.8) of λ , we have

$$\begin{aligned} \langle \lambda - \lambda_h, \varphi \rangle &= \langle f - \mathcal{A}u - \lambda_h, \varphi \rangle \\ &= \langle \mathcal{A}(u_h - u), \varphi \rangle + \langle r_h - \lambda_h, \varphi \rangle \quad \forall \varphi \in \tilde{H}^s(\Omega). \end{aligned} \quad (6.11)$$

Hence $\|\lambda - \lambda_h\|_*^2 \lesssim \|u - u_h\|^2 + \|r_h - \lambda_h\|_*^2$. So we can find an upper bound

$$\|u - u_h\|^2 + \|\lambda - \lambda_h\|_*^2 \lesssim \|r_h - \lambda_h\|_*^2 - \langle \lambda - \lambda_h, u - u_h \rangle.$$

For the second term on the right-hand side of the last inequality, since $u_h \in \mathbb{K} \subset \mathcal{K}$, we have $\langle \lambda, u - u_h \rangle \geq 0$ by the continuous variational inequality (6.4). Furthermore, with $\lambda_h \leq 0$, it is easy to see that

$$\langle \lambda_h, u - u_h \rangle = \underbrace{\langle \lambda_h, u - \chi \rangle}_{\leq 0} - \underbrace{\langle \lambda_h, u_h - \chi \rangle}_{\geq 0} \leq -\langle \lambda_h, u_h - \chi \rangle.$$

Hence

$$\|u - u_h\|^2 + \|\lambda - \lambda_h\|_*^2 \lesssim \|r_h - \lambda_h\|_*^2 - \langle \lambda_h, u_h - \chi \rangle. \quad (6.12)$$

Remark 6.1 (General Obstacle) Notice that the conformity assumption $\mathbb{K} \subset \mathcal{K}$ greatly simplifies the analysis of the term $\langle \lambda - \lambda_h, u - u_h \rangle$. For problems with general obstacles, this term also yields terms controlling the obstacle consistency error (we refer to [104] for details). We will revisit this when we discuss problems with a general obstacle later in §6.6.

By rearranging the terms in (6.11) and using the strong sector condition (2.5), we have, for any $\varphi \in \tilde{H}^s(\Omega)$, that

$$\langle r_h - \lambda_h, \varphi \rangle = \langle \mathcal{A}(u - u_h), \varphi \rangle + \langle \lambda - \lambda_h, \varphi \rangle \lesssim \|u - u_h\| \cdot \|\varphi\| + \|\lambda - \lambda_h\|_* \cdot \|\varphi\|.$$

Consequently, using the triangle inequality, we have

$$\|r_h - \lambda_h\|_*^2 \lesssim \|u - u_h\|^2 + \|\lambda - \lambda_h\|_*^2.$$

Hence $\|r_h - \lambda_h\|_*^2$ is also a lower bound of the error $\|u - u_h\|^2 + \|\lambda - \lambda_h\|_*^2$ up to a multiplicative constant.

Because of the important role of $r_h - \lambda_h$ in the error estimation (see [134, 72, 111, 104] also), we call it *Galerkin functional* and denote it by

$$\mathcal{G}_h := r_h - \lambda_h. \tag{6.13}$$

Then the previous analysis can be summarized in the following abstract lemma.

Lemma 6.2 (Abstract Error Bounds: Stationary Problems) *Let u and u_h be the solutions of (6.4) and (6.6), respectively. If $\lambda_h \leq 0$, we have the upper and lower bounds*

$$\|\mathcal{G}_h\|_*^2 \lesssim \|u - u_h\|^2 + \|\lambda - \lambda_h\|_*^2 \lesssim \|\mathcal{G}_h\|_*^2 - \langle \lambda_h, u_h - \chi \rangle. \tag{6.14}$$

6.3 Approximation of Lagrange Multipliers

In practice, it is important to find a “good” approximation λ_h , which mimics the properties of λ at the discrete level. The ideal choice would be $\lambda_h = \lambda$ of course, but this is impossible because λ is not computable. A simple-minded choice is to take $\lambda_h = 0$ and then Lemma 6.2 yields the standard upper bound for linear elliptic equations

$$\|u - u_h\|^2 \lesssim \|r_h\|_*^2.$$

However, this bound has the drawback that the residual r_h in the contact region contributes to the bound. In other words, even if u_h were the exact solution, we

would obtain a nonzero upper bound due to nonzero values of λ in the contact region. A good practical upper bound should be “localized” in the sense that only the value of the residual in the noncontact region contributes to the error bound.

6.3.1 Discrete Contact and Noncontact Sets

Before we can define the discrete Lagrange multiplier λ_h which gives a “localized” upper bound, we first need to define discrete sets that mimic the contact set $\mathcal{C} := \{u = \chi\}$ and noncontact set $\mathcal{N} := \{u > \chi\}$.

Let \mathcal{T} be a triangulation of the polygonal domain Ω and \mathcal{S} be the set of all sides or faces of triangles or tetrahedrons in \mathcal{T} . Denote by ω_z the support of the piecewise linear nodal basis functions $\{\psi_z\}_{z \in \mathcal{P}_h}$; see Figure 6.1. Let $\gamma_z \subset \mathcal{S}$ be the skeleton of ω_z , namely the set of all interior sides of ω_z which contain z ; for $d = 1$, γ_z reduces to the node z itself. Similarly, we denote ω_S be the set of triangles sharing

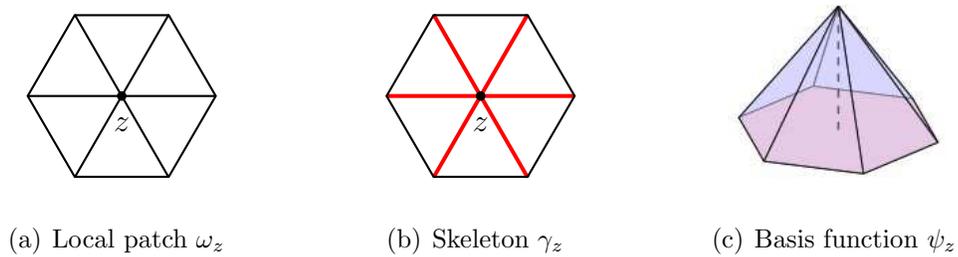


Figure 6.1: Local Patch

the side $S \in \mathcal{S}$ and ω_τ be the the union of elements surrounding $\tau \in \mathcal{T}$:

$$\omega_\tau := \cup\{\tau' \in \mathcal{T} \mid \tau' \cap \tau \neq \emptyset\}.$$

We split \mathcal{P}_h into three disjoint sets

$$\mathcal{P}_h = \mathcal{N}_h \cup \mathcal{C}_h \cup \mathcal{F}_h$$

with the noncontact nodes \mathcal{N}_h , full-contact nodes \mathcal{C}_h , and free boundary nodes \mathcal{F}_h

defined as follows:

$$\mathcal{N}_h := \{z \in \mathcal{P}_h \mid u_h > \chi \text{ in } \text{int } \omega_z\}, \quad (6.15a)$$

$$\mathcal{C}_h := \{z \in \mathcal{P}_h \mid u_h = \chi \text{ and } r_h \leq 0 \text{ in } \omega_z\}, \quad (6.15b)$$

$$\mathcal{F}_h := \mathcal{P}_h \setminus (\mathcal{N}_h \cup \mathcal{C}_h). \quad (6.15c)$$

The residual r_h contains two parts: a smooth part (interior residual) and a singular part (jump residual). Let the *interior residual* associated with \mathcal{A} to be

$$R(u_h) := f - \mathcal{A}_I u_h - c_1 \cdot \nabla u_h - c_0 u_h, \quad (6.16)$$

and the *jump residual* on the side $\tau_1 \cap \tau_2$ to be

$$J(u_h) := -c_2(\nabla u_h|_{\tau_1} \cdot \nu_1 + \nabla u_h|_{\tau_2} \cdot \nu_2), \quad (6.17)$$

where ν_i is the unit outer normal vector to the element $\tau_i \in \mathcal{T}$ for $i = 1, 2$.

Remark 6.3 (Separation of Sets) If $z \in \mathcal{N}_h$, then $u_h(z) > \chi(z)$. It is easy to see that there is no node in the neighborhood of z being a full-contact node. This is because the definition of \mathcal{C}_h requires $u_h = \chi$ in the whole star ω_z . Conversely, if $z \in \mathcal{C}_h$, then any node $x \in \mathcal{P}_h \cap \omega_z$ cannot be in \mathcal{N}_h . The noncontact nodes and the full-contact nodes are complete “separated” by the free boundary nodes.

Remark 6.4 (Sign Condition) Notice that r_h is not a discrete object, it is impossible to check the sign condition $r_h \leq 0$ in the definition (6.15b). In practice, we check $R(u_h) \leq 0$ at all quadrature nodes $x_q \in \omega_z$ and $J(u_h)|_S \leq 0$ for sides $S \subset \gamma_z$ instead.

6.3.2 Discrete Lagrange Multiplier

A first attempt for λ_h would be a piecewise linear function $\lambda_h = \sum_{z \in \mathcal{P}_h} s_z \psi_z$ in such a way that the nodal values s_z are weighted means on stars ω_z :

$$s_z := \begin{cases} \langle r_h, \psi_z \rangle / \langle 1, \psi_z \rangle & z \in \mathcal{P}_h \cap \Omega \\ 0 & z \in \mathcal{P}_h \cap \Gamma; \end{cases} \quad (6.18)$$

and s_z can be naturally divided into two parts $s_z = R_z + J_z$, where

$$R_z := \begin{cases} \langle R(u_h), \psi_z \rangle / \langle 1, \psi_z \rangle & z \in \mathcal{P}_h \cap \Omega \\ 0 & z \in \mathcal{P}_h \cap \Gamma \end{cases}$$

and

$$J_z := \begin{cases} -\langle c_2 \nabla u_h, \nabla \psi_z \rangle / \langle 1, \psi_z \rangle & z \in \mathcal{P}_h \cap \Omega \\ 0 & z \in \mathcal{P}_h \cap \Gamma. \end{cases}$$

Note that λ is zero on $\Gamma \cap \mathcal{N}$, which motivates us to define $s_z = 0$ on Γ . This definition yields $s_z \leq 0$ and $s_z = 0$ for $z \in \mathcal{N}_h$, and it is thus quite appropriate for \mathcal{N}_h but not necessarily for $z \in \mathcal{C}_h$. In fact, to achieve localization of the error estimator λ_h must equal the linear residual r_h in ω_z for $z \in \mathcal{C}_h$, thereby leading to $\lambda_h = r_h \leq 0$ in ω_z .

We can blend the two competing alternatives via the partition of unity $\{\psi_z\}_{z \in \mathcal{P}_h}$ and define formally the *discrete Lagrange multiplier*

$$\lambda_h := \sum_{z \in \mathcal{C}_h} r_h \psi_z + \sum_{z \in \mathcal{P}_h \setminus \mathcal{C}_h} s_z \psi_z. \quad (6.19)$$

As a consequence of $s_z \leq 0$ and the sign conditions in (6.15b), this definition guarantees that $\lambda_h \leq 0$ in Ω . With the choice of λ_h (6.19), the Galerkin functional vanishes in the numerical contact region in the sense of distributions (this is often called the *localization* property), i.e.

$$\mathcal{G}_h = \sum_{z \in \mathcal{P}_h} r_h \psi_z - \lambda_h = \sum_{z \in \mathcal{P}_h \setminus \mathcal{C}_h} (r_h - s_z) \psi_z. \quad (6.20)$$

Remark 6.5 (Formal Definition of λ_h and \mathcal{G}_h) The definitions of λ_h and \mathcal{G}_h are formal. Since the residual $r_h \in H^{-s}(\Omega)$ and is understood in the sense of distributions, we should view $r_h \psi_z$ also as a distribution. For any function $\varphi \in \tilde{H}^s(\Omega)$, we define

$$\langle r_h \psi_z, \varphi \rangle := \langle r_h, \varphi \psi_z \rangle.$$

Because $\varphi \psi_z \in \tilde{H}^s(\Omega)$, everything is well-defined.

Remark 6.6 (Approximation of Lagrange Multiplier) With this definition of λ_h , we can see that

$$\lambda - \lambda_h = \lambda - (r_h - \mathcal{G}_h) = (f - \mathcal{A}u) - (f - \mathcal{A}u_h) + \mathcal{G}_h = -\mathcal{A}(u - u_h) + \mathcal{G}_h.$$

Hence, using the strong sector condition (2.5) and the triangle inequality, we have

$$\|\lambda - \lambda_h\|_* \lesssim \|u - u_h\| + \|\mathcal{G}_h\|_*.$$

Therefore, if $\|u - u_h\|$ converges at the same rate as $\|\mathcal{G}_h\|_*$, the approximation error of Lagrange multipliers $\|\lambda - \lambda_h\|_*$ is of at least the same order.

6.4 Residual-type Error Estimation

We now derive a residual-type error estimator based on the abstract estimation derived in the previous section.

6.4.1 Upper Bound

In Lemma 6.2, we obtain an abstract upper bound formally. In practice, we still need to find a computable and localized upper bound of the dual norm $\|\mathcal{G}_h\|_*^2$ and a lower bound of $\langle \lambda_h, u_h - \chi \rangle$.

From Global to Local

We start with finding an upper bound of $\|\mathcal{G}_h\|_*^2$. It is equivalent to finding an upper bound of $\|\mathcal{G}_h\|_{H^{-s}(\Omega)}$. We first show that we can bound the global $H^{-s}(\Omega)$ -norm by a sum of localized norms on ω_z for $s \in [0, 1]$. Recall the definitions of local Sobolev norm $\|\cdot\|_{H_{\Gamma}^s(\omega_z)}$ and its dual norm $\|\cdot\|_{H_{\Gamma}^s(\omega_z)^*}$; see §2.1.

Lemma 6.7 (Localized Upper Bound of the Dual Norm) *Assume that $G = \sum_{z \in \mathcal{P}_h} g_z$ and $g_z \in H_{\Gamma}^s(\omega_z)^*$. For $s \in [0, 1]$ there holds*

$$\|G\|_{H^{-s}(\Omega)}^2 \leq (d+1) \sum_{z \in \mathcal{P}_h} \|g_z\|_{H_{\Gamma}^s(\omega_z)^*}^2. \quad (6.21)$$

Proof. We have for $v \in \tilde{H}^s(\Omega)$,

$$|\langle G, v \rangle| = \left| \sum_{z \in \mathcal{P}_h} \langle g_z, v \rangle \right| \leq \sum_{z \in \mathcal{P}_h} \|g_z\|_{H_\Gamma^s(\omega_z)^*} \|v\|_{H_\Gamma^s(\omega_z)} \quad (6.22)$$

$$\leq \left(\sum_{z \in \mathcal{P}_h} \|g_z\|_{H_\Gamma^s(\omega_z)^*}^2 \right)^{1/2} \left(\sum_{z \in \mathcal{P}_h} \|v\|_{H_\Gamma^s(\omega_z)}^2 \right)^{1/2} \quad (6.23)$$

Note that we have for $s = 0$ and $s = 1$ that

$$\sum_{z \in \mathcal{P}_h} \|v\|_{H_\Gamma^s(\omega_z)}^2 \leq (d+1) \|v\|_{\tilde{H}^s(\Omega)}^2 \quad (6.24)$$

since at most $d+1$ of the stars ω_z overlap on each simplex.

For any $v \in \tilde{H}^s(\Omega)$, we define an operator

$$T : \tilde{H}^s(\Omega) \rightarrow \prod_{z \in \mathcal{P}_h} H_\Gamma^s(\omega_z)$$

which restrict v to local patches, i.e.

$$T(v) := (v_z)_{z \in \mathcal{P}_h} \quad \text{with} \quad v_z(x) := \begin{cases} v(x) & x \in \omega_z \\ 0 & \text{otherwise.} \end{cases}$$

For $s = 0$ or $s = 1$, (6.24) gives $\|T(v)\|^2 \lesssim \|v\|_{\tilde{H}^s(\Omega)}^2$. By interpolation, we obtain (6.24) for all $s \in [0, 1]$, which in turn implies (6.21). \blacksquare

Although the right-hand side of (6.21) is localized, it is still not computable. The following lemma shows how to bound the negative norms by L^p norms:

Lemma 6.8 (Computable Upper Bound of Local Dual Norm) *For $z \in \mathcal{P}_h$, assume that $g_z \in L^p(\omega_z)$ satisfies $\int_{\omega_z} g_z = 0$ when $\partial\omega_z \cap \Gamma$ has measure 0. For $1 \leq p \leq 2$, let $d(\frac{1}{p} - \frac{1}{2}) < s \leq 1$. Then*

$$\|g_z\|_{H_\Gamma^s(\omega_z)^*} \lesssim h_z^{s+d(1/2-1/p)} \|g_z\|_{L^p(\omega_z)}. \quad (6.25)$$

Proof. Case i: $\partial\omega_z \cap \Gamma$ has measure 0. Then $\int_{\omega_z} g_z = 0$ by assumption. We have for $v \in H_\Gamma^s(\omega_z)$ and any constant $C_z \in \mathbb{R}$

$$|\langle g_z, v \rangle| = |\langle g_z, v - C_z \rangle| \leq \|g_z\|_{L^p(\omega_z)} \|v - C_z\|_{L^q(\omega_z)}$$

where $p^{-1}+q^{-1} = 1$ (for $p = 1$ we define $q = \infty$). Let $\hat{\omega} := h_z^{-1}\omega_z$ and $\hat{v}(\hat{x}) := v(h_z\hat{x})$, hence $\|v - C_z\|_{L^q(\omega_z)} = h_z^{d/q} \|\hat{v} - C_z\|_{L^q(\hat{\omega})}$. We have

$$\|\hat{v} - C_z\|_{L^q(\hat{\omega})} \lesssim \|\hat{v} - C_z\|_{H_{\Gamma}^s(\hat{\omega})}$$

since $H_{\Gamma}^s(\hat{\omega}) \hookrightarrow L^q(\hat{\omega})$ for $s > \frac{d}{2} - \frac{d}{q} = \frac{d}{p} - \frac{d}{2}$ (equality is true for $s = 0$ and $q = 2$, not true for $s = \frac{1}{2}$ and $d = 1, q = \infty$). We now choose the constant C_z as the mean value of \hat{v} on $\hat{\omega}_z$. For $s = 0$ we have

$$\|\hat{v} - C_z\|_{L^2(\hat{\omega}_z)} \leq \|\hat{v}\|_{L^2(\hat{\omega}_z)} = h_z^{-d/2} \|v\|_{L^2(\omega_z)}. \quad (6.26)$$

For $s = 1$ we use the second Poincaré's inequality

$$\|\hat{v} - C_z\|_{H^1(\hat{\omega}_z)} \lesssim |\hat{v}|_{H^1(\hat{\omega}_z)} = h_z^{1-d/2} |v|_{H^1(\omega_z)} \leq h_z^{1-d/2} \|v\|_{H^1(\omega_z)}. \quad (6.27)$$

Now we define an operator $T_z : L^2(\omega_z) \rightarrow L^2(\hat{\omega}_z)$ such that $T_z(v) := \hat{v} - C_z$. Then (6.26) and (6.27) give

$$\|T_z(v)\|_{L^2(\hat{\omega}_z)} \lesssim h_z^{-d/2} \|v\|_{L^2(\omega_z)} \quad \text{and} \quad \|T_z(v)\|_{H^1(\hat{\omega}_z)} \lesssim h_z^{1-d/2} \|v\|_{H^1(\omega_z)}.$$

Interpolation argument gives

$$\|\hat{v} - C_z\|_{H_{\Gamma}^s(\hat{\omega})} \lesssim h_z^{s-d/2} \|v\|_{H_{\Gamma}^s(\omega_z)} \quad \forall s \in [0, 1].$$

Case ii: $\partial\omega_z \cap \Gamma$ has positive measure. We take $C_z = 0$. Notice that (6.26) still holds for $C_z = 0$. For $s = 1$ we can now apply the first Poincaré's inequality to get (6.27). By the same argument as in Case i, we get (6.25). \blacksquare

Error Close to the Free Boundary

We now look at the second term at the right-hand side of (6.14) in Lemma 6.2. Clearly, $\langle \lambda_h, u_h - \chi \rangle$ is zero when λ_h is zero (noncontact) or $u_h = \chi$ (contact). Hence it encodes an error committed close to the free boundary.

It is trivial to see the following estimation of $\langle \lambda_h, u_h - \chi \rangle$ [104, Lemma 3.2]:

Lemma 6.9 (Lack of Monotonicity: Stationary Case)

$$\langle \lambda_h, u_h - \chi \rangle = \sum_{z \in \mathcal{F}_h} s_z d_z \quad \text{where} \quad d_z := \int_{\omega_z} (u_h - \chi) \psi_z. \quad (6.28)$$

Proof. Because $u_h = \chi$ in ω_z for any $z \in \mathcal{C}_h$ and $s_z = 0$ for $z \in \mathcal{N}_h$, we then get the result. \blacksquare

This new term is due to the nature of the contact problem since it vanishes for problems without constraint. We want to obtain an upper bound for it in terms of more standard error estimators. For contact problems, heuristically speaking, the jump residual of u_h should be a good local error indicator in the noncontact region, as suggested by the well-established a posteriori error theory for linear elliptic equations. On the other hand, in the contact region, the jump residual of $w_h := u_h - \chi$ appears to be appropriate because it matches the localization behavior (0 when full contact). In the transition region associated with the free boundary the estimate is more subtle. It seems to be reasonable to have both jump residuals of u_h and w_h .

We first consider d_z , for $z \in \mathcal{F}_h$. Similar analysis has been performed in different contexts [113, 134]. By the discrete quadratic growth¹ property [113, Lemma 6.4], we obtain that

$$\|w_h\|_{L^\infty(\omega_z)} \lesssim h_z \|J(w_h)\|_{L^\infty(\gamma_z)}, \quad (6.29)$$

because $w_h(z) = 0$ by the definition of the set of free boundary nodes \mathcal{F}_h . Hence, by the definition of d_z (6.28), we have

$$d_z = \int_{\omega_z} w_h \psi_z \leq \|w_h\|_{L^\infty(\omega_z)} \int_{\omega_z} \psi_z \lesssim h_z^{1+d} \|J(w_h)\|_{L^\infty(\gamma_z)}.$$

Using a scaling argument, we can get the following estimation

$$d_z \lesssim h_z^{\frac{3}{2} + \frac{d}{2}} \|J(w_h)\|_{L^2(\gamma_z)}. \quad (6.30)$$

From the definition of the nodal based Lagrange multiplier (6.18), we know that

$$0 \leq -s_z = -\langle r_h, \bar{\psi}_z \rangle, \quad (6.31)$$

where

$$\bar{\psi}_z := \frac{\psi_z}{\int_{\omega_z} \psi_z}$$

¹This is a statement about the quadratic growth of any non-negative function with bounded second derivatives. The continuous quadratic growth property was proved by Baiocchi [10].

is the normalized (in L^1 -norm) basis function for any $z \in \mathcal{P}_h$, i.e. $\int_{\Omega} \bar{\psi}_z = 1$. It is easy to see that

$$\|\bar{\psi}_z\|_{L^2(\omega_z)} \approx h_z^{-\frac{d}{2}} \quad \text{and} \quad \|\bar{\psi}_z\|_{L^2(\gamma_z)} \approx h_z^{-\frac{d+1}{2}}.$$

Consequently, by the Cauchy-Schwarz inequality,

$$-s_z \leq \langle r, \bar{\psi}_z \rangle \lesssim h_z^{-\frac{d+1}{2}} \|J(u_h)\|_{L^2(\gamma_z)} + h_z^{-\frac{d}{2}} \|R(u_h)\|_{L^2(\omega_z)}. \quad (6.32)$$

Now we are in a position to derive an upper bound of the free boundary error estimator term $-s_z d_z$. From the two inequalities (6.30) and (6.32) obtained above, we get

$$\begin{aligned} -s_z d_z &\lesssim h_z \|J(u_h)\|_{L^2(\gamma_z)} \|J(w_h)\|_{L^2(\gamma_z)} + h_z^{3/2} \|R(u_h)\|_{L^2(\omega_z)} \|J(w_h)\|_{L^2(\gamma_z)} \\ &\lesssim h_z \|J(u_h)\|_{L^2(\gamma_z)}^2 + h_z \|J(w_h)\|_{L^2(\gamma_z)}^2 + h_z^2 \|R(u_h)\|_{L^2(\omega_z)}^2. \end{aligned}$$

Remark 6.10 (Convergence Rate) From the inequality above, we find the new free boundary term is at least of the same order as the jump residual and the interior residual terms. Our numerical experiments show this estimate is actually pessimistic. Usually, this term is of higher order than the jump residual.

A similar calculation has been carried out by Bartels and Carstensen [15] in a different context for an averaging error estimator for obstacle problems.

Computable Upper Bound

Based on the last three lemmas, Lemma 6.7, 6.8, and 6.9, we can obtain a computable localized upper bound of the energy error. We first give a couple comments about the assumptions for the upper bound result (Theorem 6.13). Recall that $\rho = 2s$ with $s \in [0, 1]$.

Remark 6.11 ($\mathcal{A}_I u_h$ is in $L^p(\Omega)$) To make sure our estimator is actually finite, we will need $\mathcal{A}_I u_h \in L^p(\omega_z)$. In fact, if $|\hat{k}(\xi)| \leq C(1 + |\xi|)^Y$, then the pseudo-differential operator

$$\mathcal{A}_I : \tilde{W}^{s,p}(\Omega) \rightarrow W^{s-Y,p}(\Omega)$$

is continuous ([133, Theorem XI.2.5]). For piecewise linear finite element function u_h , which is in $\tilde{W}^{1+\frac{1}{p}-\varepsilon,p}(\Omega)$, we have $\mathcal{A}_I u_h \in L^p(\Omega)$ if

$$\frac{1}{p} > Y - 1.$$

Next, we give an example in one-dimension to motivate this property.

Example 6.12 (An 1d Example) Let the kernel function of \mathcal{A}_I be $k(x) = 1/|x|^{1+Y}$ and u_h be a continuous piecewise linear finite element function. It is clear that $\mathcal{A}_I u_h$ have singularities near the end points of each subinterval of the domain. Suppose that $x = 0$ is such an end point (for general case, we can show the same result by shifting the domain). A particular example of u_h and its derivatives close to $x = 0$ is shown in Figure 6.2. Near $x = 0$, the function $(\mathcal{A}_I u_h)(x)$ behaves like $|x|^{-Y+1}$.

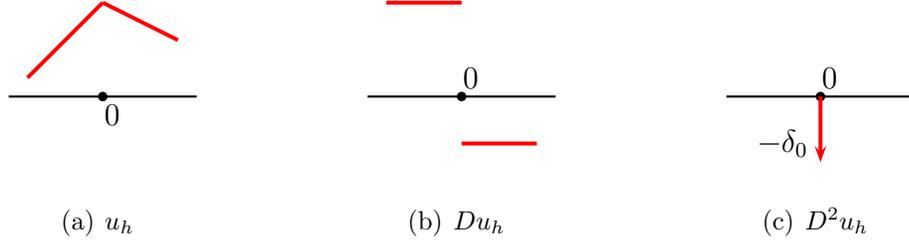


Figure 6.2: Finite element function and its derivatives

For $\mathcal{A}_I u_h$ to be in L^p , we need

$$\int_{|x|<\varepsilon} |x|^{(1-Y)p} dx < \infty.$$

And this inequality holds when $(1 - Y)p > -1$ or $\frac{1}{p} > Y - 1$ which is exactly the condition in the previous remark.

Theorem 6.13 (Upper Bound) Let $f \in L^p(\Omega)$ and $p \geq 1$ satisfy

$$Y - 1 < \frac{1}{p} < \frac{\rho}{2d} + \frac{1}{2}. \quad (6.33)$$

Then we have the following finite upper bound for the error of u_h

$$\|u - u_h\|^2 + \|\lambda - \lambda_h\|_*^2 \lesssim \sum_{z \in \mathcal{P}_h \setminus \mathcal{C}_h} (\eta_z^2 + \xi_z^2) - \sum_{z \in \mathcal{F}_h} s_z d_z, \quad (6.34)$$

where

$$\eta_z^2 := h_z \|J(u_h)\|_{L^2(\gamma_z)}^2 \quad \text{and} \quad \xi_z^2 := h_z^{2s+d-\frac{2d}{p}} \|(R(u_h) - R_z)\psi_z\|_{L^p(\omega_z)}^2. \quad (6.35)$$

Remark 6.14 (Choice of p) For any dimension d , the admissible set of p is nonempty if we have

$$Y < \frac{\rho}{2d} + \frac{3}{2}.$$

For example, if $d = 1$ and $Y = \rho$, the admissible region of p is shown in Figure 6.3. The region marked by gray indicate the possible p for different Y 's. In the gray

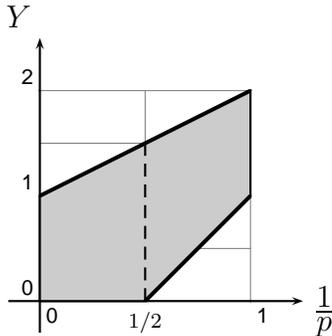


Figure 6.3: Admissible Region of p ($d = 1$)

region left to the dashed line ($0 \leq Y < 3/2$), we could choose $p = 2$; on the other hand, for $3/2 \leq Y < 2$, we need to pick some $1 < p < 2$ in the gray area.

Remark 6.15 (Localization) The space error estimator is fully localized, i.e., there is no contribution from $z \in \mathcal{C}_h$, the discrete contact set. Note that this is consistent with the absence of obstacle approximation error because $\chi_h = \chi$. Likewise, the term $-s_z d_z \geq 0$ contributes only when $z \in \mathcal{F}_h$. One may also wonder whether the sets of full-contact nodes \mathcal{C}_h and free boundary nodes \mathcal{F}_h are good approximations of the actual contact region and free boundary, respectively. We will explore this point further via numerical experiments in Chapter 8.

Proof of Theorem 6.13.

Case i ($\rho = 2$) Recall the localization property (6.20) of \mathcal{G}_h :

$$\langle \mathcal{G}_h, \varphi \rangle = \sum_{z \in \mathcal{P}_h \setminus \mathcal{C}_h} \langle (r_h - s_z) \psi_z, \varphi \rangle \quad \forall \varphi \in \tilde{H}^1(\Omega).$$

By the definitions of the residual $r_h = f - \mathcal{A}u_h$ and the interior residual $R(u_h)$ in (6.16), we have

$$\langle r_h, \varphi\psi_z \rangle = \langle R(u_h), \varphi\psi_z \rangle - \langle c_2 \nabla u_h, \nabla(\varphi\psi_z) \rangle.$$

This, and the definition (6.18) of s_z , give

$$\langle r_h - s_z, \varphi\psi_z \rangle = \langle R(u_h) - R_z, \varphi\psi_z \rangle - \langle c_2 \nabla u_h, \nabla(\varphi\psi_z) \rangle - \langle J_z, \varphi\psi_z \rangle.$$

For each node $z \in \mathcal{P}_h \setminus \mathcal{C}_h$, by the definition of J_z , we get for the third term on the right-hand side:

$$- \langle J_z, \varphi\psi_z \rangle = C_{\varphi,z} \langle c_2 \nabla u_h, \nabla\psi_z \rangle,$$

where $C_{\varphi,z} = \langle \varphi, \psi_z \rangle / \langle 1, \psi_z \rangle$ is a weighted average for interior nodes $z \in \mathcal{P}_h \cap \Omega$ and $C_{\varphi,z} = 0$ for boundary nodes $z \in \mathcal{P}_h \cap \Gamma$.

Hence integration by parts gives

$$\langle \mathcal{G}_h, \varphi \rangle = \sum_{z \in \mathcal{P}_h \setminus \mathcal{C}_h} \langle (R(u_h) - R_z)\psi_z, \varphi \rangle + \sum_{z \in \mathcal{P}_h \setminus \mathcal{C}_h} \int_{\gamma_z} J(u_h)(\varphi - C_{\varphi,z})\psi_z. \quad (6.36)$$

Applying the Cauchy-Schwarz inequality and taking

$$G = \sum_{z \in \mathcal{P}_h \setminus \mathcal{C}_h} g_z \quad \text{with} \quad g_z = (R(u_h) - R_z)\psi_z$$

in Lemma 6.7 and 6.8 (since $\int_{\omega_z} g_z = 0$ for $z \in \mathcal{P}_h \cap \Omega$), we then obtain

$$\begin{aligned} \sum_{z \in \mathcal{P}_h \setminus \mathcal{C}_h} \langle (R(u_h) - R_z)\psi_z, \varphi \rangle \\ \lesssim \left(\sum_{z \in \mathcal{P}_h \setminus \mathcal{C}_h} h_z^{2+d(1-2/p)} \|(R(u_h) - R_z)\psi_z\|_{L^p(\omega_z)}^2 \right)^{\frac{1}{2}} \|\varphi\|_{\tilde{H}^1(\Omega)}. \end{aligned}$$

For the second part of (6.36), we use standard scaling argument based on a trace theorem and the Poincaré's inequality to get (see also [108, Theorem 3.6])

$$\left| \sum_{z \in \mathcal{P}_h \setminus \mathcal{C}_h} \int_{\gamma_z} J(u_h)(\varphi - C_{\varphi,z})\psi_z \right| \lesssim \left(\sum_{z \in \mathcal{P}_h \setminus \mathcal{C}_h} h_z \|J(u_h)\|_{L^2(\gamma_z)}^2 \right)^{\frac{1}{2}} \|\varphi\|_{\tilde{H}^1(\Omega)}.$$

Then, by applying Lemma 6.2 and 6.9, we get the result.

Case ii ($\rho < 2$) In this case, the constant $c_2 = 0$ and in turn the jump term $J(u_h)$ vanishes. The upper bound follows directly from the proof above by taking $c_2 = 0$. ■

6.4.2 Lower Bound

Inspired by the global lower bound in Lemma 6.2, we first prove that the dual norm $\|\mathcal{G}_h\|_*$ can be bounded from below by the sum of local dual norms $\|\mathcal{G}_h\|_{H^{-s}(\omega_z)}$; then we show that the local error estimators η_z and ξ_z are lower bounds of $\|\mathcal{G}_h\|_{H^{-s}(\omega_z)}$ up to an oscillation term for each node $z \in \mathcal{P}_h$.

From Local to Global

We first prove a crucial lemma with the help of the norm equivalence result (Proposition 2.5).

Lemma 6.16 (Localized Lower Bound of the Dual Norm) *Let $G \in H^{-s}(\Omega)$. For $s \in [0, 1]$ there holds*

$$\sum_{z \in \mathcal{P}_h} \|G\|_{H^{-s}(\omega_z)}^2 \lesssim \|G\|_{H^{-s}(\Omega)}^2.$$

Proof. We first partition the set of nodes $\mathcal{P}_h = \bigcup_{i=1}^M \mathcal{P}_h^i$ such that the intersections $\text{int } \omega_x \cap \text{int } \omega_y = \emptyset$, for any x and y in each \mathcal{P}_h^i for $i = 1, \dots, M$.

For each $z \in \mathcal{P}_h$, let $\varphi_z \in \tilde{H}^s(\omega_z)$ satisfies

$$\langle G, \varphi_z \rangle = \|G\|_{H^{-s}(\omega_z)}^2 \quad \text{and} \quad \|\varphi_z\|_{\tilde{H}^s(\omega_z)} = \|G\|_{H^{-s}(\omega_z)}.$$

Similar to the argument as in the heat equation case [16], we have

$$\sum_{z \in \mathcal{P}_h^i} \|G\|_{H^{-s}(\omega_z)}^2 = \sum_{z \in \mathcal{P}_h^i} \langle G, \varphi_z \rangle = \langle G, \sum_{z \in \mathcal{P}_h^i} \varphi_z \rangle \leq \|G\|_{H^{-s}(\Omega)} \left\| \sum_{z \in \mathcal{P}_h^i} \varphi_z \right\|_{\tilde{H}^s(\Omega)}.$$

On the other hand, using Proposition 2.5, we obtain

$$\left\| \sum_{z \in \mathcal{P}_h^i} \varphi_z \right\|_{\tilde{H}^s(\Omega)}^2 \leq \sum_{z \in \mathcal{P}_h^i} \|\varphi_z\|_{\tilde{H}^s(\Omega)}^2 = \sum_{z \in \mathcal{P}_h^i} \|\varphi_z\|_{\tilde{H}^s(\omega_z)}^2 = \sum_{z \in \mathcal{P}_h^i} \|G\|_{H^{-s}(\omega_z)}^2.$$

The last two inequalities yield that

$$\sum_{z \in \mathcal{P}_h^i} \|G\|_{H^{-s}(\omega_z)}^2 \lesssim \|G\|_{H^{-s}(\Omega)}^2 \quad i = 1, \dots, M.$$

Hence the result. \blacksquare

Lower Bound in Terms of Jump Residual

Now we prove the lower bound of local dual norms in terms of the jump residual. Recall that $\eta_z = 0$ for Case II and III. The only non-trivial case then is the Case I ($\rho = 2$) when $c_2 \neq 0$.

Lemma 6.17 (Lower Bound of Local Dual Norms: Jump Residual) *We have the following lower bound of the local dual norm provided $\rho = 2$*

$$\eta_z^2 \lesssim \|\mathcal{G}_h\|_{H^{-1}(\omega_z)}^2 + \sum_{x \in (\mathcal{P}_h \setminus \mathcal{C}_h) \cap \omega_z} \xi_x^2 \quad \forall z \in \mathcal{P}_h \setminus \mathcal{C}_h.$$

Proof. We construct test functions explicitly as in [72]. Let $S \subset \gamma_z$ be a generic side (face) and ω_S be the union of the elements $\tau \in \mathcal{T}$ sharing side S . Consider the classical bubble functions [135]

$$b_S := \prod_{y \in \mathcal{P}_h \cap S} \psi_y \quad \text{and} \quad b_\tau := \prod_{y \in \mathcal{P}_h \cap \tau} \psi_y; \quad (6.37)$$

and define $\hat{\varphi}_S$ as

$$\hat{\varphi}_S := b_S - \sum_{y \in \mathcal{P}_h \cap \tau, \tau \subset \omega_S} \beta_{\tau,y} b_\tau \psi_y,$$

with constant coefficients $\beta_{\tau,y} \in \mathbb{R}$. We have enough freedom to choose the constants $\beta_{\tau,y}$ such that $\langle \hat{\varphi}_S, \psi_y \rangle = 0$ for $y \in \mathcal{P}_h \cap \omega_S$. Furthermore, it is clear that $\text{supp } \hat{\varphi}_S = \omega_S$ and $\hat{\varphi}_S|_S = b_S$. We can see by scaling argument and the trace theorem that

$$\|\hat{\varphi}_S\|_{\tilde{H}^1(\omega_S)} \lesssim h_z^{-\frac{1}{2}} \|\hat{\varphi}_S\|_{L^2(S)} \quad \text{and} \quad \|\hat{\varphi}_S\|_{L^q(\omega_S)} \lesssim h_z^{\frac{1}{2} + \frac{d}{2} - \frac{d}{p}} \|\hat{\varphi}_S\|_{L^2(S)}. \quad (6.38)$$

We finally set the test function $\varphi_S := J(u_h)|_S \hat{\varphi}_S$ and observe that

$$\begin{aligned} \|J(u_h)\|_{L^2(S)}^2 &\lesssim \sum_{y \in \mathcal{P}_h \cap S} \int_S J(u_h)^2 b_S \psi_y \\ &= \sum_{y \in \mathcal{P}_h \cap S} \int_S J(u_h)^2 \hat{\varphi}_S \psi_y = \sum_{y \in \mathcal{P}_h \cap S} \int_S J(u_h) \varphi_S \psi_y. \end{aligned}$$

From (6.36) and (6.38), since $C_{\varphi,z} = 0$ for $\varphi = \hat{\varphi}_S$, we see that

$$\begin{aligned}
& \|J(u_h)\|_{L^2(S)}^2 \lesssim \langle \mathcal{G}_h, \varphi_S \rangle - \sum_{x \in (\mathcal{P}_h \setminus \mathcal{C}_h) \cap \omega_S} \langle (R(u_h) - R_x)\psi_x, \varphi_S \rangle \\
& \leq \|\mathcal{G}_h\|_{H^{-1}(\omega_S)} \|\varphi_S\|_{\tilde{H}^1(\omega_S)} + \sum_{x \in (\mathcal{P}_h \setminus \mathcal{C}_h) \cap \omega_S} \|(R(u_h) - R_x)\psi_x\|_{L^p(\omega_S)} \|\varphi_S\|_{L^q(\omega_S)} \\
& \lesssim \left(\|\mathcal{G}_h\|_{H^{-1}(\omega_S)} + \sum_{x \in (\mathcal{P}_h \setminus \mathcal{C}_h) \cap \omega_S} h_x^{1+\frac{d}{2}-\frac{d}{p}} \|(R(u_h) - R_x)\psi_x\|_{L^p(\omega_S)} \right) h_z^{\frac{1}{2}} \|J(u_h)\|_{L^2(S)}.
\end{aligned}$$

By adding contributions from each side in γ_z , we obtain the result. \blacksquare

Remark 6.18 (Local Lower Bound) In case \mathcal{A} is a second order continuous and coercive differential operator ($c_I = 0$), it has been shown that $\|\mathcal{G}_h\|_{H^{-1}(\omega_z)}$ is a lower bound of the local error up to an oscillation term; see [104, Lemma 3.7]. This, in turn, yields local efficiency. Unfortunately, it is not true for global operators \mathcal{A} in our context.

Simply applying the previous two lemmas (Lemma 6.16 and 6.17) and the abstract lower bound (Lemma 6.2), we get the following global lower bound:

Theorem 6.19 (Global Lower Bound: Jump Residual) *The following global lower bound holds provided $\rho = 2$*

$$\sum_{z \in \mathcal{P}_h \setminus \mathcal{C}_h} \eta_z^2 \lesssim \|u - u_h\|^2 + \|\lambda - \lambda_h\|_*^2 + \sum_{z \in \mathcal{P}_h \setminus \mathcal{C}_h} \xi_z^2.$$

Lower Bound in Terms of Interior Residual ($p = 2$)

Now we prove the lower bound in terms of ξ_z up to an oscillation term for the rest two cases ($\rho < 2$). For simplicity, we first present results for $p = 2$ and general result can be studied analogously.

An interesting observation is that, in Case II and III, the estimator ξ_z behaves very differently than in Case I. We observe that ξ_z in numerical experiments is of higher order than the real energy error for Case I; see Table 8.14. However, in the remaining two cases, it has the same order as the energy error as we can expect from the following lemma; see Table 8.11.

Since $c_2 = 0$, we have $R(u_h)$ is actually the full residual r_h and, because of Galerkin orthogonality, we have

$$R_z = \langle R(u_h), \psi_z \rangle / \langle 1, \psi_z \rangle = 0 \quad \forall z \in \mathcal{N}_h$$

Hence, in this case, we have

$$\xi_z = h_z^s \|R(u_h)\psi_z\|_{L^2(\omega_z)} \leq h_z^s \|R(u_h)\|_{\psi_z} =: \tilde{\xi}_z.$$

Here, for short, we denote the weighted L^2 -norm

$$\|v\|_{\psi_z} := \left(\int_{\omega_z} v^2 \psi_z \right)^{\frac{1}{2}}.$$

We first introduce some new notation. Let

$$\tilde{R}_z := \sum_{x \in (\mathcal{P}_h \setminus \mathcal{C}_h) \cap \omega_z} (R(u_h) - R_x) \psi_x. \quad (6.39)$$

If all of $x \in \mathcal{P}_h \cap \omega_z$ are in the noncontact set (corresponding to a linear elliptic equation), then $\tilde{R}_z = R(u_h)$.

For each node $z \in \mathcal{P}_h$, define $\hat{\omega}_z$ to be the corresponding reference patch of ω_z . On the reference patch, we can define a finite dimensional function space $\hat{\mathbb{P}}_z$. Let the finite dimensional space \mathbb{P}_z on ω_z be a scaling transformation of the space $\hat{\mathbb{P}}_z$. For any node $z \in \mathcal{P}_h \setminus \mathcal{C}_h$, we define the oscillation to be

$$\text{osc}_z^2 := \inf_{P_z \in \mathbb{P}_z} h_z^{2s} \left(\|(R(u_h) - R_z) - P_z\|_{\psi_z}^2 + \|\tilde{R}_z - P_z\|_{\psi_z}^2 \right). \quad (6.40)$$

Remark 6.20 (Standard Oscillation Term) If all nodes in ω_z are noncontact, then it is clear that $R_z = 0$ and $\tilde{R}_z = R(u_h)$. In this case, this oscillation term reduces to the standard oscillation term $\text{osc}_z = h_z^s \inf_{P_z \in \mathbb{P}_z} \|R(u_h) - P_z\|_{\psi_z}$ for linear elliptic equations [108].

Lemma 6.21 (Lower Bound of Local Dual Norms: Interior Residual) *Let the residual $R(u_h) \in L^2(\Omega)$. Then, in Case II and III, there exists a constant $C > 0$ such that*

$$\xi_z^2 \leq C \|\mathcal{G}_h\|_{H^{-s}(\omega_z)}^2 + 2\text{osc}_z^2 \quad z \in \mathcal{P}_h \setminus \mathcal{C}_h. \quad (6.41)$$

Proof. For any $\varphi \in \tilde{H}^s(\omega_z)$, it is easy to see that

$$\langle \mathcal{G}_h, \varphi_z \rangle = \sum_{x \in (\mathcal{P}_h \setminus \mathcal{C}_h) \cap \omega_z} \langle (R(u_h) - R_x) \psi_x, \varphi_z \rangle = \langle \tilde{R}_z, \varphi_z \rangle.$$

Then we add and subtract a function $P_z \in \mathbb{P}_z$ to get

$$\langle \mathcal{G}_h, \varphi_z \rangle = \langle P_z, \varphi_z \rangle + \langle \tilde{R}_z - P_z, \varphi_z \rangle.$$

Now take the test function $\varphi_z := P_z \psi_z \in \tilde{H}^s(\omega_z)$. Then we have

$$\|P_z\|_{\psi_z}^2 = \langle P_z, \varphi_z \rangle = \langle \mathcal{G}_h, \varphi_z \rangle - \langle \tilde{R}_z - P_z, \varphi_z \rangle.$$

Hence, by the Cauchy-Schwarz inequality, we arrive at

$$\|P_z\|_{\psi_z}^2 \leq \|\mathcal{G}_h\|_{H^{-s}(\omega_z)} \cdot \|\varphi_z\|_{\tilde{H}^s(\omega_z)} + \|\tilde{R}_z - P_z\|_{\psi_z} \cdot \|P_z\|_{\psi_z}. \quad (6.42)$$

Since φ_z is finite dimensional and has the scaling property, we have the inverse estimation $\|\varphi_z\|_{\tilde{H}^s(\omega_z)} \leq Ch_z^{-s} \|\varphi_z\|_{L^2(\omega_z)} \leq Ch_z^{-s} \|P_z\|_{\psi_z}$ by using the interpolation argument and noticing that $\psi_z \leq 1$. Applying the inverse estimation to the inequality (6.42), we immediately obtain

$$\|P_z\|_{\psi_z} \leq Ch_z^{-s} \|\mathcal{G}_h\|_{H^{-s}(\omega_z)} + \|\tilde{R}_z - P_z\|_{\psi_z}. \quad (6.43)$$

On the other hand, the definition of the interior error estimator gives

$$\begin{aligned} h_z^{2s} \|(R(u_h) - R_z) \psi_z\|_{L^2(\omega_z)}^2 &\leq h_z^{2s} \|R(u_h) - R_z\|_{\psi_z}^2 \\ &\leq 2h_z^{2s} \left(\|P_z\|_{\psi_z}^2 + \|(R(u_h) - R_z) - P_z\|_{\psi_z}^2 \right). \end{aligned} \quad (6.44)$$

Then combining (6.43) and (6.44), we get the result. \blacksquare

Remark 6.22 (Convergence Rate of the Oscillation Term) For elliptic differential equations, the oscillation term is usually of higher order than the error estimator. So asymptotically, it will be small and can be ignored. However, for problems with an integro-differential operator, this is not the case. For example, if we take $\mathcal{A} = \mathcal{A}_I$ and the kernel $k(x) = 1/|x|^2$ for $d = 1$, we notice that the residual $r_h = f - \mathcal{A}_I u_h$ has logarithmic singularities towards the end points of each interval;

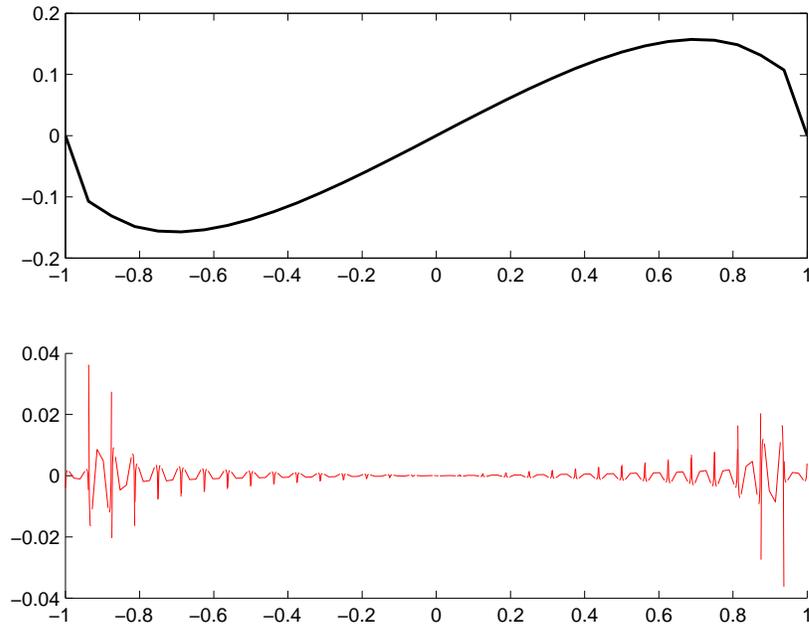


Figure 6.4: Residual is singular close to the end points of each interval. Upper: discrete solution u_h ; lower: residual $r = f - \mathcal{A}u_h$.

see Figure 6.4. Even worse, as we refine the mesh, this oscillation does not go away. Hence it is not clear whether there is any chance to have the oscillation to be of higher order. In fact, numerical experiments show that it is not the case. If you choose \mathbb{P}_z to be piecewise linear (or even quadratic) polynomials, the oscillation is of the same order as the estimator ξ_z .

Remark 6.23 (Choice of \mathbb{P}_z) In the standard oscillation term for linear finite element method, usually it is enough to take \mathbb{P}_z to be a constant function. However, constants will not help us in general in the current context. This is because the optimal value of

$$\inf_{P_z \in \text{Const}} \|R(u_h) - P_z\|_{\psi_z}$$

occurs when $P_z = R_z$, which is zero due to the Galerkin orthogonality. So we need piecewise linear polynomial space for \mathbb{P}_z at least. We shall discuss this in Chapter 8 with specific test examples.

With the help of this lemma and Lemma 6.16, we obtain, by summing local terms together, that:

Theorem 6.24 (Global Lower Bound: Interior Residual ($p = 2$)) *If $R(u_h) = f - \mathcal{A}u_h \in L^2(\Omega)$, we have the global lower bound*

$$\sum_{z \in \mathcal{P}_h \setminus \mathcal{C}_h} \left(\xi_z^2 - 2\text{osc}_z^2 \right) \lesssim \|u - u_h\|^2 + \|\lambda - \lambda_h\|_*^2.$$

Lower Bound in Terms of Interior Residual ($1 < p < 2$)

Above we only discuss the case when $p = 2$. For $3/2 < Y < 2$, in the error estimator ξ_z , we need in general $1 < p < 2$. Similar to the discussion for $p = 2$, we can also define weighted L^p -norm

$$\|v\|_{p, \psi_z} := \left(\int_{\omega_z} |v|^p \psi_z \right)^{1/p}$$

and the oscillation term

$$\text{osc}_{p,z}^2 := \inf_{P_z \in \mathbb{P}_z} h_z^{2s+d-\frac{2d}{p}} \left(\|(R(u_h) - R_z) - P_z\|_{p, \psi_z}^2 + \|\tilde{R}_z - P_z\|_{p, \psi_z}^2 \right).$$

This is consistent with the definition before for $p = 2$.

Let $\varphi_z := P_z^{\frac{p}{q}} \psi_z$. Using the same argument as in the derivation of (6.42), it is easy to see that

$$\|P_z\|_{p, \psi_z}^p = \int_{\omega_z} P_z \varphi_z = \langle \mathcal{G}_h, \varphi_z \rangle - \langle \tilde{R}_z - P_z, \varphi_z \rangle.$$

For the second term on the right-hand side, we have

$$\int_{\omega_z} (\tilde{R}_z - P_z) P_z^{\frac{p}{q}} \psi_z = \int_{\omega_z} \left[(\tilde{R}_z - P_z) \psi_z^{\frac{1}{p}} \right] \cdot \left[P_z^{\frac{p}{q}} \psi_z^{\frac{1}{q}} \right].$$

By the Cauchy-Schwarz inequality, it follows that

$$\|P_z\|_{p, \psi_z}^p \leq \|\mathcal{G}_h\|_{H^{-s}} \cdot \|\varphi_z\|_{\tilde{H}^s(\omega_z)} + \|\tilde{R}_z - P_z\|_{p, \psi_z} \cdot \|\varphi_z\|_{q, \psi_z}. \quad (6.45)$$

On the other hand, using the interpolation argument and inverse estimate as in Lemma 6.8, we have

$$\|\varphi_z\|_{\tilde{H}^s(\omega_z)} \lesssim h_z^{-s-\frac{d}{q}+\frac{d}{2}} \|\varphi_z\|_{L^q(\omega_z)} = h_z^{-s-\frac{d}{q}+\frac{d}{2}} \|P_z^{\frac{p}{q}} \psi_z\|_{L^q(\omega_z)}$$

Because $1 < p < 2$, so $2 < q < \infty$, we have $\frac{1}{q} < \frac{1}{2} < 1$ and in turn

$$\|P_z^{\frac{p}{q}} \psi_z\|_{L^q(\omega_z)} \leq \|P_z^{\frac{p}{q}} \psi_z^{\frac{1}{q}}\|_{L^q(\omega_z)} = \|P_z^{\frac{p}{q}}\|_{q, \psi_z}.$$

Plugging this inequality into (6.45), we then get

$$\|P_z\|_{p, \psi_z}^p \leq Ch_z^{-s-\frac{d}{q}+\frac{d}{2}} \|\mathcal{G}_h\|_{H^{-s}(\omega_z)} \|P_z^{\frac{p}{q}}\|_{q, \psi_z} + \|\tilde{R}_z - P_z\|_{p, \psi_z} \|P_z^{\frac{p}{q}}\|_{q, \psi_z}.$$

Notice that $\|P_z^{\frac{p}{q}}\|_{q, \psi_z} = \|P_z\|_{p, \psi_z}^{\frac{p}{q}}$ by simple calculation. Then we obtain that

$$\|P_z\|_{p, \psi_z} \leq Ch_z^{-s-\frac{d}{q}+\frac{d}{2}} \|\mathcal{G}_h\|_{H^{-s}(\omega_z)} + \|R - P_z\|_{p, \psi_z}$$

and in turn the lower bound

$$\tilde{\xi}_z^2 := h_z^{2s+d-\frac{2d}{p}} \|R(u_h) - R_z\|_{p, \psi_z}^2 \leq C \|\mathcal{G}_h\|_{H^{-s}(\omega_z)}^2 + 2\text{osc}_{z,p}^2. \quad (6.46)$$

We summarize the discussion above in the following theorem:

Theorem 6.25 (Global Lower Bound: Interior Residual ($1 < p < 2$)) *If the residual $R(u_h) = f - \mathcal{A}u_h \in L^p(\Omega)$, we have the global lower bound*

$$\sum_{z \in \mathcal{P}_h \setminus \mathcal{C}_h} \left(\xi_z^2 - 2\text{osc}_{p,z}^2 \right) \lesssim \|u - u_h\|^2 + \|\lambda - \lambda_h\|_*^2.$$

Non-negative Lower Bound

For elliptic PDEs, the oscillation term osc_z is usually of higher order than the error estimator. However it is not the case for problems with singular integral operators. The question is whether we get a non-negative lower bound or we just get some trivial inequality with a negative quantity in (6.24). For elliptic equations, a more careful analysis gives a non-negative lower bound of the error.

Lemma 6.26 (Non-trivial Lower Bound) *There exists a constant $C > 0$ such that*

$$0 \leq \sum_{z \in \mathcal{P}_h \setminus \mathcal{C}_h} \left(\tilde{\xi}_z^2 - \text{osc}_z^2 \right) \lesssim \|u - u_h\|^2, \quad (6.47)$$

where the error estimator and the oscillation term read

$$\begin{aligned} \tilde{\xi}_z &:= h_z^s \|R(u_h)\|_{\psi_z} \\ \text{osc}_z &:= h_z^s \inf_{P_z \in \mathbb{P}_z} \|R(u_h) - P_z\|_{\psi_z}. \end{aligned}$$

Proof. It is clear that, to show (6.47), we only need to prove

$$\|R(u_h)\|_{H^{-s}(\omega_z)}^2 \geq Ch_z^{2s} \left(\|R(u_h)\|_{\psi_z}^2 - \inf_{P_z \in \mathbb{P}_z} \|R(u_h) - P_z\|_{\psi_z}^2 \right). \quad (6.48)$$

Let $P_z \in \mathbb{P}_z$ be the best approximation of $R(u_h)$ in \mathbb{P}_z with respect to $\|\cdot\|_{\psi_z}$ -norm. Then, by orthogonality, we have

$$\|R(u_h)\|_{\psi_z}^2 - \|R(u_h) - P_z\|_{\psi_z}^2 = \|P_z\|_{\psi_z}^2.$$

Hence we only need to prove that $\|R(u_h)\|_{H^{-s}(\omega_z)} \gtrsim h_z^s \|P_z\|_{\psi_z} \geq 0$. In order to show this, we first prove the corresponding inequality by exploring the equivalence of norms on finite dimensional spaces and then we prove a scaling inequality to close the gap.

Step 1. Let $\hat{\omega}_z$ be the reference star with size 1 and \hat{v} be the function v on reference star after transformation. Because $\hat{\mathbb{P}}_z$ is finite dimensional, then it is well-known that

$$\|\hat{P}_z\|_{\hat{\psi}_z} \lesssim \|\hat{P}_z\|_{H^{-s}(\hat{\omega}_z)}.$$

Let $\{\hat{b}_i\}_{i=1}^M$ be the orthogonal basis of $\hat{\mathbb{P}}_z$ on $\hat{\omega}_z$. Then

$$\hat{P}_z = \sum_{i=1}^M \langle \hat{R}(\hat{u}_h), \hat{b}_i \hat{\psi}_z \rangle \hat{b}_i.$$

Consequently,

$$\|\hat{P}_z\|_{\hat{\psi}_z} \lesssim \|\hat{P}_z\|_{H^{-s}(\hat{\omega}_z)} \lesssim \sum_{i=1}^M \langle \hat{R}(\hat{u}_h), \hat{b}_i \hat{\psi}_z \rangle \|\hat{b}_i\|_{H^{-s}(\hat{\omega}_z)} \lesssim \|\hat{R}(\hat{u}_h)\|_{H^{-s}(\hat{\omega}_z)} \quad (6.49)$$

by the Cauchy-Schwarz inequality.

Step 2. To obtain the inequality (6.48), we need a scaling inequality for H^{-s} -norm. By definition, we have

$$\|R(u_h)\|_{H^{-s}(\omega_z)} := \sup_{\|v\|_{\hat{H}^s(\omega_z)}=1} \int_{\omega_z} R(u_h)v = \sup_{\|v\|_{\hat{H}^s(\omega_z)}=1} h_z^d \int_{\hat{\omega}_z} \hat{R}(\hat{u}_h)\hat{v}. \quad (6.50)$$

On the other hand, the first Poincaré inequality gives

$$\begin{aligned} \|v\|_{\hat{H}^1(\omega_z)} &= \left(\|v\|_{L^2(\omega_z)}^2 + \|\nabla v\|_{L^2(\omega_z)}^2 \right)^{1/2} \leq (Ch_z^2 + 1)^{1/2} \|\nabla v\|_{L^2(\omega_z)} \\ &\lesssim |v|_{\hat{H}^1(\omega_z)} \approx h_z^{-1+\frac{d}{2}} |\hat{v}|_{\hat{H}^1(\hat{\omega}_z)} \leq h_z^{-1+\frac{d}{2}} \|\hat{v}\|_{\hat{H}^1(\hat{\omega}_z)}. \end{aligned}$$

Furthermore, it is easy to see that $\|v\|_{L^2(\omega_z)} \leq h_z^{\frac{d}{2}} \|v\|_{L^2(\hat{\omega}_z)}$. Hence, interpolation argument on the scaling operator $\hat{T} : \hat{v} \rightarrow v$ gives the scaling inequality

$$h_z^{s-\frac{d}{2}} \|v\|_{\tilde{H}^s(\omega_z)} \lesssim \|\hat{v}\|_{\tilde{H}^s(\hat{\omega}_z)}. \quad (6.51)$$

Whence, using (6.50) and (6.51), we arrive at

$$\|\hat{R}(\hat{u}_h)\|_{H^{-s}(\hat{\omega}_z)} = \sup_{\|\hat{v}\|_{\tilde{H}^s(\hat{\omega}_z)}=1} \int_{\hat{\omega}_z} \hat{R}(\hat{u}_h)\hat{v} \lesssim h_z^{-s-\frac{d}{2}} \|R(u_h)\|_{H^{-s}(\omega_z)}. \quad (6.52)$$

Step 3. Applying the results obtained from the previous two steps, we obtain that

$$h_z^{-d} \|P_z\|_{\psi_z}^2 \lesssim \|\hat{P}_z\|_{\hat{\psi}_z}^2 \lesssim \|\hat{R}(\hat{u}_h)\|_{H^{-s}(\hat{\omega}_z)} \lesssim h_z^{-2s-d} \|R(u_h)\|_{H^{-s}(\omega_z)}^2.$$

Hence we get the lower bound (6.47). ■

We now summarize this section with a few remarks.

Remark 6.27 (Remarks on the Interior Residual) From the definition of R_z , we can see that $(R(u_h) - R_z)\psi_z$ has mean value 0. The term ξ_z behaves very differently in the case when $c_2 = 0$ than the case $c_2 \neq 0$.

- In Case I, the numerical results (Table 8.14) suggests that ξ_z is of higher order compared with the jump residual term η_z . In fact, this is not surprising at all. For example, in a special case when $\mathcal{A} = -\Delta$ (Dirichlet obstacle problem without integral operator),

$$\xi_z^2 = h_z^2 \|(f - f_z)\psi_z\|_{L^2(\omega_z)}^2$$

which is exactly the usual data oscillation term [106, 107, 108].

- On the other hand, in Case II and III, we have $0 < \rho < 2$ and the jump residual term η_z vanishes because $c_2 = 0$ and $R(u_h) = r_h$ is actually the full residual. Indeed, ξ_z does not behave like the oscillation anymore. If the constraint is nonactive, the constant R_z is zero for every node $z \in \mathcal{P}_h$. In these two cases, the term ξ_z is of the same order as the energy error in our numerical experiments (see Table 8.11).

Remark 6.28 (Other Types of Error Estimators) Several different types of error estimators have been developed for elliptic PDEs. They are residual-type estimators, hierarchical-type estimators, estimators based on local problems, and estimators based on average. See [135, 2] for details. From the discussion before, it has been shown, using an energy method, we can derive a localized residual-type upper bound of the global dual norm $\|\mathcal{G}_h\|_*$. In [127], Siebert and Veerer gave an error estimator of hierarchical-type and proved convergence of the adaptive algorithm using this estimator; they treat the contact and noncontact cases in a global sense in the upper bound proof which leads to a gap between the upper and lower bounds. Estimators based on averaging technique are considered in [15].

6.5 Time-dependent Problems

A posteriori error estimates as well as adaptive methods for linear parabolic equations have been discussed by many researchers since early 80's; for example, Bieterman and Babuška [20, 21] for 1d problems. In 90's, Erickson, Johnson, and Larsson made systematic efforts to develop adaptive methods for linear and nonlinear parabolic equations [57, 58, 59, 60, 61]. More recent advances include [119, 136, 48, 16]. A new approach based on elliptic reconstruction technique was proposed by Makridakis and Nochetto [96] for semidiscrete problems and by Lakkis and Makridakis [90] for fully-discrete numerical methods.

In spite of all above, adaptive methods for parabolic problems are less understood, especially for nonlinear problems like PVI. In this section, we consider a special case of evolution problem (Problem 4.7). To explain the idea, for the moment, we assume that χ does not change in time and it is piecewise linear in space. Furthermore, we assume that the underlying finite element space does not change in time (the mesh \mathcal{T} fixed).

Given the initial solution $u_0 \in \mathcal{K} := \{v \in \tilde{H}^s(\Omega) \mid v \geq \chi, \text{ a.e. } \Omega\}$, finding a solution $u \in L^2(0, T; L^2(\Omega)) \cap H^1(0, T; H^{-s}(\Omega))$ such that $u \in \mathcal{K}$ and

$$\langle \partial_t u + \mathcal{A}u, u - v \rangle \leq \langle f, u - v \rangle \quad \forall v \in \mathcal{K}. \quad (6.53)$$

We apply the fully-discrete numerical scheme (backward Euler linear finite element), Method 4.13.

6.5.1 Lagrange Multiplier and Galerkin Functional

We now define the continuous and discrete Lagrange multipliers and Galerkin functional for the time-dependent problem analogous to the stationary case in §6.4. We define the Lagrange multiplier λ as follows

$$\lambda(t) := f(t) - \partial_t u(t) - \mathcal{A}u(t) \in H^{-s}(\Omega), \quad t \in [0, T]. \quad (6.54)$$

At each time t_n , $n = 1, \dots, N$, we define the residual

$$r_h^n := F^n - \delta U_h^n - \mathcal{A}U_h^n,$$

and split the set of all nodes \mathcal{P}_h into three disjoint sets

$$\mathcal{P}_h = \mathcal{N}_h^n \cup \mathcal{C}_h^n \cup \mathcal{F}_h^n,$$

where noncontact, full-contact, and free boundary sets are given by

$$\mathcal{N}_h^n := \{z \in \mathcal{P}_h \mid U_h^n > \chi(t_n) \text{ in int } \omega_z\}, \quad (6.55a)$$

$$\mathcal{C}_h^n := \{z \in \mathcal{P}_h \mid U_h^n = \chi(t_n) \text{ and } r_h^n \leq 0 \text{ in } \omega_z\}, \quad (6.55b)$$

$$\mathcal{F}_h^n := \mathcal{P}_h \setminus (\mathcal{N}_h^n \cup \mathcal{C}_h^n). \quad (6.55c)$$

Notice that we define these sets for general $\chi \in H^1(\mathcal{Q})$ instead of restricting ourselves to the time-independent χ because we will discuss the general case in the next section. In this section, $\chi(t_n) = \chi$ for $n = 1, \dots, N$.

Then we can define the discrete Lagrange multiplier Λ_h^n as

$$\Lambda_h^n := \sum_{z \in \mathcal{C}_h^n} r_h^n \psi_z + \sum_{z \in \mathcal{P}_h \setminus \mathcal{C}_h^n} s_z^n \psi_z, \quad (6.56)$$

with

$$s_z^n := \begin{cases} \frac{\langle r_h^n, \psi_z \rangle}{\langle 1, \psi_z \rangle}, & z \in \mathcal{P}_h \cap \Omega \\ 0, & z \in \mathcal{P}_h \cap \Gamma. \end{cases} \quad (6.57)$$

At time t_n , the Galerkin functional is defined by

$$\mathcal{G}_h^n := r_h^n - \Lambda_h^n = \sum_{z \in \mathcal{P}_h \setminus \mathcal{C}_h} (r_h^n - s_z^n) \psi_z \quad (6.58)$$

which satisfies the localization property as well as the mean value property, i.e. $\langle (r_h^n - s_z^n) \psi_z, 1 \rangle = 0$. Notice that these properties were crucial to obtain a localized upper bound for elliptic problems in §6.4.

6.5.2 Abstract Error Bounds

Now we are ready to introduce the main steps to treat time-dependent problems. Let $\bar{\mathcal{G}}_h$ be the piecewise constant (in time) Galerkin functional (6.58), i.e.

$$\bar{\mathcal{G}}_h = \bar{F} - \delta_t U_h - \mathcal{A} \bar{U}_h - \bar{\Lambda}_h.$$

This and the definition of λ (6.54) give, for any $\varphi \in \tilde{H}^s(\Omega)$, that

$$\langle \bar{\mathcal{G}}_h, \varphi \rangle = \langle \mathcal{A}(u - \bar{U}_h), \varphi \rangle + \langle (\partial_t u - \delta_t U_h) + (\lambda - \bar{\Lambda}_h), \varphi \rangle - \langle f - \bar{F}, \varphi \rangle. \quad (6.59)$$

Taking $\varphi = u - U_h$ in (6.59) and applying Lemma 2.11, we get

$$\begin{aligned} & \frac{1}{2} \frac{d}{dt} \|u - U_h\|_{L^2(\Omega)}^2 + \frac{1}{4} \left(\| \|u - \bar{U}_h\| \|^2 + \| \|u - U_h\| \|^2 \right) \\ & \leq 2\gamma^2 \| \| \bar{U}_h - U_h \| \|^2 - \langle \lambda - \bar{\Lambda}_h, u - U_h \rangle + \langle \bar{\mathcal{G}}_h, u - U_h \rangle + \langle f - \bar{F}, u - U_h \rangle. \end{aligned} \quad (6.60)$$

Since $\langle \lambda, u - U_h \rangle \geq 0$ and $\bar{\Lambda}_h \leq 0$, as before, we obtain that

$$-\langle \lambda - \bar{\Lambda}_h, u - U_h \rangle \leq -\langle \bar{\Lambda}_h, U_h - \chi \rangle.$$

Then applying the Young's inequality with appropriate constants for the last two terms on the right-hand side of (6.60), we get

$$\begin{aligned} & \frac{1}{2} \frac{d}{dt} \|u - U_h\|_{L^2(\Omega)}^2 + \frac{1}{4} \| \|u - \bar{U}_h\| \|^2 + \frac{1}{8} \| \|u - U_h\| \|^2 \\ & \leq 2\gamma^2 \| \| \bar{U}_h - U_h \| \|^2 - \langle \bar{\Lambda}_h, U_h - \chi \rangle + 4 \| \| \bar{\mathcal{G}}_h \| \|^2 + 4 \| \| f - \bar{F} \| \|^2. \end{aligned} \quad (6.61)$$

On the other hand, rearranging terms of (6.59) and applying the strong sector condition (2.5), we have that

$$\| \| \partial_t(u - U_h) + (\lambda - \bar{\Lambda}_h) \| \|^2 \leq 12\gamma^2 \| \|u - \bar{U}_h\| \|^2 + 3 \| \| \bar{\mathcal{G}}_h \| \|^2 + 3 \| \| f - \bar{F} \| \|^2. \quad (6.62)$$

Adding the two inequalities (6.61) and (6.62), we get the upper bound after dropping all the constants:

$$\begin{aligned} \frac{d}{dt} \|u - U_h\|_{L^2(\Omega)}^2 + \left(\|u - \bar{U}_h\|^2 + \|u - U_h\|^2 \right) + \|\partial_t(u - U_h) + (\lambda - \bar{\Lambda}_h)\|_*^2 \\ \lesssim \|\bar{U}_h - U_h\|^2 - \langle \bar{\Lambda}_h, U_h - \chi \rangle + \|\bar{\mathcal{G}}_h\|_*^2 + \|f - \bar{F}\|_*^2. \end{aligned}$$

Integrating in time, we then obtain the following upper bound of the error in $L^2(0, T; \tilde{H}^s(\Omega))$ -norm. We define the error to be

$$\begin{aligned} E^2(0, T; \Omega) := \|(u - U_h)(T)\|_{L^2(\Omega)}^2 + \int_0^T \|u - \bar{U}_h\|^2 + \|u - U_h\|^2 dt \\ + \int_0^T \|\partial_t(u - U_h) + (\lambda - \bar{\Lambda}_h)\|_*^2 dt \end{aligned} \quad (6.63)$$

Lemma 6.29 (Abstract Upper Bound: Time-dependent Problems) *Let u and $\{U_h^n\}_{n=1}^N$ are solutions of the continuous and discrete variational inequalities, (1.18) and (4.19), respectively. Then we have the following upper bound:*

$$\begin{aligned} E^2(0, T; \Omega) \lesssim \|u_0 - U_h^0\|_{L^2(\Omega)}^2 + \int_0^T \|\bar{U}_h - U_h\|^2 dt \\ + \int_0^T \|\bar{\mathcal{G}}_h\|_*^2 - \langle \bar{\Lambda}_h, U_h - \chi \rangle dt + \int_0^T \|f - \bar{F}\|_*^2 dt \end{aligned} \quad (6.64)$$

Remark 6.30 (Role of Each Term in the Upper Bound) Notice that on the right-hand side of the last inequality, the first term measures the initial error; the second term is computable and measures the error due to time discretization; and the last term gives the data consistency error due to time discretization of f . The third term corresponds to space error and has been analyzed before for stationary problems.

At each time step t_n , $\|\mathcal{G}_h^n\|_*$ can be treated exactly as in the elliptic case (see §6.4). Treating term $\langle \Lambda_h^n, U_h - \chi \rangle$ is slightly different than in Lemma 6.9 though due to the time dependence. We now estimate this term following the idea in [104, Lemma 3.2].

Lemma 6.31 (Lack of Monotonicity: Evolutionary Case) *The following inequality holds*

$$\int_{t_{n-1}}^{t_n} \langle \Lambda_h^n, U_h - \chi \rangle dt \geq - \sum_{z \in \mathcal{C}_h^n \cup \mathcal{F}_h^n} \frac{k_n}{2} \langle \Lambda_h^n, (U_h^n - U_h^{n-1}) \psi_z \rangle + \sum_{z \in \mathcal{F}_h^n} k_n s_z^n d_z^n, \quad (6.65)$$

for any $n = 1, \dots, N$, with the constants

$$d_z^n := \int_{\omega_z} (U_h^n - \chi_h^n) \psi_z = \int_{\omega_z} (U_h^n - \chi_h) \psi_z \geq 0. \quad (6.66)$$

Proof. Using definition $U_h = l(t)U_h^{n-1} + (1 - l(t))U_h^n$, with $l(t)$ given in (4.9), and integrating in time yields

$$\begin{aligned} \int_{t_{n-1}}^{t_n} \langle \Lambda_h^n, U_h - \chi_h \rangle dt &= \frac{k_n}{2} \langle \Lambda_h^n, U_h^{n-1} + U_h^n - 2\chi_h \rangle \\ &= \frac{k_n}{2} \langle \Lambda_h^n, U_h^{n-1} - U_h^n \rangle + k_n \langle \Lambda_h^n, U_h^n - \chi_h \rangle. \end{aligned}$$

We finally observe that $s_z^n = 0$ for any $z \in \mathcal{N}_h^n$ and $U_h^n = \chi_h$ in ω_z for $z \in \mathcal{C}_h^n$. Therefore $\langle \Lambda_h^n, (U_h^n - \chi_h) \psi_z \rangle = s_z^n d_z^n$ for $z \in \mathcal{F}_h^n$ and zero otherwise, whence the desired estimate (6.65) follows immediately. \blacksquare

Remark 6.32 (Further Simplification) Since we assume the obstacle does not change in time, the previous lemma can be further simplified. For any node $z \in \mathcal{C}_h^n$, we have $U_h^n = \chi$ in ω_z and $U_h^{n-1} \geq \chi$. The non-positivity of s_z^n then gives

$$\int_{t_{n-1}}^{t_n} \langle \Lambda_h^n, U_h - \chi \rangle dt \geq - \sum_{z \in \mathcal{F}_h^n} \frac{k_n}{2} \langle s_z^n, (U_h^n - U_h^{n-1}) \psi_z \rangle + \sum_{z \in \mathcal{F}_h^n} k_n s_z^n d_z^n.$$

Remark 6.33 (Lower Bounds) Similar abstract lower bound in terms of $\|\mathcal{G}_h^n\|_*$ can be obtained as in §6.4.2; a lower bound in terms of the time error estimator is trivial due to the triangle inequality:

$$\int_{t_{n-1}}^{t_n} \|\overline{U}_h - U_h\|^2 dt \leq 2 \int_{t_{n-1}}^{t_n} \|\overline{U}_h - u\|^2 + \|U_h - u\|^2 dt.$$

6.5.3 Localized Error Estimators

Finally, we summarize this section by giving a computable residual-type local error estimate. Let $R(U_h^n)$ and $J(U_h^n)$ be interior and jump residual at time t_n ,

respectively, i.e.

$$\begin{aligned} R(U_h^n) &:= F^n - \delta U_h^n - \mathcal{A}_I U_h^n - c_1 \cdot \nabla U_h^n - c_0 U_h^n \\ J(U_h^n) &:= -c_2(\nabla U_h^n|_{\tau_1} \cdot \nu_1 + \nabla U_h^n|_{\tau_2} \cdot \nu_2) \end{aligned}$$

We shall use residual-type space error estimator as an example here for time-dependent problems. Other types of error estimators can also be derived without much difficulty. We define the following jump and interior indicators as in §6.4:

$$(\eta_z^n)^2 := h_z \|J(U_h^n)\|_{L^2(\gamma_z)}^2 \quad \text{and} \quad (\xi_z^n)^2 := h_z^{2s+d-\frac{2d}{p}} \|(R(U_h^n) - R_z^n) \psi_z\|_{L^p(\omega_z)}^2,$$

where $R_z^n := \langle R^n, \psi_z \rangle / \langle 1, \psi_z \rangle$ is the weighted average. Define the error estimator

$$\mathcal{E} := (\mathcal{E}_0^2 + \mathcal{E}_k^2 + \mathcal{E}_h^2 + \mathcal{E}_{kh}^2 + \mathcal{E}_D^2)^{\frac{1}{2}} \quad (6.67)$$

with

$$\begin{aligned} \mathcal{E}_0^2 &:= \|u_0 - U_h^0\|_{L^2(\Omega)}^2 && \text{initial error} \\ \mathcal{E}_k^2 &:= \sum_{n=1}^N \frac{k_n}{3} \|\|U_h^n - U_h^{n-1}\|\|^2 dt && \text{time error} \\ \mathcal{E}_h^2 &:= \sum_{n=1}^N k_n \left\{ \sum_{z \in \mathcal{P}_h \setminus \mathcal{C}_h^n} [(\eta_z^n)^2 + (\xi_z^n)^2] - \sum_{z \in \mathcal{F}_h^n} s_z^n d_z^n \right\} && \text{space error} \\ \mathcal{E}_{kh}^2 &:= \sum_{n=1}^N k_n \left\{ \sum_{z \in \mathcal{F}_h^n} |\langle s_z^n, (U_h^n - U_h^{n-1}) \psi_z \rangle| \right\} && \text{mixed error} \\ \mathcal{E}_D^2 &:= \int_0^T \|\|f - \bar{F}\|\|_*^2 dt && \text{data consistency} \end{aligned}$$

Applying Lemma 6.7 and 6.8 on $\|\|\mathcal{G}_h^n\|\|_*^2$, Lemma 6.31 and Remark 6.32 on $\int_0^T \langle \Lambda_h^n, U_h - \chi \rangle dt$, we then have the following computable and localized upper bound from the abstract upper bound (6.64):

Theorem 6.34 (Upper Bound: Evolutionary Case) *Let $f \in L^1(0, T; L^p(\Omega))$ and $p \geq 1$ satisfies*

$$Y - 1 < \frac{1}{p} < \frac{\rho}{2d} + \frac{1}{2}.$$

Then we have the following upper bound for the error

$$E^2(0, T; \Omega) \lesssim \mathcal{E}^2.$$

Remark 6.35 (Inactive Constraint) For the noncontact nodes \mathcal{N}_h^n , the variational inequality becomes an equality. This is reflected on the vanishing of all terms that account for the unilateral constraint. The resulting estimator reduces to an energy-type estimator for a linear diffusion equation. This result, however, is different from earlier versions [119, 136, 16] in that

- our new error indicators are star-based instead of element-based;
- the interior residual estimator is of higher order than the jump estimator for differential operators;
- the linear sectorial integro-differential operator \mathcal{A} is much more general than the Laplace operator.

6.6 General Obstacle

In previous sections, we derived an a posteriori error estimator for variational inequalities with the conformity assumption, i.e. $\mathbb{K} \subset \mathcal{K}$. In practice, we could have problems with an obstacle which cannot be approximated exactly by piecewise linear functions. For example, in American put option pricing problem (see §3.2), obstacles usually take a form like $\chi(x) = \max(K - e^x, 0)$ where K is a constant. We shall now consider the general case: Problem 4.7 with general obstacle χ which might depend on time also.

6.6.1 A Magic Bullet?

Since χ is known, one can make a transformation $w = u - \chi$ and rewrite the original VI as a new problem for w with a zero obstacle. It seems that difficulties associated problems with a general obstacle could be dealt with exactly as before. But actually this may not be a good idea since, as in §6.4.1, it is appropriate to look at the difference $u - \chi$ only in the contact region but not in the noncontact region.

This can be explained by a simple example. In Figure 6.5, the solution u is smooth outside of the contact region. The oscillatory obstacle χ should not affect

the mesh grading. But after transformation, $w = u - \chi$, we introduce artificial singularities because w is not smooth and local refinement is needed outside of the contact set. A related issue we want to point out here is that in the contact set, there is a kink at $x = 0$ which makes the solution u not smooth, but it is not necessary to refine more around $x = 0$ provided $x = 0$ is a mesh point. Inside of the contact region, the only thing that matters is the obstacle resolution.

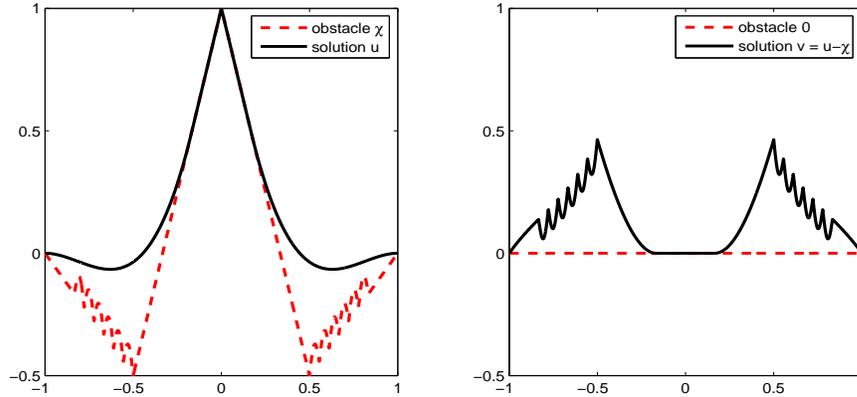


Figure 6.5: Localization Effect. Left: The obstacle χ is oscillatory outside of the contact region where the solution u is smooth. Right: After transformation $w = u - \chi$, the solution w is not smooth outside of the contact region and very small meshsize is needed there.

Based on this observation, we consider the case of general obstacles $\chi \in H^1(\mathcal{Q})$ directly instead of relying on the “magic” transformation. This generalization will not affect the estimation of $\|\mathcal{G}_h\|_*$ which is built solely upon the approximate obstacle χ_h but not related to the exact obstacle χ . We only need to revisit the estimation of

$$\int_0^T \langle \lambda - \bar{\Lambda}_h, u - U_h \rangle dt.$$

6.6.2 Obstacle Consistency Error

Therefore, in what follows, we derive a lower bound for $\int_{t_{n-1}}^{t_n} \langle \lambda - \Lambda_h^n, u - U_h \rangle dt$. To this end, we further define $\chi_h = l(t)\chi_h^{n-1} + (1 - l(t))\chi_h^n \in C([0, T]; \mathbb{V}(\Omega))$ to

be a space-time piecewise linear approximation of χ . Notice that, for numerical approximation, we only need $\{\chi_h^n\}_{n=1}^N$; the piecewise linear function χ_h is used solely for theoretical purposes.

We observe that in general $\chi_h(t) \not\equiv \chi(t)$ for $0 \leq t \leq T$. To handle this lack of consistency, we follow Veerer [134] and introduce the auxiliary function $U_h^* := \max(U_h, \chi) \in \mathcal{K}$. Since $\langle \lambda, u - U_h^* \rangle \geq 0$, we have that

$$\langle \lambda - \Lambda_h^n, u - U_h \rangle \geq \langle \Lambda_h^n, U_h^* - u \rangle + \langle \lambda - \Lambda_h^n, U_h^* - U_h \rangle. \quad (6.68)$$

We next consider each term on the right-hand side of (6.68) separately.

First Part

For the first term on the right-hand side of (6.68), we invoke

$$\Lambda_h^n \leq 0 \quad \text{and} \quad \langle \Lambda_h^n, \chi - u \rangle \geq 0$$

to obtain

$$\begin{aligned} \langle \Lambda_h^n, U_h^* - u \rangle &\geq \langle \Lambda_h^n, U_h^* - \chi \rangle \\ &= \langle \Lambda_h^n, U_h - \chi_h \rangle + \langle \Lambda_h^n, U_h^* - U_h \rangle + \langle \Lambda_h^n, \chi_h - \chi \rangle. \end{aligned} \quad (6.69)$$

Arguing as in the proof of Lemma 6.31 with the first term on the right-hand side, we deduce

$$\begin{aligned} &\int_{t_{n-1}}^{t_n} \langle \Lambda_h^n, U_h - \chi_h \rangle dt \\ &= - \sum_{z \in \mathcal{P}_h \setminus \mathcal{N}_h^n} \frac{k_n}{2} \langle s_z^n, ((U_h^n - U_h^{n-1}) - (\chi_h^n - \chi_h^{n-1})) \psi_z \rangle + \sum_{z \in \mathcal{F}_h^n} k_n s_z^n d_z^n. \end{aligned}$$

The first term on the right-hand side is the most general form of the *mixed error* in Theorem 6.34.

However, we now have two additional terms in (6.69) that account for the obstacle inconsistent approximation, as illustrated in Figure 6.6. To bound them we utilize the definition of U_h^* , which results in $U_h^* - U_h = (\chi - U_h)^+$, as well as $\Lambda_h^n \leq 0$,

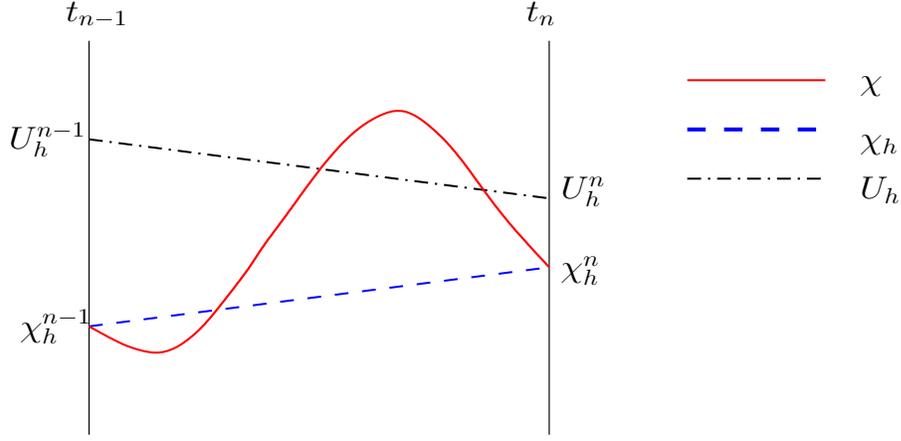


Figure 6.6: *Obstacle Consistency*: If the obstacle χ and its space-time piecewise linear approximation χ_h do not coincide in $\omega_z \times (t_{n-1}, t_n)$ for nodes $z \in \mathcal{P}_h \setminus \mathcal{N}_h^n$, then the quantities $\langle \Lambda_h^n, (\chi - U_h)^+ \psi_z \rangle$ and $\langle \Lambda_h^n, (\chi_h - \chi)^+ \psi_z \rangle$ measure the local lack of conformity. Note that these quantities vanish for $z \in \mathcal{N}_h^n$, that is for the noncontact nodes.

and end up with

$$\begin{aligned} \langle \Lambda_h^n, (U_h^* - U_h) \psi_z \rangle &\geq \langle \Lambda_h^n, (\chi - U_h)^+ \psi_z \rangle, \\ \langle \Lambda_h^n, (\chi_h - \chi) \psi_z \rangle &\geq \langle \Lambda_h^n, (\chi_h - \chi)^+ \psi_z \rangle. \end{aligned}$$

Second Part

We can also rewrite the second term on the right-hand side of (6.68) as follows:

$$\langle \lambda - \Lambda_h^n, U_h^* - U_h \rangle = \langle (\partial_t u - \delta_t U_h) + (\lambda - \Lambda_h^n), (\chi - U_h)^+ \rangle - \langle (\partial_t u - \delta_t U_h), (\chi - U_h)^+ \rangle.$$

The second term on the right-hand side is most problematic. We handle it via integration by parts in time:

$$- \int_0^T \langle \partial_t (u - U_h), (\chi - U_h)^+ \rangle = - \langle u - U_h, (\chi - U_h)^+ \rangle \Big|_0^T + \int_0^T \langle u - U_h, \partial_t (\chi - U_h)^+ \rangle dt.$$

Note that we can eliminate the first term on the right-hand side at $t = 0$ because if $\chi_0(x) > U_h^0(x)$ then $u_0(x) \geq \chi_0(x) > U_h^0(x)$ whence $\langle u_0 - U_h^0, (\chi_0 - U_h^0)^+ \rangle \geq 0$.

Upper Bound of Obstacle Consistency Error

With the estimates of (6.68) given above, we now derive an upper bound of the obstacle consistency error. After applying the Cauchy-Schwarz inequality three times, we arrive at

$$\begin{aligned}
& \int_0^T \langle \lambda - \bar{\Lambda}_h, u - U_h \rangle dt \\
& \geq - \sum_{n=1}^N \left(\sum_{z \in \mathcal{P}_h \setminus \mathcal{N}_h^n} \frac{k_n}{2} \langle \Lambda_h^n, ((U_h^n - U_h^{n-1}) - (\chi_h^n - \chi_h^{n-1})) \psi_z \rangle - \sum_{z \in \mathcal{F}_h^n} k_n s_z^n d_z^n \right) \\
& \quad + \sum_{z \in \mathcal{P}_h} \int_0^T \langle \bar{\Lambda}_h, ((\chi - U_h)^+ + (\chi_h - \chi)^+) \psi_z \rangle dt \\
& \quad - \frac{\varepsilon_1}{2} \int_0^T \left\| \partial_t(u - U_h) + (\lambda - \bar{\Lambda}_h) \right\|_*^2 dt - \frac{1}{2\varepsilon_1} \int_0^T \left\| (\chi - U_h)^+ \right\|^2 dt \\
& \quad - \frac{\varepsilon_2}{2} \left\| (u - U_h)(T) \right\|^2 - \frac{1}{2\varepsilon_2} \left\| (\chi - U_h)^+(T) \right\|^2 \\
& \quad - \int_0^T \frac{\varepsilon_3}{2} \left\| u - U_h \right\|^2 + \frac{1}{2\varepsilon_3} \left\| \partial_t(\chi - U_h)^+ \right\|_*^2 dt,
\end{aligned}$$

with $\varepsilon_1, \varepsilon_2, \varepsilon_3 > 0$ arbitrary. We finally choose appropriate $\varepsilon_1, \varepsilon_2$ and ε_3 , and insert the above estimate into (6.75) to obtain an upper bound.

We first define the error estimator which has one more term compared with Theorem 6.34 to account for the obstacle consistency error

$$\mathcal{E} := (\mathcal{E}_0^2 + \mathcal{E}_k^2 + \mathcal{E}_h^2 + \mathcal{E}_{kh}^2 + \mathcal{E}_\chi^2 + \mathcal{E}_D^2)^{\frac{1}{2}} \tag{6.70}$$

with

$$\begin{aligned}
\mathcal{E}_0^2 &:= \|u_0 - U_h^0\|_{L^2(\Omega)}^2 && \text{initial error} \\
\mathcal{E}_k^2 &:= \sum_{n=1}^N \frac{k_n}{3} \|U_h^n - U_h^{n-1}\|^2 dt && \text{time error} \\
\mathcal{E}_h^2 &:= \sum_{n=1}^N k_n \left\{ \sum_{z \in \mathcal{P}_h \setminus \mathcal{C}_h^n} [(\eta_z^n)^2 + (\xi_z^n)^2] - \sum_{z \in \mathcal{F}_h^n} s_z^n d_z^n \right\} && \text{space error} \\
\mathcal{E}_{kh}^2 &:= \sum_{n=1}^N k_n \left\{ \sum_{z \in \mathcal{C}_h^n \cup \mathcal{F}_h^n} |\langle \Lambda_h^n, ((U_h^n - U_h^{n-1}) - (\chi_h^n - \chi_h^{n-1})) \psi_z \rangle| \right\} && \text{mixed error} \\
\mathcal{E}_\chi^2 &:= \|(\chi - U_h)^+(T)\|^2 + \int_0^T \left(\|(\chi - U_h)^+\|^2 + \|\partial_t(\chi - U_h)^+\|_*^2 \right) dt && \text{obstacle consistency} \\
&\quad - \sum_{n=1}^N \sum_{z \in \mathcal{C}_h^n \cup \mathcal{F}_h^n} \int_{t_{n-1}}^{t_n} \langle \Lambda_h^n, \{(\chi - U_h)^+ + (\chi_h - \chi)^+\} \psi_z \rangle dt \\
\mathcal{E}_D^2 &:= \int_0^T \|f - \bar{F}\|_*^2 dt && \text{data consistency}
\end{aligned}$$

Theorem 6.36 (Upper Bound: General Obstacles) *For Problem 4.7 with a general obstacle $\chi \in H^1(\mathcal{Q})$, we have the following upper a posteriori bound*

$$E^2(0, T; \Omega) \lesssim \mathcal{E}^2.$$

Remark 6.37 (Obstacle Consistency) Terms involving $(\chi - U_h)^+$ are only active away from the noncontact set, a crucial localization property, and accounts for the lack of constraint consistency $U_h < \chi$ in both space and time; see Figure 6.6. The space-time situation $\chi_h > \chi$, depicted in Figure 6.6, is only detected by the term $\langle \Lambda_h^n, (\chi_h - \chi)^+ \psi_z \rangle$. In particular, if $z \in \mathcal{C}_h^n$ is a full-contact node, then this is the only nonzero local indicator. Besides justifying its presence, this argument shows that such a term can be regarded as a complement to the notion of full contact nodes which hinges on the condition $\chi_h^n = \chi(t_n)$ in ω_z ; see §6.5.1. For a kink or cusp pointing downwards the relation $\chi_h > \chi$ is not only to be expected but it might suggest that one needs strong local refinement. This is not true because asymptotically the discrete solution detaches from the obstacle and so $\langle \Lambda_h^n, (\chi_h - \chi)^+ \psi_z \rangle = 0$; see [113] for a full discussion.

6.7 Mesh Changes and Coarsening Error

Till this point, we assumed the spatial test function space \mathbb{V} does not change in time. To derive a practical adaptive algorithm for evolution problems, we need to allow mesh to change in time to give optimal approximation at each time step. This is because singularities of solutions of time-dependent problems could change their location or strength.

Mesh change is a delicate issue for evolution problems. An example has been constructed by Dupont [56] who showed changing the mesh in an uncontrolled way could lead to convergence to wrong solutions. For linear parabolic equations, coarsening error is examined by Chen and Feng [48], and Lakkis and Makridakis [90], and earlier by Nochetto et al [112] for degenerate parabolic problems. In this section, we shall consider mesh changing and coarsening error estimates.

6.7.1 Transfer Operator

Let $\Omega \subset \mathbb{R}^d$ be an open and bounded polygonal domain. We now introduce spatial quantities for $1 \leq n \leq N$ fixed. Let \mathcal{T}^n be the mesh at time t_n and \mathcal{P}_h^n be the set of all nodes of \mathcal{T}^n , including the boundary nodes. Let \mathbb{V}^n be the space of continuous piecewise linear finite element functions on \mathcal{T}^n .

For problems with general obstacles, it is not obvious how to define the transfer operator from one time step to the other because the usual linear interpolation operator or L^2 -projection operator does not always work in practice. As an example, we consider linear interpolation operator $I_{n-1}^n : \mathbb{V}^{n-1} \rightarrow \mathbb{V}^n$ as the transfer operator and show why it fails in a thought experiment in Figure 6.7.

In Figure 6.7, we suppose the exact solution u does not change in time. At time step n , the adaptive algorithm detects that the time error is quite big because of the sudden change of numerical solution U_h in the contact region and decides to reduce the time step-size to make the time error smaller. Since this effect is actually due to the resolution of the obstacle instead of the time step, reducing the time step-size does not help. Hence the adaptive algorithm would either get stuck here

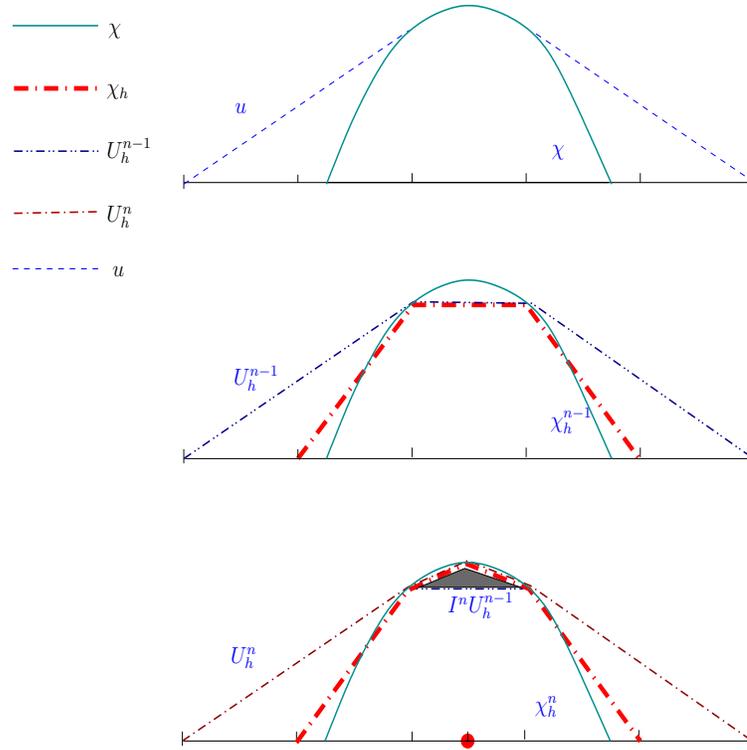


Figure 6.7: Top: exact solution at time t_n . Middle: numerical solution U_h^{n-1} for uniform mesh. Bottom: numerical solution U_h^n . Since in the contact region the numerical solution U_h^{n-1} is below χ , the adaptive algorithm detects this and refines accordingly. However $I_{n-1}^n U_h^{n-1} = U_h^{n-1}$, whence the time error (difference between U_h^n and $I_{n-1}^n U_h^{n-1}$, which is related to the gray area) does not decrease as the time step-size decreases.

if there is no control on the maximum number of iterations for time adaptation, or end up with unnecessary refinement of time step-size.

Inspired by this example, we now introduce a new transfer operator $\mathcal{I}_{n-1}^n : \mathbb{V}^{n-1} \rightarrow \mathbb{V}^n$ which circumvent this difficulty:

$$\mathcal{I}_{n-1}^n v := \sum_{z \in \mathcal{P}_h^n} \max \{ I_{n-1}^n v(z), \chi_h^n(z) \} \psi_z, \quad (6.71)$$

where $I_{n-1}^n : \mathbb{V}^{n-1} \rightarrow \mathbb{V}^n$ is the linear interpolation operator. If the obstacle does not change in time, i.e. $\chi_h^n = \chi_h^{n-1}$, then $U_h^{n-1} \geq \chi_h^n$ and \mathcal{I}_{n-1}^n reduces to the previous transfer operator I_{n-1}^n . Numerical experiments in Chapter 8 demonstrate \mathcal{I}_{n-1}^n works well in practice.

6.7.2 Residual and Galerkin Functional for Mesh Changes

We now need to introduce and modify notation to deal with mesh changes. For any sequence $\{W^n\}_{n=1}^N$, we still denote the piecewise constant interpolant \overline{W} and the piecewise linear interpolant W ; see (6.72). Furthermore we define the new piecewise linear function \widetilde{W} to be

$$\widetilde{W}(t) := l(t)\mathcal{I}_{n-1}^n W^{n-1} + (1 - l(t))W^n, \quad (6.72)$$

for any $t \in (t_{n-1}, t_n]$, $1 \leq n \leq N$, where the linear function $l(t)$ is defined in (4.9). We also denote by

$$\delta W^n := \frac{W^n - W^{n-1}}{k_n}, \quad \tilde{\delta} W^n := \frac{W^n - \mathcal{I}_{n-1}^n W^{n-1}}{k_n} \quad \forall 1 \leq n \leq N. \quad (6.73)$$

After comparing these new notation with our notation introduced in Chapter 4, we can easily find that

$$\delta_t \widetilde{W}(t) = \tilde{\delta} W^n \quad \forall t \in (t_{n-1}, t_n].$$

The definition of residual is also modified due to mesh changes

$$r_h^n := F^n - \tilde{\delta} U_h^n - \mathcal{A} U_h^n,$$

as in the definition of nonlinear defect measure \mathcal{G}_h^n , the Galerkin functional

$$\mathcal{G}_h^n := r_h^n - \Lambda_h^n = \sum_{z \in \mathcal{P}_h \setminus \mathcal{C}_h} (r_h^n - s_z^n) \psi_z$$

that now incorporates the new definition of r_h^n .

We split the set \mathcal{P}_h^n into three disjoint sets as before (but with the new definition of r^n):

$$\mathcal{P}_h^n = \mathcal{N}_h^n \cup \mathcal{C}_h^n \cup \mathcal{F}_h^n$$

with the noncontact nodes \mathcal{N}_h^n , full-contact nodes \mathcal{C}_h^n , and free boundary nodes \mathcal{F}_h^n defined as in (6.55).

6.7.3 Coarsening Error Estimate

We apply the energy method used in §6.5; see (6.59). From the definition (6.54) of the Lagrange multiplier λ , it follows that for any $\varphi \in \tilde{H}^s(\Omega)$

$$\langle \bar{\mathcal{G}}_h, \varphi \rangle = \langle (\mathcal{A}(u - \bar{U}_h) + (\partial_t u - \delta_t \tilde{U}_h) + (\lambda - \bar{\Lambda}_h), \varphi) - \langle f - \bar{F}, \varphi \rangle.$$

By taking $\varphi = u(t) - U_h(t)$ in the last equation, we obtain

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|u - U_h\|^2 + \langle \mathcal{A}(u - \bar{U}_h), u - U_h \rangle \\ = \langle \bar{\mathcal{G}}_h, u - U_h \rangle + \langle \delta_t U_h - \delta_t \tilde{U}_h, u - U_h \rangle \\ - \langle \lambda - \bar{\Lambda}_h, u - U_h \rangle + \langle f - \bar{F}, u - U_h \rangle. \end{aligned} \quad (6.74)$$

We now proceed as in Lemma 6.29, namely we Integrate both sides of the equality on $[0, T]$ and use the coercivity inequality (2.9) and the Cauchy-Schwarz inequality to get the following inequality:

$$\begin{aligned} E^2(0, T; \Omega) \\ \lesssim \|u_0 - U_h^0\|^2 & \text{initial error} \\ + \sum_{n=1}^N \int_{t_{n-1}}^{t_n} \|\bar{U}_h - U_h\|^2 + \langle \delta_t U_h - \delta_t \tilde{U}_h, u - U_h \rangle dt & \text{evolution error} \\ + \sum_{n=1}^N \int_{t_{n-1}}^{t_n} \|\bar{\mathcal{G}}_h\|_*^2 dt & \text{spatial error} \\ - \int_0^T \langle \lambda - \bar{\Lambda}_h, u - U_h \rangle dt & \text{mixed error} \\ + \int_0^T \|f - \bar{F}\|_*^2 dt. & \text{data oscillation} \end{aligned} \quad (6.75)$$

Remark 6.38 (Coarsening Error) Note that, comparing with Lemma 6.29, we now have the new term $\int_0^T \langle \delta_t U_h - \delta_t \tilde{U}_h, u - U_h \rangle dt$ on the right-hand side that accounts for mesh evolution. The remaining terms can be handled as in previous sections.

We now discuss the difference between the case when mesh changes and the fixed mesh case, especially the new term. It is easy to show, by triangular inequality, that

$$\begin{aligned} \int_{t_{n-1}}^{t_n} \|\bar{U}_h - U_h\|^2 dt &= \frac{1}{3} k_n \|U_h^n - U_h^{n-1}\|^2 \\ &\leq \frac{2}{3} k_n \left(\|U_h^n - \mathcal{I}_{n-1}^n U_h^{n-1}\|^2 + \|\mathcal{I}_{n-1}^n U_h^{n-1} - U_h^{n-1}\|^2 \right). \end{aligned}$$

Furthermore, we have

$$\|\mathcal{I}_{n-1}^n U_h^{n-1} - U_h^{n-1}\| \leq \|\mathcal{I}_{n-1}^n U_h^{n-1} - I_{n-1}^n U_h^{n-1}\| + \|I_{n-1}^n U_h^{n-1} - U_h^{n-1}\|.$$

Hence it follows that

$$\begin{aligned} \int_{t_{n-1}}^{t_n} \|\bar{U}_h - U_h\|^2 dt &\lesssim k_n \left(\|U_h^n - \mathcal{I}_{n-1}^n U_h^{n-1}\|^2 \right. \\ &\quad \left. + \|I_{n-1}^n U_h^{n-1} - U_h^{n-1}\|^2 \right. \\ &\quad \left. + \|\mathcal{I}_{n-1}^n U_h^{n-1} - I_{n-1}^n U_h^{n-1}\|^2 \right). \end{aligned} \tag{6.76}$$

Notice that the three terms in (6.76) represent three different parts of the error: the first term is the time error; the second term is the coarsening error; and the last term contributes to the obstacle consistency error. These three terms will contribute in \mathcal{E}_k , \mathcal{E}_c , and \mathcal{E}_χ , respectively.

To handle the new term, we recall that

$$\delta_t U_h - \delta_t \tilde{U}_h = \frac{\mathcal{I}_{n-1}^n U_h^{n-1} - U_h^{n-1}}{k_n}$$

and use the Cauchy-Schwarz inequality to get

$$\begin{aligned}
& \int_{t_{n-1}}^{t_n} \langle \delta_t U_h - \tilde{\delta}_t U_h, U_h - u \rangle dt \\
& \leq \int_{t_{n-1}}^{t_n} \frac{1}{k_n} \|\mathcal{I}_{n-1}^n U_h^{n-1} - U_h^{n-1}\|_* \|U_h - u\| dt \\
& \leq \int_{t_{n-1}}^{t_n} \frac{1}{2\varepsilon k_n^2} \|\mathcal{I}_{n-1}^n U_h^{n-1} - U_h^{n-1}\|_*^2 dt + \int_{t_{n-1}}^{t_n} \frac{\varepsilon}{2} \|U_h - u\|^2 dt \\
& \leq \int_{t_{n-1}}^{t_n} \frac{1}{\varepsilon k_n^2} \|\mathcal{I}_{n-1}^n U_h^{n-1} - I_{n-1}^n U_h^{n-1}\|_*^2 + \frac{1}{\varepsilon k_n^2} \|I_{n-1}^n U_h^{n-1} - U_h^{n-1}\|_*^2 dt \\
& \quad + \int_{t_{n-1}}^{t_n} \frac{\varepsilon}{2} \|U_h - u\|^2 dt, \tag{6.77}
\end{aligned}$$

for any positive constant ε . We can choose appropriate ε to absorb the last term on the right-hand side of (6.7.4). Then we are left with two new terms, namely

$$\frac{1}{k_n} \|\mathcal{I}_{n-1}^n U_h^{n-1} - I_{n-1}^n U_h^{n-1}\|_*^2 \quad \text{and} \quad \frac{1}{k_n} \|I_{n-1}^n U_h^{n-1} - U_h^{n-1}\|_*^2.$$

These terms accounts for the obstacle consistency error and coarsening error, respectively. So we add these two terms to \mathcal{E}_χ and \mathcal{E}_c , respectively.

6.7.4 Final A Posteriori Upper Bound

We combine the inequalities (6.76) and with the estimate in the previous section and choose appropriate constant ε to arrive at the following upper bound of the error $E(0, T; \Omega)$.

Theorem 6.39 (Final Upper Bound) *For Problem 4.7 with a general obstacle $\chi \in H^1(\mathcal{Q})$, we have the following upper a posteriori bound for adaptive mesh*

$$E^2(0, T; \Omega) \lesssim \mathcal{E}^2,$$

where the error estimator is given by

$$\mathcal{E} := (\mathcal{E}_0^2 + \mathcal{E}_k^2 + \mathcal{E}_h^2 + \mathcal{E}_{kh}^2 + \mathcal{E}_c^2 + \mathcal{E}_\chi^2 + \mathcal{E}_D^2)^{\frac{1}{2}}.$$

The various estimators account for different discretization effects and are listed and described below:

Initial Error Estimate

$$\mathcal{E}_0^2 := \|u_0 - U_h^0\|_{L^2(\Omega)}^2$$

This part of the error estimator is due to the initial mesh and approximation of the initial condition u_0 . It can never be reduced once the initial mesh has been fixed.

Time Error Estimate

$$\mathcal{E}_k^2 := \sum_{n=1}^N k_n \left\| U_h^n - \mathcal{I}_{n-1}^n U_h^{n-1} \right\|^2$$

This part measures the error because of the evolution of the solution u . Philosophically it is only a good approximation of the evolution error when the space error is small, i.e. U_h^n is close enough to the real solution $u(t_n)$.

Space Error Estimate

$$\mathcal{E}_h^2 := \sum_{n=1}^N k_n \left\{ \sum_{z \in \mathcal{P}_h^n \setminus \mathcal{C}_h^n} [(\eta_z^n)^2 + (\xi_z^n)^2] - \sum_{z \in \mathcal{F}_h^n} s_z^n d_z^n \right\}$$

where we modify the residual-type error estimators in (6.70) as follows

$$\eta_z^n := \left\| h^{\frac{1}{2}} J(U_h^n) \right\|_{L^2(\gamma_z)} \quad \text{and} \quad \xi_z^n := \left\| h^{s+\frac{d}{2}-\frac{d}{p}} (R(U_h^n) - R_z^n) \psi_z \right\|_{L^p(\omega_z)}.$$

This is because we may have different local meshsize in different stage of evolution. The constants s_z^n and d_z^n are defined in (6.57) and (6.66), respectively. We can separate the contribution into three parts $\mathcal{E}_h^2 = \mathcal{E}_{h,1}^2 + \mathcal{E}_{h,2}^2 + \mathcal{E}_{h,3}^2$ where

$$\begin{aligned} \mathcal{E}_{h,1}^2 &:= \sum_{n=1}^N \sum_{z \in \mathcal{P}_h^n \setminus \mathcal{C}_h^n} k_n (\eta_z^n)^2 \\ \mathcal{E}_{h,2}^2 &:= \sum_{n=1}^N \sum_{z \in \mathcal{P}_h^n \setminus \mathcal{C}_h^n} k_n (\xi_z^n)^2 \\ \mathcal{E}_{h,3}^2 &:= - \sum_{n=1}^N \sum_{z \in \mathcal{P}_h^n \setminus \mathcal{F}_h^n} k_n s_z^n d_z^n. \end{aligned}$$

Mixed Error Estimate

$$\mathcal{E}_{kh}^2 := \sum_{n=1}^N k_n \left\{ \sum_{z \in \mathcal{C}_h^n \cup \mathcal{F}_h^n} |\langle \Lambda_h^n, ((U_h^n - \mathcal{I}_{n-1}^n U_h^{n-1}) - (\chi_h^n - \mathcal{I}_{n-1}^n \chi_h^{n-1})) \psi_z \rangle| \right\}$$

This part contributes not only to error due to the space discretization but also to evolutionary error.

Coarsening Error Estimate

$$\begin{aligned} \mathcal{E}_c^2 := & \sum_{n=1}^N k_n \left\| U_h^{n-1} - I_{n-1}^n U_h^{n-1} \right\|^2 + \sum_{n=1}^N \frac{1}{k_n} \left\| U_h^{n-1} - I_{n-1}^n U_h^{n-1} \right\|_*^2 \\ & + \sum_{n=1}^N k_n \sum_{z \in \mathcal{C}_h^n \cup \mathcal{F}_h^n} \langle \Lambda_h^n, \{ (I_{n-1}^n U_h^{n-1} - U_h^{n-1}) - (I_{n-1}^n \chi_h^{n-1} - \chi_h^{n-1}) \} \psi_z \rangle \end{aligned}$$

This quantifies the coarsening error. Mesh coarsening leads to information loss and thus the need to control it not to spoil the overall approximation.

Obstacle Consistency Error Estimate

$$\begin{aligned} \mathcal{E}_\chi^2 := & \|(\chi - U_h)^+(T)\|^2 + \int_0^T \left(\left\| (\chi - U_h)^+ \right\|^2 + \left\| \partial_t (\chi - U_h)^+ \right\|_*^2 \right) dt \\ & + \sum_{n=1}^N k_n \left\| \mathcal{I}_{n-1}^n U_h^{n-1} - I_{n-1}^n U_h^{n-1} \right\|^2 + \sum_{n=1}^N \frac{1}{k_n} \left\| \mathcal{I}_{n-1}^n U_h^{n-1} - I_{n-1}^n U_h^{n-1} \right\|_*^2 \\ & - \sum_{n=1}^N \sum_{z \in \mathcal{C}_h^n \cup \mathcal{F}_h^n} \int_{t_{n-1}}^{t_n} \langle \Lambda_h^n, \{ (\chi - U_h)^+ + (\chi_h - \chi)^+ \} \psi_z \rangle dt \end{aligned}$$

This part measures the discrepancy between the numerical obstacle χ_h and the real obstacle χ .

Data Oscillation Estimate

$$\mathcal{E}_D^2 := \int_0^T \left\| f - \overline{F} \right\|_*^2 dt$$

This part of the estimator gives information of the approximation of the data f .

Chapter 7

Adaptive and Multilevel Algorithms

It is well known that the standard finite element approximation on a quasi-uniform grid converges optimally with respect to the number of degrees of freedom provided the solution is sufficiently smooth. However, sometimes solutions might not be smooth enough for the standard finite element method to achieve optimal convergence rate. Furthermore, the strength and locations of singularities are sometimes not known a priori. This rules out the possibility to design a priori optimal meshes. In particular, for American option pricing problems, the solution is singular close to the maturity in time and the strike price in space; in some cases, the space derivative of the log-price has jumps across the free boundary (whose location is unknown). With this motivation in mind, in this chapter we design a practical adaptive time-space mesh refinement strategy based on the a posteriori error estimators proposed in Chapter 6. The rest of the chapter is organized as follows. We first give a brief introduction to adaptive finite element methods for stationary as well as evolutionary variational inequalities in Section 7.1. We then discuss major steps of the adaptive algorithm in §7.2, 7.3, 7.4, and 7.5.

7.1 Introduction

After more than thirty years of extensive development, adaptive methods are now standard tools in science and engineering. Adaptive mesh refinement is important to deal with multiscale phenomena and to reduce the size of linear systems that

arise from finite element discretizations. In many practical applications, solutions of PDEs are singular. Furthermore, location and strength of singularities are not known in general. The goal of adaptive methods is to generate graded meshes in space and time that automatically adapt to the problem at hand such that certain error is smaller than a tolerance with minimal computational cost.

7.1.1 Adaptive Algorithm for Static Problem

Generally, the adaptive FEM for static problems generates graded meshes and iterations in the form

$$\boxed{\text{SOLVE} \rightarrow \text{ESTIMATE} \rightarrow \text{MARK} \rightarrow \text{REFINE/COARSEN.}} \quad (7.1)$$

In finite element methods, a finite dimensional test function space is associated with a given mesh. The **SOLVE** step finds the discrete solution of the finite dimensional approximate problem. Usually this finite dimensional problem is solved by some iterative method. The **ESTIMATE** procedure quantifies the error size. Since we cannot compute the exact error of the solution, we need to find computable local error indicators to estimate the local error of the discrete solution. As soon as the local error indicator has been computed by **ESTIMATE**, the procedure **MARK** uses their magnitude to determine regions of the domain that may undergo mesh refinement or coarsening. A simple flowchart is given in Figure 7.1. To design a good adaptive finite element method, reliable and efficient *a posteriori* error estimation is essential. To learn more about adaptive algorithm design as well as implementation issues, we refer to the book by Schmidt and Siebert [123].

7.1.2 Adaptive Algorithm for Evolution Problems

For time-dependent problems, we need to add an outer loop to the procedure above to take care of the time variable and its adaptive control of step-size. In ALBERTA [123], for general time-dependent problems, the following algorithm is used:

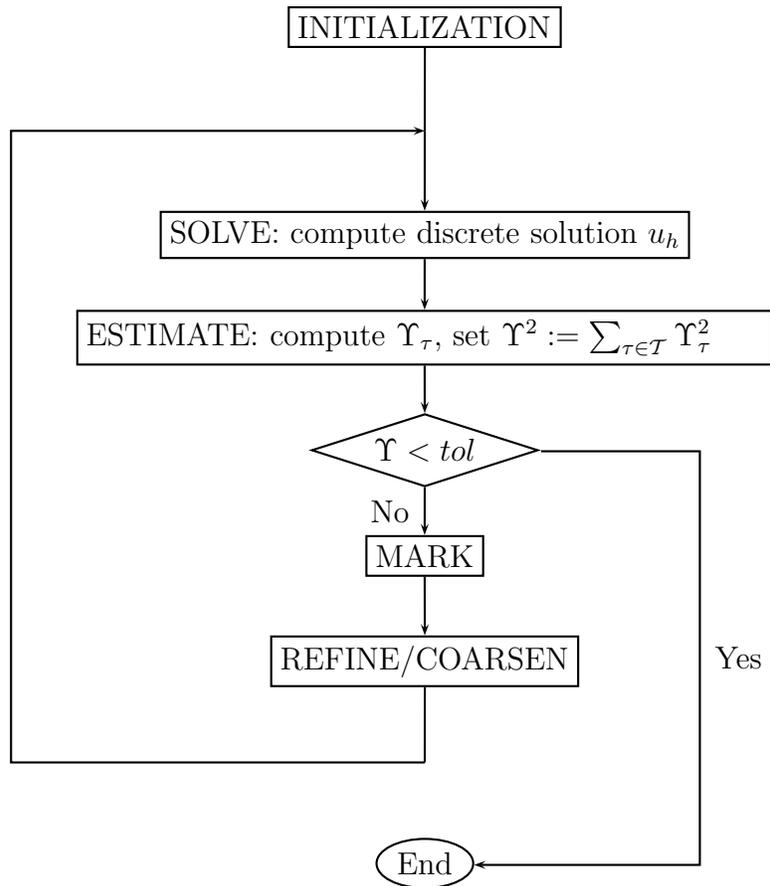


Figure 7.1: Flowchart of adaptive algorithm for static problems

Algorithm 7.1 (Adaptive Algorithm for Evolution Problems) Start with k_0 , \mathcal{T}_0 , U_h^0 .

(i) Compute initial error indicators for Υ_{init} . If $\Upsilon_{\text{init}}(\tau)$ is too large, refine τ .

Repeat (i) if necessary.

For $n \geq 0$ and $t_n \leq T$

(a) solve for U_h^n and compute error indicators for $\tau \in \mathcal{T}_n$

if Υ_{time}^n is large, reduce time step k_n and goto (a)

(b) for every $\tau \in \mathcal{T}_n$

if $\Upsilon_{\text{space}}^n(\tau)$ is too large, refine τ

if $\Upsilon_{\text{space}}^n(\tau) + \Upsilon_{\text{coarse}}^n(\tau)$ is too small, coarsen τ (if possible)

(c) if the mesh was changed

solve for U_h^n and compute error indicators again

if Υ_{time}^n is too large, reduce k_n and goto (a)

if $(\sum_{\tau \in \mathcal{T}_n} (\Upsilon_{\text{space}}^n(\tau))^2)^{\frac{1}{2}}$ is too large, goto (b)

otherwise, accept \mathcal{T}_n and U_h^n

(d) if Υ_{time}^n is small, enlarge k_{n+1}

(e) let $t_{n+1} = t_n + k_{n+1}$ and $n = n + 1$

Algorithm 7.1 is a modification of the algorithm originally proposed by Nochetto et al. [112] for the Stefan problem.

7.1.3 Convergence and Optimality

Even though adaptivity has been a successful tool of engineering and scientific computing for more than three decades, the convergence analysis is rather recent. Dörfler [54] introduced a crucial marking strategy, which will be discussed in §7.3, and proved strict energy error reduction for the Laplacian provided the initial mesh is sufficiently fine. Morin, Nochetto, and Siebert [106, 107] showed that energy error reduction cannot be expected in general by a counter-example, studied the role of data oscillation, and prove convergence without assumptions on the initial mesh. Later Mekchay and Nochetto [101] generalized this convergence result to general

second order elliptic operators.

Quasi-optimal convergence rates for adaptive finite element method for the Laplace equation were first shown by Binev, Dahmen and DeVore [22] with the help of an artificial coarsening step. In [22], the energy error decay in terms of number of degrees of freedom (DOF) is proved to be quasi-optimal, namely as dictated by nonlinear approximation theory [53]. The coarsening step was later removed by Stevenson [129], still for the Laplacian, at the expense of an inner loop to reduce oscillation. More recently, Cascon et al. [44] proposed a simple and practical adaptive algorithm, which avoids marking by oscillation, and proved a contraction property and quasi-optimal convergence rate for general second-order elliptic equations.

For obstacle problems, convergence and optimality are still in their early stages. To the best of our knowledge, the only existing convergence result (without rate) was given by Siebert and Veerer [127] for piecewise linear constraints. This topic deserves further study. For elliptic problems with integral operators as well as time-dependent problems, convergence and optimality are still to be developed. For linear parabolic problems, Chen and Feng [48] gave an adaptive algorithm allowing time-space adaptation and proved error reduction at one time step; the compound effect in time is however missing.

7.2 Estimate

The ESTIMATE step provides local information of the error which guide the adaptive algorithm to generate *optimal* meshes. An accepted principle for adaptive algorithms is the error equidistribution, i.e. local error on each element has about the same magnitude. Since error is not known, the next best thing is to equidistribute the local error indicator instead of real local error. A posteriori error estimations discussed in the previous chapter can guide us to design local error indicators.

We first define the following nodal-based local error indicators:

- Initial error indicator:

$$\Upsilon_0(\tau) = \|u_0 - U_h^0\|_{L^2(\tau)}.$$

- Space error indicator:

$$\Upsilon_h^n(z) := \frac{1}{\sqrt{T}} \left\{ (\Upsilon_{h,j}^n(z))^2 + (\Upsilon_{h,i}^n(z))^2 + (\Upsilon_{h,f}^n(z))^2 \right\}^{\frac{1}{2}}$$

where we define the nodal-based error indicators in (6.70) as follows

$$\begin{aligned} \text{jump residual} \quad \Upsilon_{h,j}^n(z) &:= \begin{cases} \|h^{\frac{1}{2}} J(U_h^n)\|_{L^2(\gamma_z)} & z \in \mathcal{F}_h^n \cup \mathcal{N}_h^n \\ 0 & z \in \mathcal{C}_h^n \end{cases} \\ \text{interior residual} \quad \Upsilon_{h,i}^n(z) &:= \begin{cases} \|h^{s+\frac{d}{2}-\frac{d}{p}} (R(U_h^n) - R_z^n) \psi_z\|_{L^p(\omega_z)} & z \in \mathcal{F}_h^n \cup \mathcal{N}_h^n \\ 0 & z \in \mathcal{C}_h^n \end{cases} \\ \text{free boundary term} \quad \Upsilon_{h,f}^n(z) &:= \begin{cases} -s_z^n d_z^n & z \in \mathcal{F}_h^n \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

- Time error indicator: Since the time error estimator is not local, we use the following heuristic local time error indicator

$$\Upsilon_k^n := \frac{1}{\sqrt{T}} \left\| \|U_h^n - \mathcal{I}_{n-1}^n U_h^{n-1}\| \right\|.$$

- Coarsening error indicator

$$\Upsilon_c^n(\tau) := \frac{1}{\sqrt{T}} \left\| \|U_h^{n-1} - I_{n-1}^n U_h^{n-1}\| \right\|_{\tau}.$$

- Obstacle consistency error indicators

$$\begin{aligned} \Upsilon_{\chi,h}^n(\tau) &:= \frac{1}{\sqrt{T}} \left\| \|(\chi_h^n - I_{n-1}^n U_h^{n-1})^+\| \right\|_{\tau} \\ \Upsilon_{\chi,k}^n &:= \frac{1}{\sqrt{T}} \left(\int_{t_{n-1}}^{t_n} \left\| \|(\chi - U_h)^+\| \right\|^2 dt \right)^{\frac{1}{2}}. \end{aligned}$$

Remark 7.2 (From Nodal-based to Element-based Indicators) Note that in Algorithm 7.1, we use element-based error indicators. However, we define nodal-based space error indicators above. In fact, we can define element-based space error indicator easily by

$$\Upsilon_h^n(\tau) := \max_{z \in \mathcal{P}_h^n \cap \tau} \Upsilon_h^n(z)$$

or we define it by averaging

$$\Upsilon_h^n(\tau) := \frac{\sum_{z \in \mathcal{P}_h^n \cap \tau} \Upsilon_h^n(z)}{d+1}.$$

Remark 7.3 (Negative Norm Estimators) We do not implement the error estimator terms $\|\partial_t(\chi - U_h)^+\|_*$ and $\|\overline{F} - f\|_*$ in dual norms. We would expect the first one to be at least of the same order as $\|(\chi - U_h)^+\|$ (see example 8.1.4 for numerical evidence) and the second term to be of higher order than $\mathcal{O}(h)$.

Now we can define error indicators needed in Algorithm 7.1:

$$\begin{aligned} \Upsilon_{\text{init}}(\tau) &:= \Upsilon_0(\tau) \\ \Upsilon_{\text{time}} &:= \Upsilon_k^n + \Upsilon_{\chi,k}^n \\ \Upsilon_{\text{space}}(\tau) &:= \Upsilon_h^n(\tau) + \Upsilon_{\chi,h}^n(\tau) \\ \Upsilon_{\text{coarse}}(\tau) &:= \Upsilon_c^n(\tau). \end{aligned}$$

7.3 Mark

The MARK step is based on the local error indicator given by ESTIMATE. The marking strategy could be based on the elements, edges, or nodes. Here we only consider the element based error indicators defined in previous section. To achieve error equidistribution, it is clear that elements with a large local error indicator should be refined, while elements with a small indicator need to be coarsened. There are several marking strategies have been proposed in the literature. We now review them very briefly.

7.3.1 Maximum Strategy

A very simple strategy is to mark those elements with an error indicator close to the largest indicator. More precisely, given a threshold $\theta \in (0, 1)$, we mark all elements $\tau \in \mathcal{T}$ with

$$\Upsilon_\tau \geq \theta \max_{\tau \in \mathcal{T}} \Upsilon_\tau$$

for refinement. See [123, Algorithm 1.18].

7.3.2 Equidistribution Strategy

This marking strategy is based on an average idea. Assume the number of mesh elements in \mathcal{T} is $\#\mathcal{T}$. Then we refine all elements with error indicator

$$\Upsilon_\tau \geq \theta \frac{\sum_{\tau \in \mathcal{T}} \Upsilon_\tau}{\#\mathcal{T}},$$

with a parameter $\theta \in (0, 1)$. See [123, Algorithm 1.19].

7.3.3 Dörfler's Marking Strategy

It is not clear whether an adaptive algorithm converges or even terminates within a prescribed tolerance. Dörfler [54] proposed a marking strategy with guaranteed energy error reduction provided the initial mesh is fine enough; it is the so-called *guaranteed error reduction strategy* (GERS). The idea of GERS is to mark a portion of elements such that their contribution exceeds a percentage of the total, namely $\theta \sum_{\tau \in \mathcal{T}} \Upsilon_\tau$ where $\theta \in (0, 1)$ is a fixed parameter. To introduce as few degrees of freedom as possible, we should mark those elements with largest local indicators. For details, see [123, Algorithm 1.20].

7.4 Refine/Coarsen

Several refinement strategies in 2d and 3d are widely used. One such an example is regular refinement or red-green refinement [19], which divides every triangle into four in 2d (see Figure 7.2) and every tetrahedron into eight tetrahedra in 3d. One problem with this strategy in adaptive mesh refinement is the hanging nodes (leading to non-conforming meshes) introduced by local refinement. Additional refinement (the so-called green closure) is necessary to remove the hanging nodes (this becomes difficult in 3d though). One complication is that, before further refinement, the green refinement has to be removed to keep shape-regularity.

An alternative way is the bisection scheme introduced by Mitchell [103] for 2d and Bänsch [14] (iterative algorithm) and Kossaczky [88] (recursive algorithm) for 3d. The recursive bisection scheme for 2d and 3d are proved to terminate in



Figure 7.2: Regular refinement. Left: red refinement and hanging nodes; Right: green closure.

finite steps and keep shape-regularity (see [103, 88]). In 2d, one can either choose to bisect the longest edge (Longest Edge Bisection) or to bisect the edge opposite to the newest vertex of each element (Newest Vertex Bisection).

We only consider the newest vertex bisection in 2d and the corresponding bisection method by Kossaczky in 3d. Next we describe the newest vertex bisection for 2d as well as the corresponding coarsening algorithm in detail.

7.4.1 Newest Vertex Bisection in 2d

We first give a brief review of the newest vertex bisection method. Given a shape-regular grid or triangulation \mathcal{T} of $\Omega \subset \mathbb{R}^2$, we label one vertex of each element $\tau \in \mathcal{T}$ as the *newest vertex*. The opposite edge of newest vertex is called the *refinement edge*. This process is called a *labeling* of \mathcal{T} .

Starting with a labeled initial grid \mathcal{T}_0 , newest vertex bisection follows the rules:

1. An element (*father*) is bisected to generate two new elements (*children*) by connecting the newest vertex with the midpoint of the refinement edge;
2. The new vertex created at the midpoint of the refinement edge is labeled as the newest vertex of each child.

Once the labeling is done for an initial grid, the subsequent grids inherit labels according to the second rule so that the bisection process can proceed. Sewell [126] showed that all the descendants of an original element fall into at most four similarity

classes and hence grids obtained by newest vertex bisection is uniformly shape-regular.

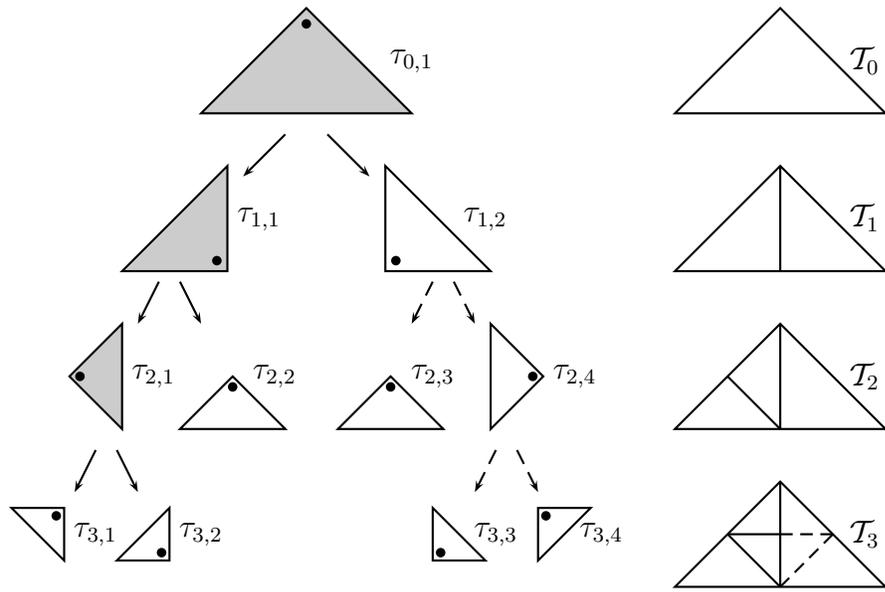


Figure 7.3: Bisection tree (left) and its corresponding grids (right).

We now give an example to illustrate the bisection procedure. In Figure 7.3, we start from an initial grid \mathcal{T}_0 with only one element $\tau_{0,1}$. The ‘dot’ close to a vertex indicates that that vertex is the newest vertex of that element. The generation of each element in the initial grid is defined to be 0; once an element is bisected, the generations of both children (the new elements) are defined as one plus the generation of their father (the old element). From now on, the generation of an element $\tau \in \mathcal{T}$ will be denoted by $g(\tau)$. We denote by $\tau_{i,j}$ the j -th element of generation i , namely $i = g(\tau_{i,j})$.

Suppose the adaptive method marks the element $\tau_{0,1}$ for bisection (we indicate a marked element by drawing it in light gray). After one step of newest vertex bisection, the new grid \mathcal{T}_1 contains two elements $\tau_{1,1}$ and $\tau_{1,2}$ which are the *siblings*. Suppose $\tau_{1,1}$ is bisected to produce the grid \mathcal{T}_2 and later $\tau_{2,1}$ to give rise to \mathcal{T}_3 . However, when $\tau_{2,1}$ is bisected, to keep conformity, we need to bisect $\tau_{1,2}$ twice according to the rules of newest vertex bisection. The dashed lines in the tree as well as in the grid in Figure 7.3 mean they are generated due to the conformity

requirements. From the discussion above, it is easy to see that the bisection algorithm generates nested meshes with the hierarchical structure of binary trees; each binary tree corresponds to an element of the initial triangulation \mathcal{T}_0 (often called *macro-elements*).

7.4.2 Coarsening Algorithm

The bisection procedure is fully revertible using a recursive coarsening algorithm developed in [88]. Let us still use Figure 7.3 to illustrate the algorithm. In the final grid \mathcal{T}_3 , suppose we want to coarsen the element $\tau_{2,3}$, the algorithm will first find its neighbor $\tau_{3,4}$ and it should be intelligent enough to tell that these two elements are not siblings with the same father and cannot be glued together. So the algorithm will then try to coarsen $\tau_{3,4}$ first. This can be done in a recursive manner. The element $\tau_{3,3}$ is found to be the sibling of $\tau_{3,4}$. Once the algorithm glue $\tau_{3,3}$ and $\tau_{3,4}$ together to get $\tau_{2,4}$ back again, the grid becomes not conforming. To keep conformity, the other neighbor (not sibling) of $\tau_{3,4}$, i.e. $\tau_{3,1}$, and its own sibling should be glued together (if there is a problem with this step as before, do the same recursive step for $\tau_{3,1}$ first). Once this conformity step has been completed, the algorithm returns to $\tau_{2,3}$ and glue it with its sibling $\tau_{2,4}$ to obtain \mathcal{T}_2 . To allow the algorithm to traverse easily to its neighbors and so on, the bisection tree is needed (for details, see, for example, [88, 123]).

7.4.3 Compatible Bisection

We denote the set of nodes (including boundary nodes) of a grid \mathcal{T} by \mathcal{P}_h and the set of edges or sides by \mathcal{S}_h . We denote the cardinality, i.e. number of nodes in \mathcal{P}_h , by $\#\mathcal{P}_h$. Let \mathcal{T} be a labeled grid. For any $\tau \in \mathcal{T}$, let S_τ be the refinement edge of τ and let

$$F(\tau) = \begin{cases} \tau' & S_\tau \subset \tau' \in \mathcal{T} \\ \emptyset & S_\tau \subset \partial\Omega \end{cases}$$

be the element of \mathcal{T} (if exists) which shares the same refinement edge of S_τ with τ .

An element τ is called *compatible* if $F(\tau) = \emptyset$ or $F(F(\tau)) = \tau$. A labeled grid \mathcal{T} is called *compatible* if every element in \mathcal{T} is compatible and the labeling of \mathcal{T} is, in turn, called a *compatible labeling*. Given a compatible initial grid \mathcal{T}_0 , we define

$$\mathcal{S}(\mathcal{T}_0) := \{\mathcal{T} \mid \mathcal{T} \text{ is obtained from } \mathcal{T}_0 \text{ by newest vertex bisections}\}.$$

and a subset of $\mathcal{S}(\mathcal{T}_0)$

$$\mathcal{A}(\mathcal{T}_0) := \{\mathcal{T} \in \mathcal{S}(\mathcal{T}_0) \mid \mathcal{T} \text{ is conforming}\}.$$

Notice that the difference between $\mathcal{S}(\mathcal{T}_0)$ and $\mathcal{A}(\mathcal{T}_0)$ is that a grid in $\mathcal{S}(\mathcal{T}_0)$ could be non-conforming. We shall consider the coarsening of grids in the class $\mathcal{A}(\mathcal{T}_0)$.

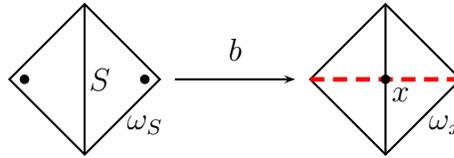


Figure 7.4: Compatible bisection b .

For a compatible element τ , its refinement edge is called a *compatible edge*. Let ω_S be the patch of elements sharing the side $S \in \mathcal{S}$. If S is compatible, we call the bisection of ω_S a *compatible bisection* and denote by b . More precisely, let x be the midpoint of S , then b is understood as a map $b : \omega_S \rightarrow \omega_x$, where the patch ω_S consists of coarser elements and ω_x of fine elements; see Figure 7.4. In two dimensions, a compatible bisection b only has two possible configurations. One is bisecting an interior compatible edge. In this case, the patch ω_S is a quadrilateral. Another case is bisecting a boundary compatible edge and ω_S is a triangle. See Figure 7.5.

7.4.4 Bisection Grids Revisited

Let $\overline{\mathcal{T}}_l \in \mathcal{A}(\mathcal{T}_0)$ be the grid generated from a compatible initial grid \mathcal{T}_0 after l times of uniform refinement (meaning refine each element once every time). Apparently, from the previous subsection, $\overline{\mathcal{T}}_l$ can be viewed as full binary trees (one

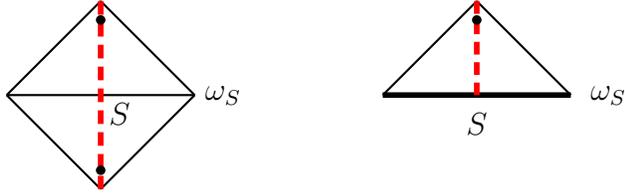


Figure 7.5: Compatible bisection of S . Left: interior edge; right: boundary edge.

tree for each element in the initial grid). Bisection guarantees that the sequence $\{\bar{\mathcal{T}}_l\}$ is shape-regular and quasi-uniform [103]. Assuming that the initial mesh \mathcal{T}_0 is shape-regular with meshsize \bar{h}_0 , we can see the meshsize of $\bar{\mathcal{T}}_l$ is quasi-uniform. We denote the meshsize of \mathcal{T}_l by \bar{h}_l .

A triangulation $\mathcal{T} \in \mathcal{A}(\mathcal{T}_0)$ can be viewed as the result of a sequence of compatible bisections applied on the initial grid \mathcal{T}_0 with compatible initial labeling [45, 47]. Formally, we can denote it by

$$\mathcal{T} = \mathcal{T}_0 + b_1 + \cdots + b_m.$$

Now we use the grid \mathcal{T}_3 in Figure 7.3 as an example to illustrate this. We can view \mathcal{T}_3 as the result of applying four compatible bisections, b_1, \dots, b_4 , on \mathcal{T}_0 ; see Figure 7.6. The sequence $\{b_1, b_2, b_3, b_4\}$ is called a *compatible bisection sequence*. Notice that

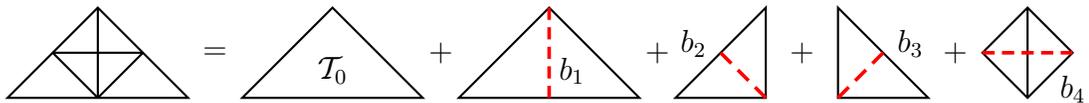


Figure 7.6: Decomposition of a bisection grid.

the order of b_2 and b_3 could be interchanged without changing the final grid. This means that there might be several different adaptive paths resulting in a particular final bisection grid in adaptive algorithms. The order of the bisection sequence does not imply generation information of bisections.

Let $L := \max_{\tau \in \mathcal{T}} g(\tau)$ be the maximum generation among all elements in $\mathcal{T} \in \mathcal{A}(\mathcal{T}_0)$. Then $\bar{\mathcal{T}}_L$ is a set of full binary trees (one for each macro-element) of depth $L + 1$. On the other hand, a locally refined mesh \mathcal{T} of depth $\leq L + 1$ is a

subtree of $\overline{\mathcal{T}}_L$ and can be embedded into $\overline{\mathcal{T}}_L$. With our notation, it is easy to see that $h_{\min}(\mathcal{T}) \approx \bar{h}_L$.

Remark 7.4 (Simple Bisection and Coarsening Algorithms) Exploiting this new view of bisection grids, Chen and Zhang [47] proposed a simple coarsening strategy for 2d problems. This coarsening strategy is implemented in the package AFEM@matlab [46].

7.5 Solve

It has been shown in Chapter 4 that we need to solve a discrete variational inequality (4.20) at each time step. As we discussed in §4.3.3 the discrete variational inequality (4.20) can be written as the following finite-dimensional linear complementarity problem (LCP)

$$\mathbf{A}\vec{\mathbf{U}} \geq \vec{\mathbf{F}}, \quad \vec{\mathbf{U}} \geq \vec{\mathbf{X}}, \quad (\mathbf{A}\vec{\mathbf{U}} - \vec{\mathbf{F}})^T (\vec{\mathbf{U}} - \vec{\mathbf{X}}) = 0; \quad (7.2)$$

see also (4.21). The subject of finite-dimensional variational inequalities and complementarity problems and their applications in engineering and economics have received intensive attention for over more than three decades. We refer to the review paper by Ferris and Pang [70] and the references therein for a comprehensive overview of the importance of linear and nonlinear complementarity problems in various application areas. For more general variational inequalities, we refer to the monograph by Facchinei and Pang [63, 64].

Here we will only mention some new methods designed especially for discretization of obstacle problems. A classical way to solve LCP is the projected successive over-relaxation (PSOR) method by Cryer [52]. For elliptic symmetric obstacle problems, different multigrid and domain decomposition techniques have been developed (see Tai [130] and the references therein for a quick review). Among them, typical examples include the full approximation scheme (FAS) [28], monotone multigrid (MMG) methods [97, 85, 87], multigraph interior point methods [13], and subspace correction methods [131, 9, 130].

7.5.1 Subspace Correction Methods for Obstacle Problems

Multigrid and domain decomposition methods have been studied extensively for linear partial differential equations. Multigrid methods and conjugate gradient methods with multilevel preconditioners are among the most efficient numerical methods for solving linear systems arising from elliptic PDEs. They can be analyzed under the general framework of space decomposition and subspace correction; see Xu [144] and the references therein for details.

Usually, subspace correction methods can be divided into two categories: parallel subspace correction (PSC) methods and successive subspace correction (SSC) methods. PSC methods are also called additive methods because they make corrections in each subspace simultaneously. They are suitable for parallel computing and preconditioning because of this nature. On the contrary, SSC methods make corrections in one subspace at a time and are often called multiplicative methods. Detailed information on the convergence theory as well as implementation for both PSC and SSC can be found in Xu [145].

Recently, the subspace correction framework has been extended to nonlinear convex minimization problems by Tai and Xu [132]. They considered a nonlinear convex optimization problem and proved global linear convergence rate for PSC and SSC under some assumptions on the subspace decomposition. Later this technique has been applied to develop domain decomposition and multigrid methods for variational inequalities [131, 9]. Furthermore, a constraint decomposition technique was introduced by Tai [130] to improve the efficiency of the methods. In this section, we discuss the constraint decomposition methods for obstacle problems.

We consider the energy minimization problem

$$\min_{v \in \mathbb{K}} \mathcal{J}(v), \tag{7.3}$$

where $\mathcal{J} : \mathbb{K} \subset \mathbb{V} \rightarrow \mathbb{R}$ is the convex functional defined in Problem 1.3 over the finite dimensional convex set

$$\mathbb{K} := \{v \in \mathbb{V}(\mathcal{T}) \mid v \geq 0\}.$$

Note that the algorithms discussed in this section could be generalized to problems with more general obstacles.

We decompose the space \mathbb{V} into a sum of subspaces \mathbb{V}_i , i.e.

$$\mathbb{V} = \mathbb{V}_1 + \cdots + \mathbb{V}_m = \sum_{i=1}^m \mathbb{V}_i. \quad (7.4)$$

Once we have the space decomposition (7.4), we can further decompose the feasible set \mathbb{K} as follows

$$\mathbb{K} = \mathbb{K}_1 + \cdots + \mathbb{K}_m = \sum_{i=1}^m \mathbb{K}_i \quad \mathbb{K}_i \subset \mathbb{V}_i \quad (i = 1, \dots, m), \quad (7.5)$$

where \mathbb{K}_i are convex and closed in \mathbb{V}_i .

There are two possibilities to construct numerical methods: one based on (7.4) and the other based on (7.5). To simplify the presentation, we only consider SSC versions of the algorithms; PSC versions can be constructed similarly (see [9, 130] for details).

We first look at the first possibility: an algorithm based on (7.4).

Algorithm 7.5 (Successive Space Correction Method) Given an initial guess $u \in \mathbb{K}$:

Let $w^{(0)} = u$

For $i = 1 : m$

$$d_i = \operatorname{argmin} \{ \mathcal{J}(w^{(i-1)} + d_i) \mid w^{(i-1)} + d_i \in \mathbb{K} \text{ and } d_i \in \mathbb{V}_i \}$$

$$\text{Let } w^{(i)} = w^{(i-1)} + d_i$$

End For

Let $w = w^{(m)}$ and use w as the initial guess to start the iteration again.

This is a natural extension of the SSC algorithm for unconstrained convex minimization problems [132]. On each subspace, we need to keep the new iteration $w^{(i)}$ in the feasible set \mathbb{K} . To do this, the computational cost at each iteration might be big even if \mathbb{V}_i is only low dimensional (as would correspond to a coarse mesh).

One can then modify this algorithm using the feasible set decomposition, or equivalently constraint decomposition (7.5).

Algorithm 7.6 (SSC Constraint Decomposition Method) Given an initial guess $u \in \mathbb{K}$:

Decompose $u = \sum_{i=1}^m u_i$, $u_i \in \mathbb{K}_i$ and let $w^{(0)} = u$

For $i = 1 : m$

$d_i = \operatorname{argmin} \{ \mathcal{J}(w^{(i-1)} + d_i) \mid u_i + d_i \in \mathbb{K}_i \text{ and } d_i \in \mathbb{V}_i \}$

Let $w^{(i)} = w^{(i-1)} + d_i$

End For

Let $w = w^{(m)}$ and use w as the initial guess to start the iteration again.

Remark 7.7 (Local Obstacle) The idea of using local obstacle to reduce the computational cost of local problems is not new. It has been explored by Mandel [97] and then extended by Kornhuber [85, 86]. However, the constraint decomposition method is essentially different from the monotone multigrid methods in its philosophy. We will discuss this later in Remarks 7.14, 7.15 and 7.21.

Remark 7.8 (Feasibility) For both Algorithm 7.5 and 7.6, we need a feasible initial guess to start with. It is clear that each iteration $w^{(i)}$ ($i = 1, \dots, m$) stays in the feasible set \mathbb{K} because of (7.5).

The main difference between Algorithm 7.5 and 7.6 relies on the fact that, for the latter, we only solve a minimization problem in $\mathbb{K}_i \subset \mathbb{V}_i$ at each iteration. This is usually just an one-dimensional minimization problem and is cheap to solve. On the other hand, the conditions $u_i \in \mathbb{K}_i$ ($i = 1, \dots, m$) is more restrictive for decomposition of u than $\sum_{i=1}^m u_i \in \mathbb{K}$ of course. We only consider Algorithm 7.6 here in this thesis.

7.5.2 Convergence Rate of SSC-CDM Methods

We shall prove the linear convergence rate of the SSC constraint decomposition method (SSC-CDM), Algorithm 7.6. This presentation follows the idea of Tai [130] except tuned to the way Algorithm 7.6 is written (which is different than [130]).

First of all, we make two assumptions on the decomposition: the first is stability of the decomposition and the second is the strengthened Cauchy-Schwarz (SCS) inequality.

Assumption 7.9 (Assumptions on Decomposition) We assume that

1. For any $u, v \in \mathbb{K}$, there exist a constant $C_1 > 0$ and decompositions $u = \sum_{i=1}^m u_i$ with $u_i \in \mathbb{K}_i$, $v = \sum_{i=1}^m v_i$ with $v_i \in \mathbb{K}_i$ such that

$$\left(\sum_{i=1}^m \|u_i - v_i\|^2 \right)^{\frac{1}{2}} \leq C_1 \|u - v\|; \quad (7.6)$$

2. There exists $C_2 > 0$ such that

$$\sum_{i,j=1}^m |\langle \mathcal{J}'(w_{ij} + v_i) - \mathcal{J}'(w_{ij}), \tilde{v}_j \rangle| \leq C_2 \left(\sum_{i=1}^m \|v_i\|^2 \right)^{\frac{1}{2}} \left(\sum_{j=1}^m \|\tilde{v}_j\|^2 \right)^{\frac{1}{2}}, \quad (7.7)$$

for any $w_{ij} \in \mathbb{V}$, $v_i \in \mathbb{V}_i$, and $\tilde{v}_j \in \mathbb{V}_j$.

Remark 7.10 (Stable Decomposition) The counterpart of the first assumption for unconstrained case is usually called stability of the subspace decomposition. This is a statement about lack of redundancy in the decomposition, i.e. the decomposition is *almost* orthogonal.

Remark 7.11 (Strengthened Cauchy-Schwarz Inequality) The second assumption is the so-call strengthened Cauchy-Schwarz inequality for nonlinear problems.

If these two assumptions in Assumption 7.9 are satisfied, then the SSC-CDM is globally convergent and has linear convergence rate.

Theorem 7.12 (Convergence Rate of SSC-CDM) *If Assumption 7.9 is satisfied, then Algorithm 7.6 converges and*

$$\frac{\mathcal{J}(w) - \mathcal{J}(u^*)}{\mathcal{J}(u) - \mathcal{J}(u^*)} \leq 1 - \frac{1}{(\sqrt{1 + C_0} + \sqrt{C_0})^2}, \quad (7.8)$$

where u^* is the solution of (7.3) and $C_0 = 2C_2 + C_1^2 C_2^2$.

Remark 7.13 (Measure of Error) Here, the error is measured by $\mathcal{J}(u) - \mathcal{J}(u^*)$. This is natural for energy minimization problem. In fact, by definition,

$$\begin{aligned}\mathcal{J}(u) - \mathcal{J}(u^*) &= \frac{1}{2} \|u\|^2 - \frac{1}{2} \|u^*\|^2 - \langle f, u - u^* \rangle \\ &= \frac{1}{2} \|u - u^*\|^2 + a(u^*, u - u^*) - \langle f, u - u^* \rangle \\ &= \frac{1}{2} \|u - u^*\|^2 - \langle \lambda(u^*), u - u^* \rangle.\end{aligned}$$

For any feasible u , the second term on the right-hand side $\langle \lambda(u^*), u - u^* \rangle$ is non-positive. Hence, if $\mathcal{J}(u) - \mathcal{J}(u^*) = 0$, then $\|u - u^*\| = 0$.

Remark 7.14 (Global and Monotone Convergence) Notice that the previous theorem implies that energy \mathcal{J} is strictly decreasing in Algorithm 7.6. Furthermore, the convergence rate is globally linear starting from any feasible initial guess. This is different than the asymptotic linear convergence rate of monotone multigrid methods [85, 86].

Remark 7.15 (Non-degeneracy Assumption) There is no need to assume that the strict complementarity condition is satisfied by the discrete problem (non-degenerate assumption) as for monotone multigrid methods [85, Lemma 2.2]. Numerical experiments show the method is stable for degenerate problems also; see Table 8.18.

Remark 7.16 (General Convex Minimization) For our purpose, we only consider Problem 1.3 here. The methods discussed here can be generalized to convex minimization problems with strongly convex and Gâteaux differentiable objective functionals.

We now give several lemmas in preparation to prove Theorem 7.12.

Lemma 7.17 (First Order Optimal Condition) *For each $i = 1, \dots, m$, we have*

$$\left\langle \mathcal{J}'(w^{(i)}), \tilde{d}_i - d_i \right\rangle \geq 0 \quad \forall u_i + \tilde{d}_i \in \mathbb{K}_i.$$

Proof. Note that both $u_i + d_i$ and $u_i + \tilde{d}_i$ are in \mathbb{K}_i . Therefore $u_i + (1 - \alpha)d_i + \alpha\tilde{d}_i \in \mathbb{K}_i$ for any $0 \leq \alpha \leq 1$ since \mathbb{K}_i is a convex set. We then consider the minimization problem

$$\min_{0 \leq \alpha \leq 1} \mathcal{J}(w^{(i-1)} + (1 - \alpha)d_i + \alpha\tilde{d}_i).$$

From the first order optimality condition, it is then clear, for $i = 1, \dots, m$, that

$$\langle \mathcal{J}'(w^{(i)}), \tilde{d}_i - d_i \rangle \geq 0 \quad \forall u_i + \tilde{d}_i \in \mathbb{K}_i.$$

Hence we have the desired inequality. \blacksquare

Lemma 7.18 (Monotonicity) *In Algorithm 7.6, the energy is decreasing and*

$$\mathcal{J}(u) - \mathcal{J}(w) \geq \frac{1}{2} \sum_{i=1}^m \|d_i\|^2.$$

Proof. For any $v, \tilde{v} \in \mathbb{K}$, it is easy to see that

$$\mathcal{J}(\tilde{v}) - \mathcal{J}(v) = \langle \mathcal{J}'(v), \tilde{v} - v \rangle + \frac{1}{2} \|\tilde{v} - v\|^2. \quad (7.9)$$

For $i = 1, \dots, m$, we have that $w^{(i-1)}$ and $w^{(i)}$ are both in \mathbb{K} . Hence, by applying (7.9) and Lemma 7.17 with $\tilde{d}_i = 0$, we get

$$\mathcal{J}(w^{(i-1)}) - \mathcal{J}(w^{(i)}) = -\langle \mathcal{J}'(w^{(i)}), d_i \rangle + \frac{1}{2} \|d_i\|^2 \geq \frac{1}{2} \|d_i\|^2.$$

Then $\mathcal{J}(u) - \mathcal{J}(w) = \sum_{i=1}^m \mathcal{J}(w^{(i-1)}) - \mathcal{J}(w^{(i)})$ gives the lower bound of the energy reduction. \blacksquare

This lemma ensures the algorithm will result in strict energy reduction when $d_i \neq 0$. To prove the convergence theorem, we are going to bound $\|d_i\|$ from below by the error in energy. The following lemma basically says if one cannot make any progress in a step, i.e. $\sum_{i=1}^m \|d_i\|^2 = 0$, then one has obtained the exact solution; otherwise, one can always reduce the energy using Algorithm 7.6.

Lemma 7.19 (Error in Energy) *Suppose $u^* \in \mathbb{K}$ is the optimal solution. The error in energy after one loop of CDM-SSC method satisfies*

$$\mathcal{J}(w) - \mathcal{J}(u^*) \leq C_2 \sum_{i=1}^m \|d_i\|^2 + C_1 C_2 \left(\sum_{i=1}^m \|d_i\|^2 \right)^{1/2} \|u - u^*\|.$$

Proof. We first recall that Assumption 7.9 (1) implies the existence of decompositions $u^* = \sum_{i=1}^m u_i^*$ and $u = \sum_{i=1}^m u_i$ with $u_i^*, u_i \in \mathbb{K}_i$ satisfying (7.6). Taking $\tilde{v} = u^*$ and $v = w$ in (7.9), we arrive at

$$\mathcal{J}(w) - \mathcal{J}(u^*) \leq \langle \mathcal{J}'(w), w - u^* \rangle.$$

On the other hand, by Lemma 7.17, we obtain

$$\langle \mathcal{J}'(w^{(i)}), (u_i^* - u_i) - d_i \rangle \geq 0,$$

whence

$$\begin{aligned} \langle \mathcal{J}'(w), w - u^* \rangle &= \sum_{i=1}^m \langle \mathcal{J}'(w), u_i + d_i - u_i^* \rangle \\ &\leq \sum_{i=1}^m \langle \mathcal{J}'(w) - \mathcal{J}'(w^{(i)}), u_i + d_i - u_i^* \rangle \\ &= \sum_{i=1}^m \sum_{j=i}^m \langle \mathcal{J}'(w^{(j)}) - \mathcal{J}'(w^{(j-1)}), u_i + d_i - u_i^* \rangle. \end{aligned}$$

Using the strengthened Cauchy-Schwarz inequality (7.7), we then have

$$\langle \mathcal{J}'(w), w - u^* \rangle \leq C_2 \left(\sum_{i=1}^m \|d_i\|^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^m \|(u_i - u_i^*) + d_i\|^2 \right)^{\frac{1}{2}}.$$

Hence a consequence of the above inequality, the triangle inequality and the stability of the decomposition (7.6) is

$$\langle \mathcal{J}'(w), w - u^* \rangle \leq C_2 \left(\sum_{i=1}^m \|d_i\|^2 \right)^{\frac{1}{2}} \left[\left(\sum_{i=1}^m \|d_i\|^2 \right)^{\frac{1}{2}} + C_1 \|u - u^*\| \right].$$

This in turn gives the upper bound of the error in energy. ■

Now we are ready to prove the main convergence theorem.

Proof of Theorem 7.12. From Lemma 7.19, we can see that

$$\mathcal{J}(w) - \mathcal{J}(u^*) \leq C_2 \sum_{i=1}^m \|d_i\|^2 + C_1 C_2 \left(\sum_{i=1}^m \|d_i\|^2 \right)^{1/2} \|u - u^*\|.$$

Using the generalized triangle inequality, $ab \leq \frac{1}{2\varepsilon} a^2 + \frac{\varepsilon}{2} b^2$ with a constant $0 < \varepsilon < 1$,

the monotonicity Lemma 7.18, and (7.9) with $\tilde{v} = u$ and $v = u^*$, we obtain

$$\begin{aligned} \mathcal{J}(w) - \mathcal{J}(u^*) &\leq C_2 \sum_{i=1}^m \|d_i\|^2 + \left(\frac{C_1^2 C_2^2}{2\varepsilon} \sum_{i=1}^m \|d_i\|^2 + \frac{\varepsilon}{2} \|u - u^*\|^2 \right) \\ &\leq \left(2C_2 + \frac{C_1^2 C_2^2}{\varepsilon} \right) (\mathcal{J}(u) - \mathcal{J}(w)) + \varepsilon (\mathcal{J}(u) - \mathcal{J}(u^*)) \\ &\leq \frac{C_0}{\varepsilon} (\mathcal{J}(u) - \mathcal{J}(w)) + \varepsilon (\mathcal{J}(u) - \mathcal{J}(u^*)). \end{aligned}$$

Hence, it is easy to see that

$$\frac{\mathcal{J}(w) - \mathcal{J}(u^*)}{\mathcal{J}(u) - \mathcal{J}(u^*)} \leq \frac{C_0 \varepsilon^{-1} + \varepsilon}{1 + C_0 \varepsilon^{-1}} = \frac{C_0 + \varepsilon^2}{C_0 + \varepsilon}.$$

To minimize the right-hand side $f(\varepsilon) := (C_0 + \varepsilon^2)/(C_0 + \varepsilon)$, we find

$$f'(\varepsilon) = \frac{\varepsilon^2 + 2C_0\varepsilon - C_0}{(\varepsilon + C_0)^2}$$

and there exists a unique minimizer of $f(\varepsilon)$, $\varepsilon_* = \sqrt{C_0^2 + C_0} - C_0 \in (0, 1)$. By picking the optimal ε_* , we obtain the convergence result (7.8). \blacksquare

7.5.3 SSC-CDM on Adaptive Grids

We have proved in the previous subsection that the SSC-CDM method converges linearly if the space and constraint decompositions satisfy the assumptions in Assumption 7.9. In this section, we construct subspace decompositions for continuous piecewise linear finite element space $\mathbb{V} = \mathbb{V}(\mathcal{T})$ vanishing on the boundary of the polygonal domain Ω on an adaptive grid obtained by newest vertex bisection, \mathcal{T} . This is new because the original paper by Tai [130] assumes quasi-uniformity of the underlying meshes.

In Algorithm 7.6, once a space decomposition $\mathbb{V} = \sum_{i=1}^m \mathbb{V}_i$ is introduced, we need to decompose the feasible set $\mathbb{K} = \sum_{i=1}^m \mathbb{K}_i$ first and then decompose the current iterative solution u such that

$$u = \sum_{i=1}^m u_i \quad \text{and} \quad u_i \in \mathbb{K}_i \subset \mathbb{V}_i.$$

If there is no constraint, i.e. $\mathbb{K} = \mathbb{V}$, then it is clear that we can take $\mathbb{K}_i = \mathbb{V}_i$ for $i = 1, \dots, m$. The SSC-CDM algorithm is then reduced to the SSC method for unconstrained convex optimization problem in [132].

There are two ways to decompose the space \mathbb{V} which are proved to be efficient in practice: one is domain decomposition (DD) type, the other is multigrid (MG) type. Both were discussed in [130]. Here we shall focus on multigrid decomposition and remove the quasi-uniform assumption on the underlying grid as posed in [130]. Then we can apply this algorithm for symmetric elliptic variational inequalities on adaptive meshes.

Space and Constraint Decomposition

From now on, we assume that $\mathcal{T} \in \mathcal{A}(\mathcal{T}_0)$ can be decomposed in the following way as discussed in §7.4.3

$$\mathcal{T} = \mathcal{T}_0 + b_1 + \cdots + b_m,$$

where b_i 's are compatible bisections. We first introduce the multigrid space decomposition for \mathbb{V} . We denote the intermediate grids by

$$\mathcal{T}_i := \mathcal{T}_0 + b_1 + \cdots + b_i \quad i = 1, \dots, m,$$

and observe that $\mathcal{T}_i \in \mathcal{A}(\mathcal{T}_0)$. Define the nodal basis $\psi_{i,x} \in \mathbb{V}(\mathcal{T}_i)$ at node $x \in \mathcal{T}_i$. For the same geometric node x , we could have different nodal basis functions on different grids.

It is easy to see that there is a one-to-one correspondence between the compatible bisection b_i and a compatible refinement edge $S_i \in \mathcal{S}_h(\mathcal{T}_i)$. In turn, we also have a one-to-one correspondence between b_i and x_i , the middle point of S_i , when x_i first occur. Denote the support of ψ_{i,x_i} by ω_{i,x_i} and the subspaces associated with b_i by

$$\mathbb{V}_i := \{\psi_{i,x} \mid x \in \mathcal{P}_h(\mathcal{T}_i) \cap \omega_{i,x_i}\}. \quad (7.10)$$

If $\mathbb{V}_0 = \mathbb{V}(\mathcal{T}_0)$ is the space corresponding to the initial mesh \mathcal{T}_0 , then we have a subspace decomposition

$$\mathbb{V} = \sum_{i=0}^m \mathbb{V}_i.$$

Based on this subspace decomposition, there are infinitely many possibilities to decompose the feasible set \mathbb{K} when the subspace decomposition is fixed. We

do not consider the optimal way to choose such a constraint decomposition. We decompose $\mathbb{K} := \{v \in \mathbb{V} \mid v \geq 0\}$ into

$$\mathbb{K} = \sum_{i=0}^m \mathbb{K}_i \quad \text{and} \quad \mathbb{K}_i := \{v \in \mathbb{V}_i \mid v \geq 0\}. \quad (7.11)$$

We shall use the following notation for various kinds of local patches:

- $\omega_{i,x} := \bigcup \{\tau \mid x \in \tau, \tau \in \mathcal{T}_i\};$
- $\tilde{\omega}_{i,x} := \bigcup \{\omega_{i,y} \mid y \in \mathcal{P}(\mathcal{T}_i) \cap \omega_{i,x}\};$
- $\omega_{i,\tau} := \bigcup \{\omega_{i,y} \mid y \in \mathcal{P}(\mathcal{T}_i) \cap \tau\};$
- $\omega_i := \omega_{i,x_i};$
- $\tilde{\omega}_i := \tilde{\omega}_{i,x_i}.$

SSC-CDM Algorithm on Adaptive Grid

With the subspace and constraint decompositions discussed above, we can construct a practical SSC-CDM algorithm. The main difference between the SSC-CDM for the constrained minimization problems and the SSC method for unconstrained problems is that, in the former, we need to actually decompose each iterative solution $u \in \mathbb{K}$; on the contrary, in the latter, the decomposition is only for theoretical purposes. In fact, in SSC methods, one can think there is a decomposition of each iteration u . However the particular choice of decomposition will not change the next iteration w . On the contrary, for constrained minimization, the decomposition of u will affect the local obstacle in each subspace. This is because we need to compute

$$d_i = \operatorname{argmin} \{ \mathcal{J}(w^{(i-1)} + d_i) \mid u_i + d_i \geq 0 \text{ and } d_i \in \mathbb{V}_i \} \quad i = 1, \dots, m$$

to obtain $w^{(i)}$. We can see from the formula above that u_i is only needed to verify the constraint $u_i + d_i \geq 0$.

We first introduce a decomposition of u and then apply it to the SSC-CDM algorithm on adaptive grids. For $i = 1, \dots, m$ and any function $u \in \mathbb{V}$, we define an

operator $Q_i : \mathbb{V} \rightarrow \mathbb{V}(\mathcal{T}_{i-1})$ such that, for any node $x \in \mathcal{P}_h(\mathcal{T}_{i-1})$,

$$Q_i u(x) := \min_{y \in \omega_{i,x}} u(y) \quad (7.12)$$

Having defined $Q_i u$ at all nodes $\mathcal{P}_h(\mathcal{T}_{i-1})$ by (7.12), the rest of values of $Q_i u$ can then be obtained by interpolation since $Q_i u \in \mathbb{V}(\mathcal{T}_{i-1})$. Notice that Q_i 's are nonlinear operators, i.e. $Q_i u - Q_i v \neq Q_i(u - v)$.

Lemma 7.20 (Stability of Q_i) *Let $u, v \in \mathbb{V}$. For any node $x \in \mathcal{P}_h(\mathcal{T}_i)$ and any element $\tau \in \mathcal{T}_i$, we have*

$$h_\tau^{-1} \|Q_{i+1} u - Q_{i+1} v\|_{L^2(\tau)} \leq C_{d,\tau} \|u - v\|_{H^1(\omega_{i,\tau})},$$

where the constant $C_{d,\tau}$ depends on the meshsize

$$C_{d,\tau} := \begin{cases} C & d = 1 \\ C(1 + |\ln(h_\tau/h_{\min})|)^{\frac{1}{2}} & d = 2 \\ C(h_\tau/h_{\min})^{\frac{1}{2}} & d = 3. \end{cases} \quad (7.13)$$

Here C is a generic constant which is independent of the meshsize.

Proof. From the definition of Q_i 's, we have, for any $u, v \in \mathbb{V}$, that

$$\begin{aligned} \|Q_{i+1} u - Q_{i+1} v\|_{L^2(\tau)} &\lesssim \sum_{y \in \mathcal{P}_h(\mathcal{T}_i) \cap \tau} \|u - v\|_{L^\infty(\omega_{i,y})} |\tau| \\ &\lesssim h_\tau^{\frac{d}{2}} \|u - v\|_{L^\infty(\omega_{i,\tau})}. \end{aligned}$$

The result then follows directly from scaling argument and the classical discrete Sobolev inequality between L^∞ and H^1 ; see [27]. \blacksquare

Next we define a decomposition of u (see Figure 7.7):

$$u = \sum_{i=0}^m u_i, \quad (7.14)$$

where

$$u_m := u - Q_m u, \quad u_i := Q_{i+1} u - Q_i u \quad (i = 1, \dots, m-1), \quad u_0 = Q_1 u. \quad (7.15)$$

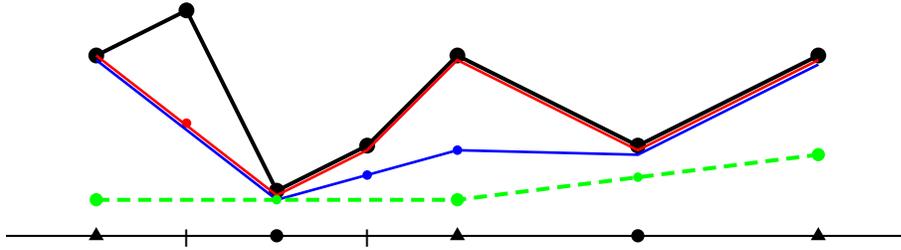


Figure 7.7: Decomposition of u .

Comparing these with the definitions (7.10) of \mathbb{V}_i and (7.11) of \mathbb{K}_i , we can easily see that

$$u_i \in \mathbb{K}_i \quad i = 0, 1, \dots, m.$$

We have specified all ingredients of Algorithm 7.6 and it can be now applied to symmetric elliptic obstacle problems. In practice, we can further decompose each \mathbb{V}_i by natural nodal basis decomposition

$$\mathbb{V}_i = \sum_{x \in \mathcal{P}_h(T_i) \cap \omega_i} \text{span}\{\psi_{i,x}\}.$$

Then at each step, we only need to solve a univariable simple constrained minimization problem which is easy.

Remark 7.21 (Different Philosophy Between SSC-CDM and MMG) Now we discuss a little bit about the different philosophy between the SSC-CDM method and the monotone multigrid methods (MMG).

- In MMG, we give the maximum freedom to high frequency corrections. This will in turn restrict the freedom of the low frequency corrections. Close to the free boundary, the standard MMG methods behaves more like a Gauss-Siedel method and has multigrid performance when the contact region has been resolved. To speed up the convergence, Kornhuber [85] suggested a modified MMG method. This modification, on the other hand, causes computational overhead.

- For the SSC-CDM method, the convergence theorem actually suggests we leave more freedom to the lower frequency corrections. Hence we give as little freedom as possible to the high frequency search directions. Heuristically, this is more natural because the fine grid corrections take care of oscillations (high frequency error) and leave the smooth part of error to the coarse grid corrections.

Mesh Dependent Reduction Factor

We proved linear convergence of SSC-CDM algorithms in Theorem 7.12. However, the reduction factor depends on the constants C_1 and C_2 . It is possible that the reduction factor goes quickly to 1 as we keep refining the mesh. For linear elliptic PDEs, multigrid and multilevel preconditioning techniques are usually used to construct algorithms with a reduction factor independent on the mesh-size. It is critical to prove the mesh independence of the reduction factor under subspace correction framework for uniformly refined meshes [146]. On the other hand, for adaptive meshes, uniform convergence is proved on newest vertex bisection grids in 2d by Chen and Wu [143] recently. Chen et al. [45] proved that a space decomposition is stable and optimal on graded bisection grids provided it is stable and optimal on quasi-uniform bisection grids.

Now we consider mesh dependence of the SSC-CDM method on bisection grids. We have presented a general convergence theory in Theorem 7.12. The convergence rate is globally linear but the reduction rate depends on the constants C_1 and C_2 in Assumption 7.9. The second assumption, the strengthened Cauchy-Schwarz inequality, depends solely upon the property of the space decomposition. We can show it is mesh independent using [132, §4.2.2] and [45, Theorem 5.2]. On the other hand, the estimation of C_1 is non-standard and problematic because we do not have the freedom to choose a ‘good’ decomposition. The decomposition is restricted due to the constraint $u_i \in \mathbb{K}_i$. We shall see that C_1 degenerates quickly in 3d and depends mildly on the smallest meshsize in 1d and 2d.

Lemma 7.22 (Estimation of C_1) For the multilevel decomposition defined in (7.14), we have the constant C_1 satisfies

$$C_1 \approx \begin{cases} |\ln(h_{\min})| & d = 1 \\ |\ln(h_{\min})|(1 + |\ln(h_{\min})|)^{\frac{1}{2}} & d = 2 \\ |\ln(h_{\min})|(h_{\min})^{-\frac{1}{2}} & d = 3. \end{cases}$$

Proof. Suppose $u = \sum_{i=0}^m u_i$ and $v = \sum_{i=0}^m v_i$. Recall that $u_i - v_i$ is supported on $\tilde{\omega}_i$. Using inverse estimation, we obtain that

$$\|u_i - v_i\|^2 \lesssim h_i^{-2} \|u_i - v_i\|_{L^2(\tilde{\omega}_i)}^2.$$

On the other hand, from Lemma 7.20, it is easy to see that

$$\|u_i - v_i\|_{L^2(\tau)}^2 \lesssim C_{d,\tau} h_\tau^2 \|u - v\|_{H^1(\omega_{i,\tau})}^2 \quad \forall \tau \in \tilde{\omega}_i.$$

We then regroup patches with respect to the generation of bisections and use shape-regularity of the bisection grids as well as the finite overlapping property of $\tilde{\omega}_j$ for same generation to get

$$\begin{aligned} \sum_{i=0}^m \|u_i - v_i\|^2 &= \sum_{l=0}^L \sum_{g_j=l} \|u_j - v_j\|^2 \lesssim \sum_{l=0}^L \bar{h}_l^{-2} \sum_{g_j=l} \|u_j - v_j\|_{L^2(\tilde{\omega}_j)}^2 \\ &\lesssim C_d \sum_{l=0}^L \|u - v\|_{H^1(\Omega)}^2 \lesssim C_d L \|u - v\|^2, \end{aligned}$$

where the constant C_d is

$$C_d := \begin{cases} C & d = 1 \\ C(1 + |\ln(h_{\min})|)^{\frac{1}{2}} & d = 2 \\ C(h_{\min})^{-\frac{1}{2}} & d = 3. \end{cases} \quad (7.16)$$

Since we are using bisection grids, $L \approx |\ln(h_{\min})|$, we obtain the final estimate.

■

Chapter 8

Numerical Experiments

In this chapter, we design numerical experiments to test various aspects of the a priori and a posteriori error estimations and the adaptive algorithm proposed in previous chapters. These include:

- A priori convergence rate (compare with Chapter 5);
- Asymptotic behavior of the error estimators (compare with Chapter 6);
- Reliability and efficiency of the error estimators (compare with Chapter 6);
- Localization property of the space error estimator (compare with Chapter 6);
- Approximation of the free boundary;
- Performance of the adaptive algorithm (compare with Chapter 7);
- Linear convergence rate of the discrete solver: SSC-CDM (compare with §7.5);
- Mesh dependence of the reduction rate for SSC-CDM (compare with §7.5);
- Application on American option pricing.

The goal of these numerical tests is to confirm the theories developed in previous chapters as well as provide more insight for future research.

The rest of this chapter is organized as follows. First we design benchmark test examples to test asymptotic convergence rates of the error and the error estimators in §8.1 (differential operators) and §8.2 (integral operators). Then we apply

the adaptive algorithm to solve the test problems and compare the performance of adaptive refinement strategy with the standard uniform refinement in §8.3. Finally, we examine the convergence behavior of the discrete solver (SSC-CDM) in §8.4.

The numerical experiments are done with adaptive finite element toolboxes ALBERTA of Schmidt and Siebert [123] or AFEM@matlab of Chen and Zhang [46]. Experiments are performed on a desktop PC with Pentium IV 2.4GHz and 1GB RAM.

We shall keep the notation as consistent as possible with the notation used in previous chapters. Here is a list of important quantities for quick reference:

- E : total error. For elliptic problems, it is the energy error; for parabolic problems, it is the L^2 -energy error.
- \mathcal{E} : total error estimator; see §6.7.4.
- \mathcal{E}/E : the *effectivity index* of error estimator \mathcal{E}
- N : number of time steps.
- DOF: number of degrees of freedom in space.
- EOC: experimental order of convergence (based on last two experiments).

8.1 Asymptotic convergence rates (Part I: Differential Problems)

The main purpose of the section is to design and perform test examples to confirm the theoretical results in Chapters 5 and 6.

8.1.1 1d Tent Obstacle: Case $\chi_h = \chi$

We take $\mathcal{A} := -\frac{\partial^2}{\partial x^2}$, the domain $\Omega := (-1.0, 1.0)$, the time interval $[0.5, 1.0]$, and the noncontact and contact sets to be $\mathcal{N} := \{|x| > t/6\}$ and $\mathcal{C} := \{|x| \leq t/6\}$.

If the obstacle is $\chi(x) = 1 - 3|x|$, then the exact solution u and forcing function f are given by

$$u(x, t) = \begin{cases} 36t^{-2}x^2 - (3 + 12t^{-1})|x| + 2 & \text{in } \mathcal{N} \\ 1 - 3|x| & \text{in } \mathcal{C}, \end{cases}$$

$$f(x, t) = \begin{cases} -12t^{-2}(6t^{-1}x^2 - |x| + 6) & \text{in } \mathcal{N} \\ -72t^{-2} & \text{in } \mathcal{C}. \end{cases}$$

Function u is depicted in Figure 8.1 at times $t = 0.5, 0.75$, and 1.0 .

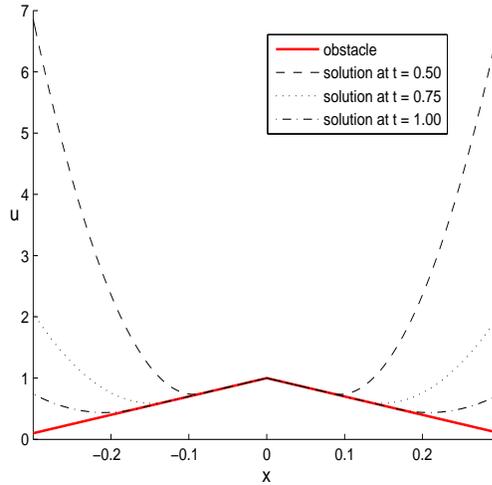


Figure 8.1: 1d Tent Obstacle: Exact solution $u(\cdot, t)$ at times $t = 0.5, 0.75, 1.0$. The obstacle χ is piecewise linear with a kink at $x = 0$, belonging to all partitions. This implies $\chi_h = \chi$.

To test the asymptotic convergence rates of both the proposed error estimator \mathcal{E} and exact error E , we halve time step k and space meshsize h in each experiment and report the results in Table 8.1 and Figure 8.2. To investigate the decay of each component $\mathcal{E}_{h,i}$ of the space estimator \mathcal{E}_h we fix the time-step to be $k = 2.5 \times 10^{-4}$, so small that the error is dominated by the space discretization. Table 8.2 displays their behavior under uniform space refinement: the estimator $\mathcal{E}_{h,1}$ exhibits optimal order 1 and dominates the other two terms.

We display in Figure 8.3 the nodal-based space error estimator $\Upsilon_h^n(z)$ at dif-

N	DOF	\mathcal{E}_h	\mathcal{E}_k	\mathcal{E}_{kh}	\mathcal{E}	E	Effectivity
64	127	2.256e+0	2.121e+0	2.731e-2	3.097e+0	7.347e-1	4.219
128	255	1.138e+0	1.059e+0	9.686e-3	1.555e+0	3.700e-1	4.202
256	511	5.716e-1	5.294e-1	3.338e-3	7.791e-1	1.857e-1	4.202
512	1023	2.864e-1	2.646e-1	1.181e-3	3.900e-1	9.301e-2	4.184
1024	2047	1.434e-1	1.323e-1	4.148e-4	1.951e-1	4.655e-2	4.184
EOC		0.998	1.000	1.510	0.999	0.999	–

Table 8.1: 1d Tent Obstacle Problem ($\chi_h = \chi$): The space and time estimators $\mathcal{E}_h, \mathcal{E}_k$, decrease with optimal order 1, but the mixed estimator \mathcal{E}_{kh} is of higher order. The ratio between total estimator \mathcal{E} and energy error E is quite stable and of moderate size.

1d tent obstacle example ($\chi_h = \chi$)				2d oscillating moving circle example			
DOF	$\mathcal{E}_{h,1}$	$\mathcal{E}_{h,2}$	$\mathcal{E}_{h,3}$	DOF	$\mathcal{E}_{h,1}$	$\mathcal{E}_{h,2}$	$\mathcal{E}_{h,3}$
129	2.282	3.034e-1	4.929e-2	145	1.094	1.323	1.194e-2
257	1.144	1.073e-1	1.823e-2	545	5.660e-1	4.974e-1	3.936e-3
513	5.729e-1	3.792e-2	6.295e-3	2113	2.880e-1	1.817e-1	1.368e-3
1025	2.866e-1	1.341e-2	2.250e-3	8321	1.453e-1	6.532e-2	4.652e-4
2049	1.434e-1	4.740e-3	7.903e-4	33025	7.295e-2	2.329e-2	1.617e-4
EOC				EOC			
0.999				0.994			
1.500				1.488			
1.509				1.525			

Table 8.2: Decay of each component $\mathcal{E}_{h,i}$ of the space estimator \mathcal{E}_h for a fixed time-step $k = 2.5 \times 10^{-4}$ so small that the time estimator \mathcal{E}_k is insignificant. Left: 1d tent obstacle problem 8.1.1; Right: 2d oscillating moving obstacle problem 8.3.3. In both cases the nodal-based estimator $\mathcal{E}_{h,1}$ exhibits the expected order 1 whereas the other two superconverge.

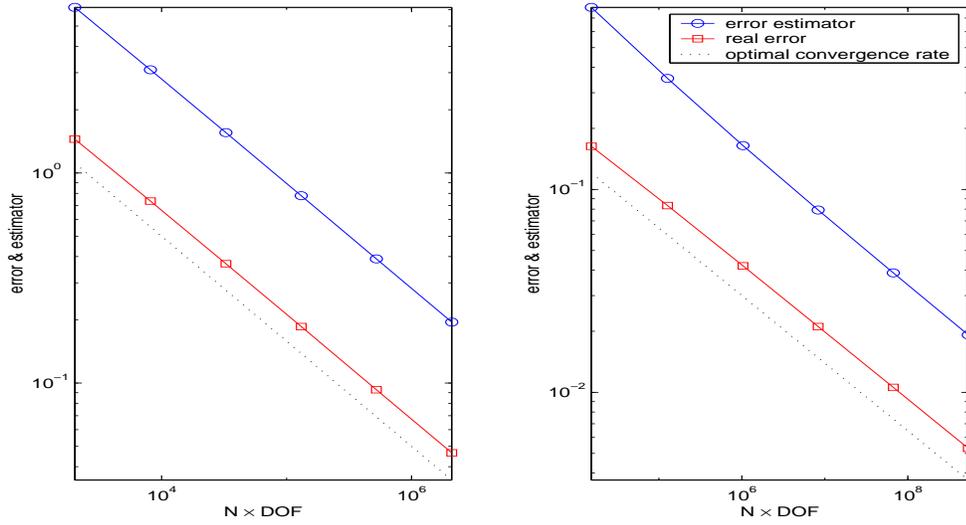


Figure 8.2: Error estimator \mathcal{E} and energy error E vs. total number of degrees of freedom ($N \cdot \text{DOF}$) for 1d Tent Obstacle Example 8.1.1 with $\chi_h = \chi$ (left) and 2d Oscillating Moving Circle Problem 8.3.3 (right). Since $N \cdot \text{DOF} \simeq \frac{1}{kh^d} \simeq \frac{1}{h^{d+1}}$, provided $k \simeq h$, the optimal error decay is $\mathcal{O}(h) = \mathcal{O}((N \cdot \text{DOF})^{-\frac{1}{d+1}})$ and is indicated by the dotted lines with slopes $-1/2$ (left) for $d = 1$ and $-1/3$ (right) for $d = 2$. This shows optimal decay of both \mathcal{E} and E .

ferent stages $t_n = 0.6, 0.8, 1.0$ of the evolution. We see that $\Upsilon_h^n(z)$ vanishes at full-contact nodes $z \in \mathcal{C}_h^n$, as predicted by theory, and that the exact free-boundary is captured within one element. This is further documented in Table 8.3 which shows exact and approximate free boundary locations at times $t_n = 0.6, 0.8, 1.0$.

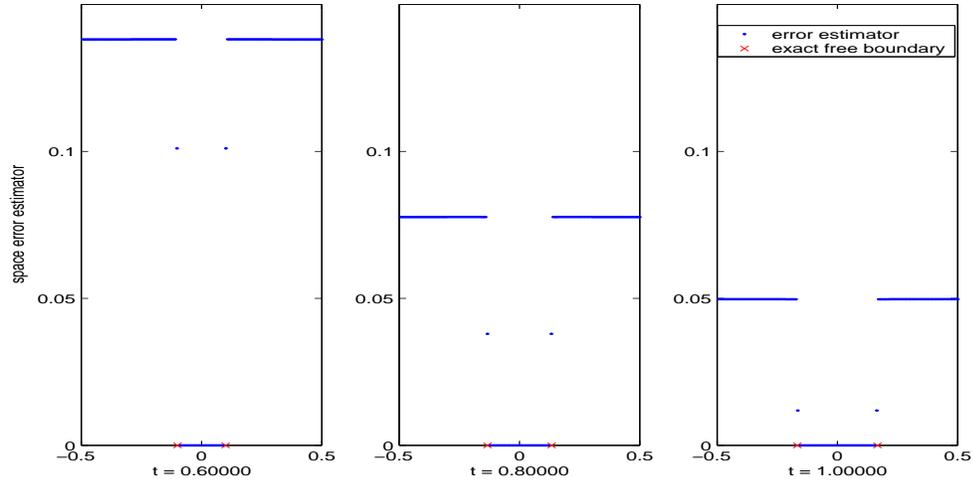


Figure 8.3: 1d Tent Obstacle Problem: Nodal-based space error estimator $\Upsilon_h^n(z)$ at times $t_n = 0.6, 0.8, 1.0$ for $\text{DOF} = 255$ and $k = 2.5 \times 10^{-4}$. The localization property that $\Upsilon_h^n(z)$ vanishes at the full-contact nodes $z \in \mathcal{C}_h^n$ is clearly visible, along with the fact that free-boundary approximation takes place within one element (see Table 8.3).

Time	Exact Free Boundary	Approx Free Boundary
0.6	$\pm 1.0000 \times 10^{-1}$	$\pm 1.0156 \times 10^{-1}$
0.8	$\pm 1.3333 \times 10^{-1}$	$\pm 1.3328 \times 10^{-1}$
1.0	$\pm 1.6667 \times 10^{-1}$	$\pm 1.6406 \times 10^{-1}$

Table 8.3: 1d Tent Obstacle Problem ($\chi_h = \chi$): Since the meshsize is $h \approx 7.8 \times 10^{-3}$ the FEM captures the exact free boundary within one element.

8.1.2 1d Tent Obstacle: Case $\chi_h \neq \chi$

In general, we cannot expect the underlying mesh to match the singular behavior of the obstacle, as in Example 8.1.1, even for piecewise linear obstacles. This happens, for instance, when the obstacles change in time. The question thus arises whether or not the proposed error estimator \mathcal{E} is able to capture the correct behavior of the solution when a singularity is not resolved by the mesh.

To answer this question, we modify Example 8.3.2 by the shift $v(x - \frac{1}{3}, t)$ for $v = u, \chi, f$ but keep the same meshes and time steps as before. In this case, the kink at $x = 1/3$ is never a mesh point and $\chi_h \neq \chi$. Since χ is almost in $\tilde{H}^{3/2}$ we expect a rate of convergence 0.5 in \tilde{H}^1 . This is confirmed by the results of Table 8.4, which also shows that the only estimator that detects this reduced order is \mathcal{E}_χ , the obstacle consistency error estimator. We observe that \mathcal{E}_h and \mathcal{E}_k dominate at the beginning and it takes quite awhile to reach the asymptotic regime.

N	DOF	\mathcal{E}_h	\mathcal{E}_k	\mathcal{E}_{kh}	\mathcal{E}_χ	\mathcal{E}	E	Effectivity
1024	2047	1.434e-1	1.548e-1	4.154e-4	9.882e-2	2.330e-1	8.175e-2	2.850
2048	4095	7.172e-2	7.741e-2	1.466e-4	6.988e-2	1.266e-1	5.050e-2	2.507
4096	8191	3.587e-2	3.871e-2	5.181e-5	4.941e-2	7.229e-2	3.282e-2	2.203
8192	16383	1.794e-2	1.935e-2	1.831e-5	3.494e-2	4.378e-2	2.213e-2	1.978
16384	32767	8.970e-3	9.676e-3	6.471e-6	2.471e-2	2.801e-2	1.527e-2	1.834
EOC		1.000	1.000	1.501	0.500	0.644	0.535	–

Table 8.4: 1d Tent Obstacle Problem ($\chi_h \neq \chi$): the kink is not resolved by the underlying meshes with uniform mesh refinement. The only estimator that detects the reduced order 0.5 is \mathcal{E}_χ . The total estimator is dominated by \mathcal{E}_h and \mathcal{E}_k at the beginning but eventually \mathcal{E}_χ takes over. This combined effect is reflected in the behavior of the total estimator \mathcal{E} .

We wonder whether making a suitable local mesh refinement near the kink may restore the optimal linear rate. We conduct an experiment consisting of locally refined meshes only at the kink location, where the meshsize is h^2 , whereas it remains uniform and equal to h elsewhere. The interpolation error in \tilde{H}^1 becomes now

proportional to h , both at the kink location and elsewhere, because the error in W_∞^1 is $\mathcal{O}(1)$ and $\mathcal{O}(h)$, respectively. This heuristic argument is corroborated by the results of Table 8.5, which illustrates the potentials of mesh refinement to achieve optimal complexity along with the importance of \mathcal{E}_χ .

N	DOF	\mathcal{E}_h	\mathcal{E}_k	\mathcal{E}_{kh}	\mathcal{E}_χ	\mathcal{E}	E	Effectivity
32	45	8.710	4.915	1.787e-1	1.398e-1	1.000e+1	2.802	3.570
64	93	4.477	2.467	5.648e-2	9.882e-2	5.113	1.431	3.573
128	191	2.267	1.236	1.947e-2	3.494e-2	2.583	7.230e-1	3.572
256	382	1.141	6.186e-1	6.631e-3	1.747e-2	1.298	3.634e-1	3.572
512	767	5.723e-1	3.095e-1	2.326e-3	8.735e-3	6.507e-1	1.822e-1	3.572
EOC		0.995	0.999	1.511	1.000	0.996	0.996	–

Table 8.5: 1d Tent Obstacle Problem ($\chi_h \neq \chi$): The underlying partition is locally refined at the kink location, where the meshsize is h^2 , but is otherwise uniform with meshsize h . This restores the optimal linear rate for both \mathcal{E}_χ and E , as well as the total estimator \mathcal{E} (compared with the reduced rate reported in Table 8.4 for uniform meshes).

8.1.3 1d American Option

In American option pricing problems, we start from an initial condition u_0 , as in (8.4), which is in the Sobolev space $\tilde{H}^{\frac{3}{2}-\epsilon}$ for any $\epsilon > 0$ but not in any smoother regularity class. The a priori error estimates in Chapter 5 imply a rate of convergence $\mathcal{O}(k^{1/2})$ for $u_0 \in \tilde{H}^1$ and $\mathcal{O}(k)$ for $u_0 \in \tilde{H}^2$. Given the fractional regularity right in the middle between \tilde{H}^1 and \tilde{H}^2 , we expect, from interpolation theory, that the convergence rate with uniform time-step would be about $\mathcal{O}(k^{3/4})$. Our experiments confirm this expectation.

We take an American put option problem on a single stock with strike price $K = 100$, maturity time $T = 0.5$ year, volatility $\sigma = 0.4$, interest rate $r = 6\%$, and forcing $f = 0$. We choose the space domain to be $\Omega = (-1, 7)$. Table 8.6 displays all four estimators and \mathcal{E}_k has indeed the expected rate of about 0.75.

N	DOF	\mathcal{E}_h	\mathcal{E}_k	\mathcal{E}_{kh}	\mathcal{E}_χ	\mathcal{E}
128	511	4.353e-2	1.149e-1	3.240e-3	3.843e-1	4.035e-1
256	1023	2.172e-2	7.023e-2	1.147e-3	2.434e-1	2.543e-1
512	2047	1.091e-2	5.026e-2	4.035e-4	1.195e-1	1.301e-1
1024	4095	5.461e-3	2.940e-2	1.416e-4	7.581e-2	8.150e-2
2048	8191	2.736e-3	1.751e-2	4.980e-5	4.931e-2	5.240e-2
EOC		0.997	0.748	1.505	0.620	0.637

Table 8.6: 1d American Put Option Problem: Uniform time and space partitions yield suboptimal rates for \mathcal{E}_k and \mathcal{E}_χ due to the fractional regularity of the initial condition, which is about $\tilde{H}^{3/2}$. This explains the order of about 0.75 of \mathcal{E}_k , that accounts for the initial transient regime, but not quite the suboptimal order of \mathcal{E}_χ .

We now explore the effect of refining the time partition to restore the optimal convergence rate. We design an algebraically graded time grid

$$t_n = \left(\frac{n}{N}\right)^\beta \quad \forall 1 \leq n \leq N,$$

with $\beta > 0$ to be determined so that the time error estimator $\mathcal{E}_k \approx \mathcal{O}(N^{-1})$. The time-step k_n reads

$$k_n = \left(\frac{n}{N}\right)^\beta - \left(\frac{n-1}{N}\right)^\beta \approx \frac{\beta}{N} \left(\frac{n}{N}\right)^{\beta-1} \quad \Rightarrow \quad k_n \approx \frac{\beta}{N} t_n^{1-1/\beta}.$$

We recall the regularizing effect for linear parabolic problems, namely,

$$\|\partial_t u(\cdot, t)\|_{H^1} \approx \|u(\cdot, t)\|_{H^3} \lesssim t^{-3/4}$$

provided the initial condition $u_0 \in \tilde{H}^{3/2}$. We proceed heuristically and assume the same asymptotic behavior to be valid for parabolic variational inequalities. We next formally replace $\|U_h^n - U_h^{n-1}\| \lesssim \int_{t_{n-1}}^{t_n} \|\partial_t u(\cdot, t)\| dt$ in the definition of \mathcal{E}_k to get

$$\mathcal{E}_k^2 \approx \sum_{n=1}^N \int_{t_{n-1}}^{t_n} \|\partial_t u(\cdot, t)\|^2 k_n^2 dt \approx \frac{\beta}{N} \int_0^T t^{-3/2+2(1-1/\beta)} dt \approx \mathcal{O}(N^{-1}),$$

provided $\beta > 4/3$. This argument can be made rigorous for linear parabolic equations upon using Theorem 4.5 of [137] and carefully approximating the solution on

the first time interval. To test this heuristic argument for parabolic variational inequalities, we take $\beta = 1.5$ and report the results in Table 8.7. We see that this properly chosen time partition restores the optimal convergence rate not only for \mathcal{E}_k but also for \mathcal{E}_χ . Moreover, this argument explains why uniform time stepping, i.e. $\beta = 1$, yields a suboptimal convergence rate for the time estimator \mathcal{E}_k (see Table 8.6).

N	DOF	\mathcal{E}_h	\mathcal{E}_k	\mathcal{E}_{kh}	\mathcal{E}_χ	\mathcal{E}
80	1023	2.386e-2	8.152e-2	1.945e-3	1.833e-1	2.021e-1
160	2047	1.159e-2	4.397e-2	6.693e-4	8.679e-2	9.798e-2
320	4095	5.657e-3	2.235e-2	2.313e-4	4.385e-2	4.954e-2
640	8191	2.793e-3	1.137e-2	8.030e-5	2.238e-2	2.526e-2
1280	16383	1.388e-3	5.899e-3	2.787e-5	1.162e-2	1.310e-2
EOC		1.018	0.975	1.526	0.970	0.972

Table 8.7: 1d American Put Option Problem: Algebraically graded time partition $t_n = \left(\frac{n}{N}\right)^{3/2}$ and uniform space mesh. This grading restores the optimal linear convergence rate of both \mathcal{E}_k and \mathcal{E}_χ (compared with Table 8.6).

8.1.4 1d American Option with Moving Obstacle

To test the asymptotic behavior of the obstacle consistency term $\|\partial_t(\chi - U_h)^+\|_*$, which we omitted in \mathcal{E} , we modify the previous American option problem in the following way: from time $t = 0$ to 0.5, we still have the same American option pricing problem as in §8.3.1. From time $t = 0.5$ to 1.0, we raise the obstacle at a constant rate $\xi \in \mathbb{R}_+$. In other words, the obstacle in the previous example has been replaced by:

$$\chi(x, t) := (K - e^x)^+ [1 + \xi(t - 0.5)^+] \quad x \in (-1, 7), t \in [0, T].$$

In this way, we exclude the initial transient region from our consideration and the singular point $\log(K)$ is always a mesh point. Also we choose the speed ξ moderate to prevent the free boundary point to recede to $\log(K)$. As in the

analysis for the last example, the uniform space mesh and algebraically graded time partition should give optimal convergence rate. We report in Table 8.8 the mixed error estimator terms \mathcal{E}_{kh} , \mathcal{E}_χ and

$$\mathcal{E}_\chi^o := \left(\int_0^T \|\partial_t(\chi - U_h)^+\|_{L^2(\Omega)}^2 dt \right)^{\frac{1}{2}}.$$

Since it is difficult to compute the dual norm in the term

$$\mathcal{E}_\chi^* := \left(\int_0^T \|\|\partial_t(\chi - U_h)^+\|\|_*^2 dt \right)^{\frac{1}{2}},$$

we compute the term \mathcal{E}_χ^o with the L^2 -norm instead.

From Table 8.8 we see that the experimental convergence rate of \mathcal{E}_χ^o is 1.0. Since $\mathcal{E}_\chi^* \lesssim \mathcal{E}_\chi^o$, the numerical results show evidence that \mathcal{E}_χ^* is of at least the same order as the obstacle consistency term \mathcal{E}_χ and this justifies the comments in Remark 7.3. On the other hand, we see that the convergence rate of the mixed error term \mathcal{E}_{kh} is greater than 1.0 and becomes closer to 1.0 as the obstacle moves faster and faster. Notice that we omit in Table 8.8 the space error estimator \mathcal{E}_h and time error estimator \mathcal{E}_k , which also converge at the optimal rate 1.0.

ξ		0.01			0.1			1.0		
N	DOF	\mathcal{E}_{kh}	\mathcal{E}_χ	\mathcal{E}_χ^o	\mathcal{E}_{kh}	\mathcal{E}_χ	\mathcal{E}_χ^o	\mathcal{E}_{kh}	\mathcal{E}_χ	\mathcal{E}_χ^o
40	511	1.024e-2	3.116e-1	1.250e-1	1.128e-2	3.321e-1	1.250e-1	8.120e-2	5.144e-1	1.280e-1
80	1023	3.623e-3	1.559e-1	6.311e-2	4.015e-3	1.660e-1	6.310e-2	3.613e-3	2.565e-1	6.421e-2
160	2047	1.255e-3	7.801e-2	3.157e-2	1.523e-3	8.304e-2	3.157e-2	1.895e-2	1.280e-1	3.231e-2
320	4095	4.472e-4	3.902e-2	1.584e-2	5.554e-4	4.153e-2	1.584e-2	8.698e-3	6.398e-2	1.611e-2
EOC		1.489	0.999	0.995	1.455	1.000	0.995	1.123	1.000	1.005

Table 8.8: Modified American Put Option Problem: Algebraically graded time partition $t_n = \left(\frac{n}{N}\right)^{3/2}$ and uniform space mesh.

8.1.5 2d Oscillating Moving Circle

Let the operator be $\mathcal{A} := -\Delta$, the domain be $\Omega = (-1, 1)^2$, the time interval be $[0, 0.25]$, and the noncontact and contact sets be $\mathcal{N} := \{|x - c(t)|_2 > r_0(t)\}$ and

$\mathcal{C} := \{|x - c(t)|_2 \leq r_0(t)\}$ with

$$r_0(t) = 1/3 + 0.3 \sin(4\omega\pi t), \quad c(t) = r_1 (\cos(\omega\pi t), \sin(\omega\pi t))^T,$$

and $r_1 = 1/3, \omega = 4.0$. The obstacle is $\chi \equiv 0$, and the exact solution u and forcing function f are

$$u(x, t) = \begin{cases} \frac{1}{2} (|x - c(t)|_2^2 - r_0(t)^2)^2 & \text{in } \mathcal{N} \\ 0 & \text{in } \mathcal{C}, \end{cases}$$

$$f(x, t) = \begin{cases} 4(r_0^2(t) - 2|x - c(t)|_2^2 - \frac{1}{2}(|x - c(t)|_2^2 - r_0^2(t))((x - c(t)) \cdot c'(t) + r_0(t)r_0'(t))) & \text{in } \mathcal{N} \\ -4r_0^2(t)(1 - |x - c(t)|_2^2 + r_0^2(t)) & \text{in } \mathcal{C}. \end{cases}$$

The free boundary is an oscillating circle with radius $r_0(t)$ and center $c(t)$ moving counterclockwise along the circle of radius r_1 centered at the origin. The initial and boundary conditions are given by u .

We halve both time-step k and space meshsize h in each experiment and report the results in Table 8.9 and Figure 8.2; we observe optimal linear convergence rate. We also investigate in Table 8.2 the decay of the space estimators alone. We fix the time-step $k = 2.5 \times 10^{-4}$ and halve the meshsize size in each experiment. We observe optimal linear decay of $\mathcal{E}_{h,1}$ but higher order of convergence for $\mathcal{E}_{h,2}, \mathcal{E}_{h,3}$.

N	DOF	\mathcal{E}_h	\mathcal{E}_k	\mathcal{E}_{kh}	\mathcal{E}	E	Effectivity
64	1985	3.432e-1	8.110e-2	2.219e-3	3.527e-1	8.328e-2	4.237
128	8065	1.597e-1	4.008e-2	8.087e-4	1.646e-1	4.204e-2	3.922
256	32513	7.664e-2	1.996e-2	2.899e-4	7.920e-1	2.111e-2	3.745
512	130561	3.749e-2	9.965e-3	1.037e-4	3.879e-2	1.058e-2	3.663
1024	523265	1.853e-2	4.980e-3	3.691e-5	1.919e-2	5.297e-3	3.623
EOC		1.017	1.001	1.490	1.015	0.998	–

Table 8.9: 2d Oscillating Moving Circle Problem: The space and time estimators $\mathcal{E}_h, \mathcal{E}_k$, decrease with optimal order 1, but the mixed estimator \mathcal{E}_{kh} is of higher order. The effectivity index, the ratio between total estimator \mathcal{E} and energy error E , is quite stable and of moderate size.

In Figure 8.4, we show the nodal-based space error indicator $\Upsilon_h^n(z)$ on the cross section $x_2 = 0$ at different stages of the evolution $t_n = 0.02, 0.05, 0.18$. For

the same times and cross section, we also compare the exact and approximate free boundaries in Table 8.10. Their difference is well within one meshsize.

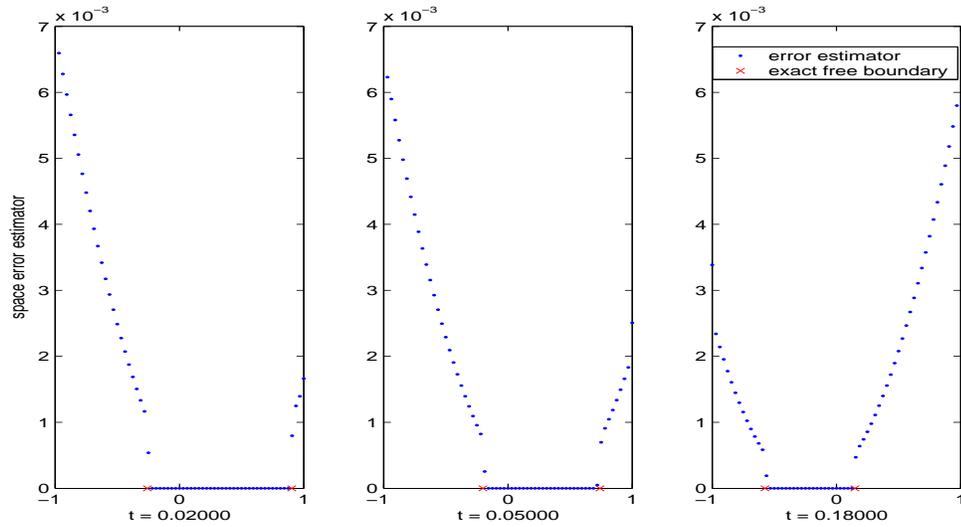


Figure 8.4: 2d Oscillating Moving Circle Problem: nodal-based error estimator $\Upsilon_h^n(z)$ in the cross section $x_2 = 0$ for $\text{DOF} = 8065, k = 2.5 \times 10^{-4}$ and $t_n = 0.02, 0.05, 0.18$. Note the vanishing of $\Upsilon_h^n(z)$ for full-contact nodes and the monotone behavior for the rest.

Time	Exact Free Boundaries	Approx Free Boundaries
0.02	$\{-2.5788 \times 10^{-1}, 9.0361 \times 10^{-1}\}$	$\{-2.5000 \times 10^{-1}, 9.0625 \times 10^{-1}\}$
0.05	$\{-2.0083 \times 10^{-1}, 7.4017 \times 10^{-1}\}$	$\{-1.8750 \times 10^{-1}, 7.1875 \times 10^{-1}\}$
0.18	$\{-5.7430 \times 10^{-1}, 1.4942 \times 10^{-1}\}$	$\{-5.6250 \times 10^{-1}, 1.5625 \times 10^{-1}\}$

Table 8.10: 2d Oscillating Moving Circle: Exact and approximate free boundaries on the cross section $x_2 = 0$. Their differences are less than one meshsize, which is about 2.2×10^{-2} .

8.2 Asymptotic convergence rates (Part II: Integral Problems)

Till now, we have not done any test on problems with an integral operator. In this part, we test the behavior of the local error estimators on elliptic and parabolic equations and inequalities with an integral operator. As an example, we employ an hyper-singular elliptic operator which mimics the behavior of the integral operator in the CGMY model in 1d: $\Omega = (a, b)$, $\mathcal{A}_I : \tilde{H}^{Y/2}(\Omega) \rightarrow H^{-Y/2}(\Omega)$

$$\mathcal{A}_I u(x) := \int_{\Omega} k(x-y)u(y) dy \quad \text{and} \quad k(x) := \frac{1}{|x|^{1+Y}}. \quad (8.1)$$

Refer to §3.15 for the meaning of this singular integral. We take $p = 2$ in (6.35) and let

$$\mathcal{E}_{h,1}^2 := \sum_{z \in \mathcal{P}_h \setminus \mathcal{C}_h} \eta_z^2 \quad \text{and} \quad \mathcal{E}_{h,2}^2 := \sum_{z \in \mathcal{P}_h \setminus \mathcal{C}_h} \xi_z^2.$$

Remark 8.1 (Quadrature for Singular Integration) Let $a = x_0 < x_1 < \dots < x_N = b$ be the mesh points of $\Omega = (a, b)$. Since the residual r_h is singular at the ends of each interval, we subdivide $[x_{i-1}, x_i]$ of length h_i into the following way: Let $P > 0$ be an integer and $\rho = 0.1$. We introduce additional points at distance $\rho^j h_i$ from the left and right endpoints, for $j = 1, \dots, P$. This divides the interval in $1 + 2P$ subintervals. On each of these intervals, Q -point Gauss-Legendre rule is applied for numerical integration. Also the condition $r \leq 0$ in the definition of \mathcal{C}_h is checked pointwise at each of the $(1 + 2P)Q$ quadrature points. It is known that the quadrature error decrease exponentially fast with respect to PQ (see [124]). In all our numerical tests, $P = 1$ and $Q = 2$.

8.2.1 Elliptic Equations

In this example, we consider problem (1.10). Let $\Omega = (-1, 1)$ and $Y = 1$. It is easy to see that if the solution $u > \chi$, then the variational inequality becomes a variational equation. To test the asymptotic behavior of the error estimators, we choose $\chi = -\infty$ and construct a problem with exact solution available.

Pure Integral Operator Case

Take $\mathcal{A} = \mathcal{A}_I$ and $f(x) = \frac{15}{8} - \frac{15}{2}x^2 + 5x^4$. The exact solution for this problem is $u = \frac{1}{\pi}(1 - x^2)^{5/2}$. The exact solution u is smooth and therefore the convergence rate in the energy norm $\|u - u_h\|$ (in this case, the energy norm is equivalent to $\tilde{H}^{1/2}(\Omega)$ -norm) is expected to be $\text{DOF}^{-1.5}$ for uniform mesh. The numerical test (see Table 8.11) shows that both energy error and error estimator $\mathcal{E}_{h,2}$ converge at the optimal rate; note that $\mathcal{E}_{h,1} = 0$ and $\mathcal{E}_{h,3} = 0$. Furthermore, the effectivity index of \mathcal{E} is almost a constant (around 2.5).

DOF	$\ u - u_h\ $	$\mathcal{E} = \mathcal{E}_{h,2}$	Effectivity
7	4.0418e-002	1.7125e-001	4.2370
15	1.3021e-002	6.2052e-002	4.7655
31	4.4597e-003	2.2014e-002	4.9362
63	1.5618e-003	7.7849e-003	4.9846
127	5.5069e-004	2.7527e-003	4.9986
255	1.9455e-004	9.7327e-004	5.0027
EOC	1.501	1.500	–

Table 8.11: Elliptic equation with pure integral operator $\mathcal{A} = \mathcal{A}_I$ (uniform mesh, expected convergence rate 1.5). EOC is the experimental convergence rate based on last two iterations, which agrees with the expected value 1.5.

In Remark 6.23, we have discussed that the oscillation term behaves differently in the integro-differential equations than in the usual elliptic equations. The choice of \mathbb{P}_z is important. In particular, the usual choice of \mathbb{P}_z being the space of constant, does not help. The next simplest choice of \mathbb{P}_z is piecewise linear functions on ω_z . On the other hand, we would like to have a meaningful lower bound. To this end, we want to have a relatively small oscillation term with respect to the error estimator. We have seen, in the differential case, that the oscillation terms are of higher order in §8.1. Hence, in that case, the oscillation term is negligible asymptotically. In contrast, for problems with integral operators, the singularities of the residual on each element do not go away as the elements are refined. We thus have the oscillation

term of the same order as the error estimator asymptotically. Fortunately, if we enrich the finite dimensional space \mathbb{P}_z , we could make the oscillation term smaller and smaller. For example, we could choose \mathbb{P}_z to be piecewise linear functions and denote the corresponding oscillation term by osc_1 ; we can also add singular functions such as $\log(|x-z|)$ to the basis of \mathbb{P}_z to obtain a smaller oscillation osc_2 ; note that for $Y = 1$ the singularities of the residual r_h are logarithmic. We report both oscillation terms in Table 8.12.

DOF	\mathcal{E}	osc_1	osc_2
7	1.7125e-01	1.6803e-01	1.4660e-02
15	6.2052e-02	6.1627e-02	3.4651e-03
31	2.2014e-02	2.1953e-02	1.0112e-03
63	7.7849e-03	7.7786e-03	3.3479e-04
127	2.7527e-03	2.7521e-03	1.1610e-04
255	9.7327e-04	9.7322e-04	4.0828e-05
EOC	1.500	1.500	1.508

Table 8.12: Elliptic integral equation: asymptotic convergence rates of the oscillation term with $\mathcal{A} = \mathcal{A}_I$ and uniform meshes. Even though the asymptotic decay of \mathcal{E} and osc is the same, adding singular functions mimicking the residual behavior may reduce osc by an order of magnitude (compare osc_1 and osc_2).

Although, we can only prove the global efficiency of the proposed error estimator \mathcal{E} , we notice that ξ_z also captures the local behavior of the pointwise error based on comparison of the nodal-based error indicator and the pointwise error in Figure 8.5. This observation justifies in some sense why the proposed error estimator should work well for driving adaptive algorithms.

In the above problem, we constructed the exact solution for $Y = 1$. For $Y \neq 1$, we can still check the asymptotic convergence rate of the error estimator and we report the convergence rate of \mathcal{E} for different Y in Table 8.13. From the approximation theory standpoint, we would expect the convergence rates to be $2 - \frac{Y}{2}$ and the numerical experiments corroborate this theoretical expectation.

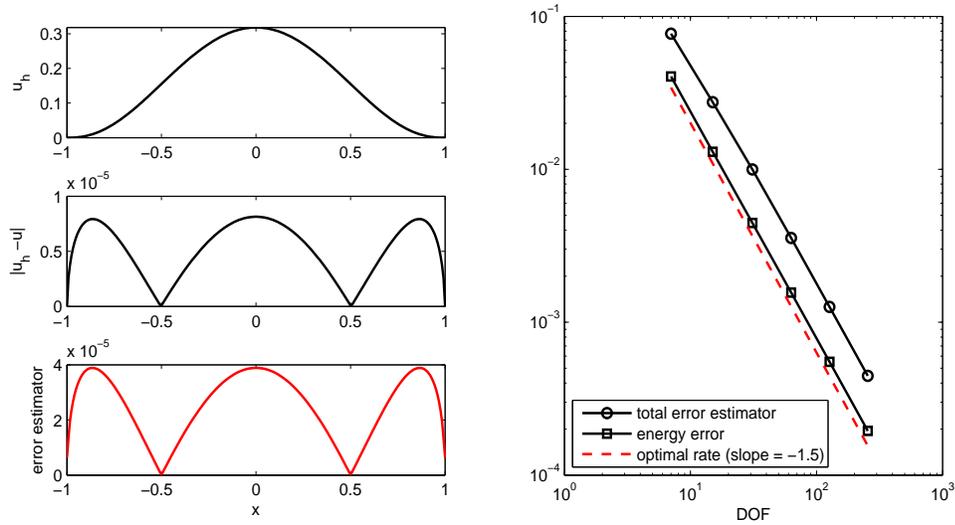


Figure 8.5: Elliptic equation with integral operator $\mathcal{A} = \mathcal{A}_I$ (uniform mesh): upper left, solution; middle left: pointwise error; lower left, nodal-based error estimator; right, asymptotic convergence rates for energy error and estimator $\mathcal{E}_{h,2}$.

$Y = 0.5$		$Y = 1.5$		$Y = 1.9$	
DOF	\mathcal{E}	DOF	\mathcal{E}	DOF	\mathcal{E}
7	1.7789e-01	7	2.7871e-01	7	2.8523e-01
15	6.7131e-02	15	1.4110e-01	15	1.1953e-01
31	1.9521e-02	31	6.4127e-02	31	5.2421e-02
63	5.8037e-03	63	2.8051e-02	63	2.4796e-02
127	1.7413e-03	127	1.1923e-02	127	1.1907e-02
255	5.3718e-04	255	5.0268e-03	255	5.7361e-03
EOC	1.737	EOC	1.246	EOC	1.046

Table 8.13: Elliptic equation with pure integral operator $\mathcal{A} = \mathcal{A}_I$ (uniform mesh). The expected convergence rate, for smooth solutions, is $2 - \frac{Y}{2}$ and is corroborated by the experiments.

Integro-differential Operator Case

In this case, we take $\mathcal{A} = -\Delta + \mathcal{A}_I$ and choose an appropriate right-hand side function f so that the exact solution is exactly the same as in the previous example. The energy error (equivalent to $\tilde{H}^1(\Omega)$ -norm) as well as the error estimators are reported in Table 8.14. In this case, $\rho = 2$ and $Y = 1$. We see that the jump residual term η converges at the optimal convergence rate ($\text{DOF}^{-1.0}$) just as the energy error itself. On the other hand, $\mathcal{E}_{h,2}$ is of higher order as we expected (see Remark 6.27). As in the last example, Figure 8.6 shows the nodal-based error indicator captures the local behavior of the pointwise error.

DOF	$\ u - u_h\ $	$\mathcal{E}_{h,1}$	$\mathcal{E}_{h,2}$	\mathcal{E}	Effectivity
7	1.2483e-001	3.7080e-001	3.7090e-002	3.7265e-001	2.9853
15	5.9891e-002	1.9670e-001	9.6446e-003	1.9694e-001	3.2883
31	2.9484e-002	1.0010e-001	2.4849e-003	1.0013e-001	3.3962
63	1.4647e-002	5.0323e-002	6.2787e-004	5.0327e-002	3.4361
127	7.3015e-003	2.5203e-002	1.5751e-004	2.5204e-002	3.4519
255	3.6455e-003	1.2608e-002	3.9419e-005	1.2608e-002	3.4585
EOC	1.002	0.999	1.998	0.999	–

Table 8.14: Elliptic equation with integro-differential operator $\mathcal{A} = -\Delta + \mathcal{A}_I$ (uniform mesh, expected convergence rate 1.0). The experimental convergence rate EOC, based on last two iterations, agrees with the expected value 1.0.

8.2.2 Elliptic Variational Inequalities

Take $\mathcal{A} = \mathcal{A}_I$ with $Y = 0.2$ and consider the problem (1.10) with $f = 0$ and the obstacle

$$\chi(x) = \max(0.5 - |x|, 0).$$

There are singularities at both the end points and the free boundary points. To resolve the singularities at the boundary points, we can employ algebraically graded. This still gives a suboptimal convergence rate (see Figure 8.7) due to the singular-

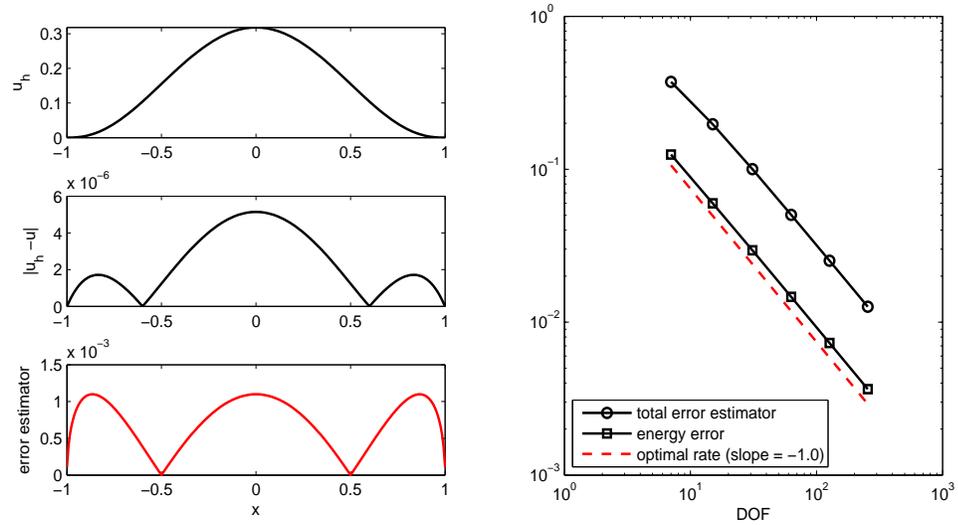


Figure 8.6: Elliptic equation with integro-differential operator $\mathcal{A} = -\Delta + \mathcal{A}_I$ (uniform mesh): upper left, solution; middle left: pointwise error; lower left, nodal-based error estimator; right, asymptotic convergence rates for energy error and estimator $\mathcal{E}_{h,2}$.

ities at the free boundary points (see regularity result [128]). Using the classical adaptive algorithm of §7.1.1 (see also [123]) driven by the local error indicator ξ_z , we restore the optimal convergence rate (see Figure 8.8). From Figure 8.8, we can see that the adaptive algorithm automatically generated locally refined mesh near the singularities (both the end points and the free boundary).

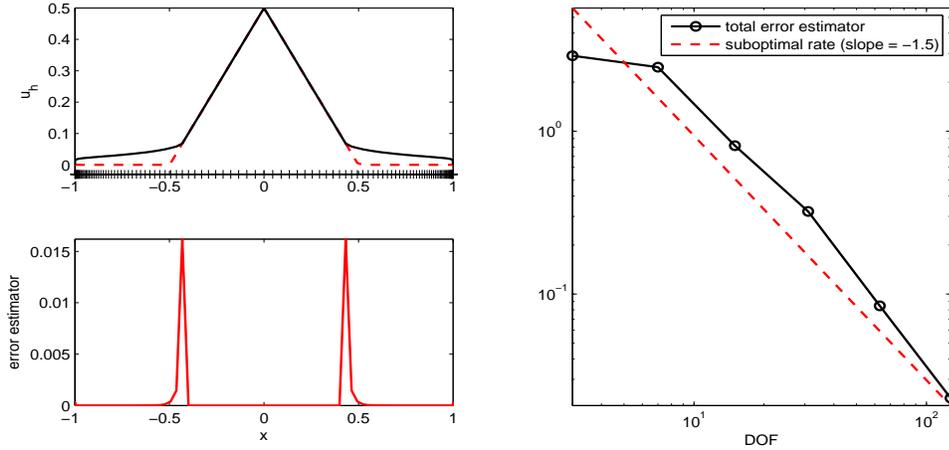


Figure 8.7: Elliptic Variational Inequality (algebraically graded mesh towards end points): upper left, solution (black solid) and obstacle (red dashed); lower left, nodal-based error estimator in logarithmic scale $\log(\xi_z)$; right, convergence rate.

8.2.3 Parabolic Variational Inequalities

In this example, we examine the time-dependent problem (1.18). To mimic the butterfly American-style option, we take $\mathcal{A} = \mathcal{A}_I$, $f = 0$, $\chi(x) = \max(\frac{1}{2} - |x|, 0)$, and $u_0 = \chi$. The solution as well as space error estimator at $t = 0.0625$ and $t = 0.5$ are shown in Figure 8.9. In [104], a heuristic argument has been given for the suboptimal convergence rate for the energy error if uniform time-steps are employed. This does not apply now because we have a weaker energy norm and the initial singularity is not strong enough to be seen. To resolve the singularities at both ends, we use algebraically graded meshes toward the end points and uniform time partition. The convergence rates for both time and space error estimators are optimal and have

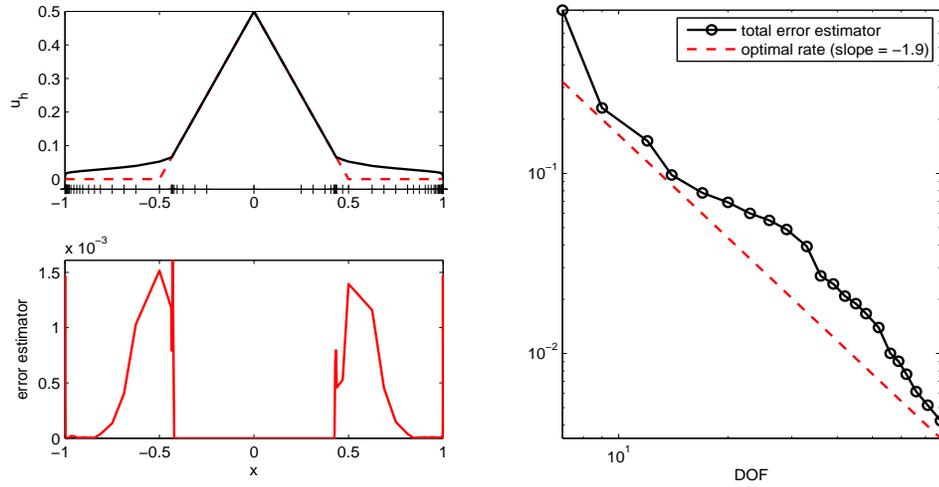


Figure 8.8: Elliptic Variational Inequality (Adaptive Method): upper left, solution (black solid), obstacle (red dashed), and associated mesh points; lower left, nodal-based error estimator in logarithmic scale $\log(\xi_z)$; right, convergence rate.

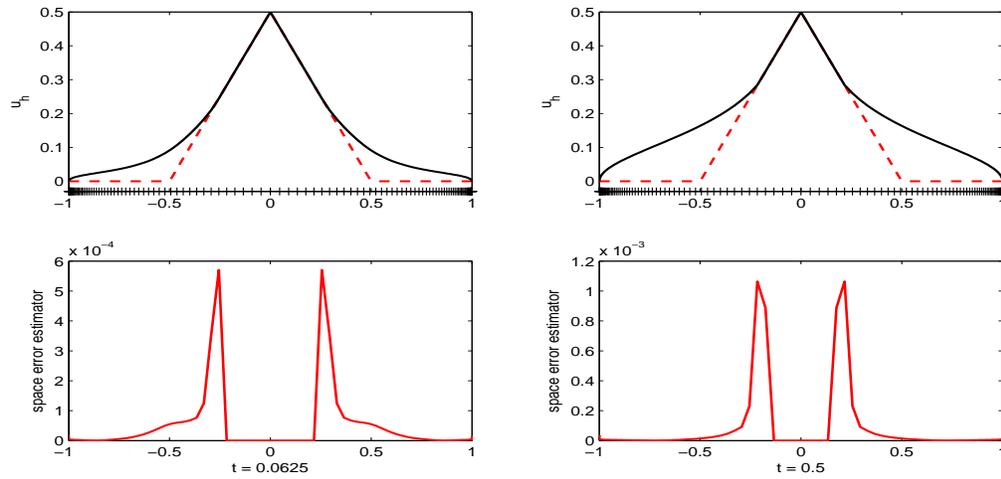


Figure 8.9: Numerical solution and space error estimator associated with it for parabolic variational inequality Example 8.2.3 ($N = 128$ and $\text{DOF} = 127$).

been reported in Figure 8.10. An interesting observation is the failure of fast pasting

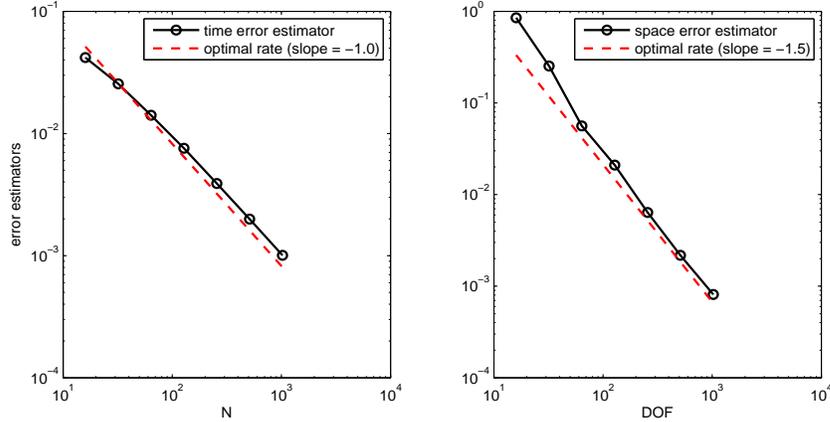


Figure 8.10: Asymptotic convergence rates for time (left) and space (right) error estimators for the parabolic variational inequality of Example 8.2.3.

for this case. From Figure 8.11, we can see that the free boundary point jumps from 0.5 (initially) to 0.34 (after one time step). This is the case even if one chooses extremely small time steps.

8.3 Adaptivity

In previous section, enough numerical evidence has been collected that the proposed error estimators are reliable and efficient. Now the question is what we can gain by using adaptive mesh refinement instead of uniform refinement. In this section, we compare adaptivity and uniform mesh refinement.

8.3.1 1d American Option

Under the standard assumption of a frictionless market without arbitrage, one can formulate the 1d American option as an optimal stopping problem and find that the option contract price $V(S, t)$ satisfies a parabolic variational inequality problem. Using the time to maturity $\tilde{t} = T - t$ and $x = \log S$ as independent variables, the

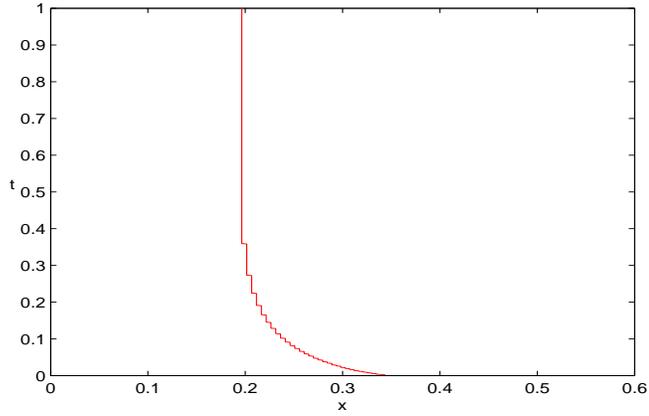


Figure 8.11: Jump of the free boundary point: time-step $k = 1/1024$ and space meshsize $h = 1/1024$. The approximate free boundary point (red curve) jumps from 0.5 (initially) to 0.34 (after one time step).

function $u(x, \tilde{t}) := V(e^x, T - \tilde{t})$ satisfies the following differential inequality (we will write t instead of \tilde{t} from now on):

$$\frac{\partial u}{\partial t} + \mathcal{A}u = \frac{\partial u}{\partial t} - \frac{\sigma^2}{2} \frac{\partial^2 u}{\partial x^2} + \left(\frac{\sigma^2}{2} - r \right) \frac{\partial u}{\partial x} + ru \geq 0 \quad \text{for } x \in \mathbb{R} \text{ and } 0 < t < T \quad (8.2)$$

with the obstacle constraint

$$u(x, t) \geq \chi(x) \quad \text{for } x \in \mathbb{R} \text{ and } 0 < t < T \quad (8.3)$$

and the initial condition

$$u(x, 0) = u_0(x) = H(e^x) = \max(K - e^x, 0) \quad \text{for } x \in \mathbb{R} \quad (8.4)$$

where $\chi(x) = u_0(x)$ is the payoff function in the log of the asset price. The solution $u(x, t)$ has a singular behavior in both time and space close to $t = 0$ and $x = \log K$ (i.e., time close to maturity and price close to strike price).

In American option pricing problem, we start from an initial solution which is in a Sobolev space, $\tilde{H}^{\frac{3}{2}-\epsilon}$, for $\epsilon > 0$. From the results in [110], we can conclude that $u_0 \in \tilde{H}^1$ implies the error in $L^2(\tilde{H}^1)$ -norm converges with order $O(k^{1/2})$ and $u_0 \in \tilde{H}^2$ implies order $O(k)$. And now, given the fractional regularity right in

between \tilde{H}^1 and \tilde{H}^2 , we expect, from interpolation theory, that the convergence order with uniform time-step would be about $O(k^{3/4})$. Numerical experiments in §8.3.1 (see also [104, Example 5.4]) confirm this expectation and by using a priori designed graded time steps the optimal convergence rate can be restored as pointed in [104].

For numerical experiments, we take an American put option problem on a single stock with strike price $K = 100$, maturity time $T = 0.5$ year, volatility $\sigma = 0.4$ and interest rate $r = 6\%$. We choose space domain to be $(-1, 7)$. The results (see Figure 8.12) show that if we choose the transfer operator to be the ordinary interpolation operator I_{n-1}^n , time step size k^n goes to about 10^{-50} even if we set the maximum number of iterations for time step size adaptation to be 20. This is exactly the effect we expected as in the thought experiment conducted in §6.7.1. In contrast, if we choose \mathcal{I}_{n-1}^n as in (6.71), the adaptive program gives reasonable time-steps which increase as time does.

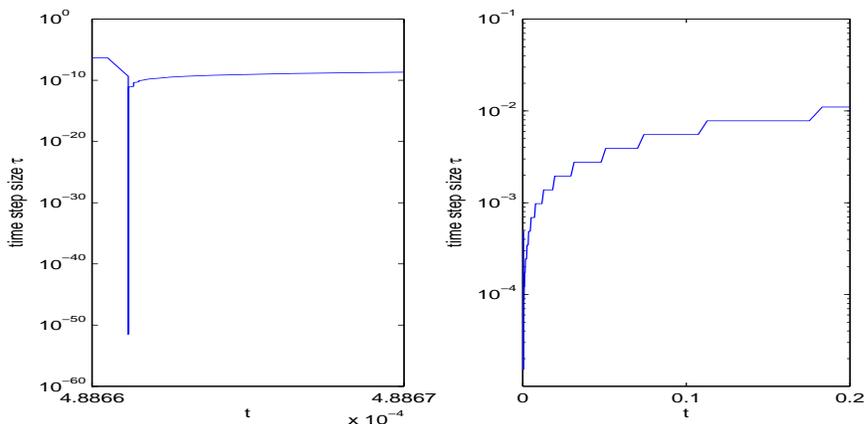


Figure 8.12: 1d American Option: time step size k^n . Left: using interpolation operator I_{n-1}^n , time-steps k^n decrease dramatically at the beginning because of the effect explained in §2. Right: using operator \mathcal{I}_{n-1}^n yields adaptively generated graded time-steps k^n .

Figure 8.13 shows that uniform refinement gives a suboptimal convergence rate, due to the singularity close to $t = 0$, but the adaptive algorithm restores

the optimal convergence rate. Furthermore, we see from Figure 8.12 (Right) that time-steps are automatically graded as t approaches 0.

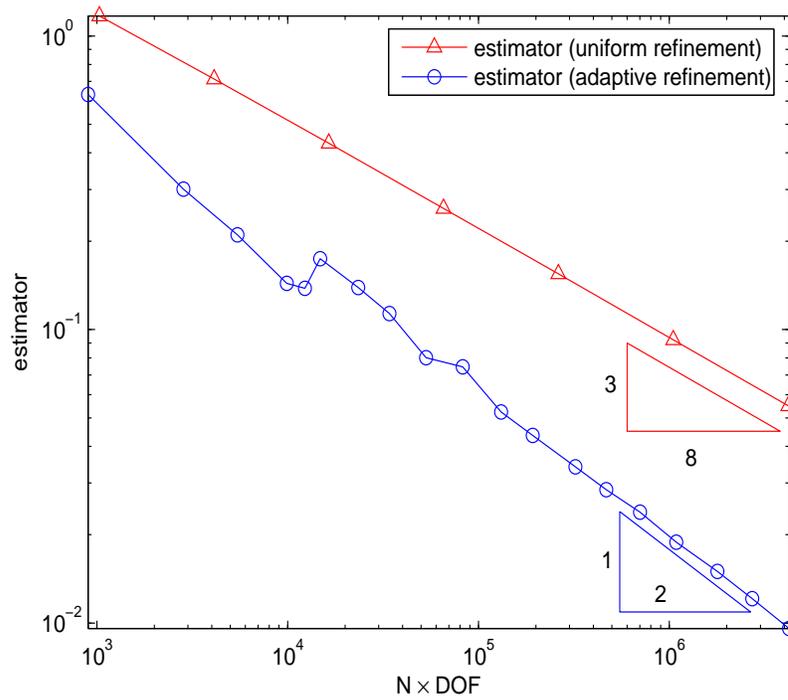


Figure 8.13: 1d American Option: error estimator in $L^2(\tilde{H}^1)$ -norm. Adaptive refinement achieve faster convergence rate, which is optimal $\mathcal{O}((N \cdot \text{DOF})^{-1/2})$.

8.3.2 1d Tent Obstacle

We use the same test example as in §8.1.2. In this case, the singular point $x = 1/3$ is never a mesh point if starting from a single macro element $[0.0, 1.0]$ and bisection method for refinement. Table 8.4 in §8.1.2 demonstrates uniform refinement gives suboptimal convergence rate (see also [104, Table 4]). By using the adaptive algorithm, we can recover the optimal convergence rate and both error estimator and real error converge at almost the same rate (see Figure 8.14).

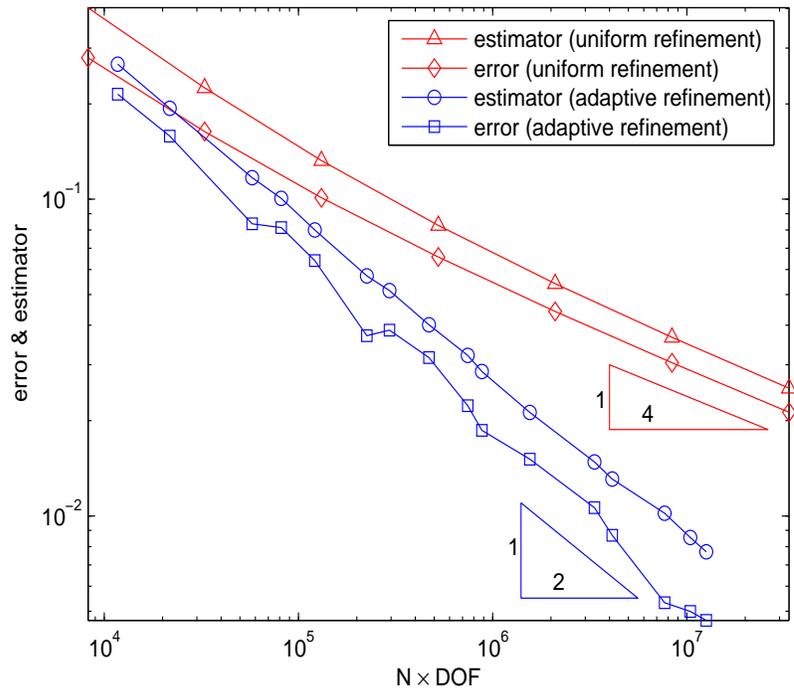


Figure 8.14: 1d Tent Obstacle: error estimator and exact error in $L^2(\tilde{H}^1)$ -norm. For both uniform and adaptive refinements, the a posteriori error estimator converges at the same rate as the exact error asymptotically. Adaptive refinement achieves faster convergence $\mathcal{O}((N \cdot \text{DOF})^{-1/2})$, which is optimal rate.

8.3.3 2d Tent Obstacle

This is an example with operator $\mathcal{A} := -\Delta$ and obstacle

$$\chi(x) = \begin{cases} 2|x| & \text{if } |x| \leq \frac{1}{2} \\ 2 - 2|x| & \text{otherwise,} \end{cases} \quad (8.5)$$

which is obtained by revolving a 1d tent similar to the 1d tent around the z-axis.

The exact solution is known:

$$u(x, t) = \begin{cases} \frac{(|x| - 1)^2}{1 - F(t)} + 1 - F(t) & \text{if } |x| > F(t) \\ \frac{|x|^2}{1 - F(t)} + 1 - F(t) & \text{if } |x| < 1 - F(t) \\ \chi(x, t) & \text{if } F(t) \leq |x| \leq 1 - F(t), \end{cases} \quad (8.6)$$

where $F(t) = \frac{3}{5} + \frac{3}{10}t$.

The numerical simulation is done in a square domain $\Omega = [-1, 1]^2$ for $t \in [0, 0.25]$ with exact initial and boundary conditions. Because in this problem, the exact solution is no longer in $\tilde{H}^2(\Omega)$, the uniform refinements give a suboptimal convergence rate. On the other hand, the adaptive program converges at an optimal rate (see Figure 8.16).

8.4 Convergence of Discrete Solver

In this section, we design several examples to test the discrete solver discussed in §7.5. We choose the simplest setting $\mathcal{A} = -\Delta$ throughout this section. Consider the following elliptic variational inequality problem (4.1). For comparison, we use projected SOR to find the “exact” solution by an overkill computation.

8.4.1 Smooth Constraint

We first take the example in [130]. Let $\Omega = [-2, 2]^2$, $f = 0$ and

$$\chi = \begin{cases} \sqrt{1 - |x|^2} & |x| \leq 1 \\ -1 & \text{otherwise.} \end{cases}$$

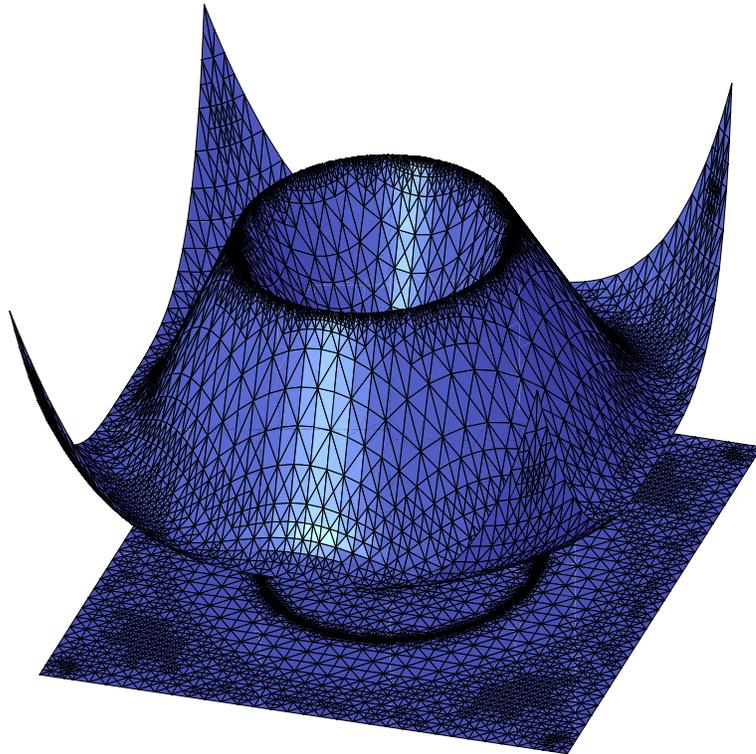


Figure 8.15: 2d Tent Obstacle: graph and grids of the numerical solution of adaptive method at time $t = 0.75$. There is a circular kink at $|x| = 0.5$, which requires fine mesh for obstacle resolution.

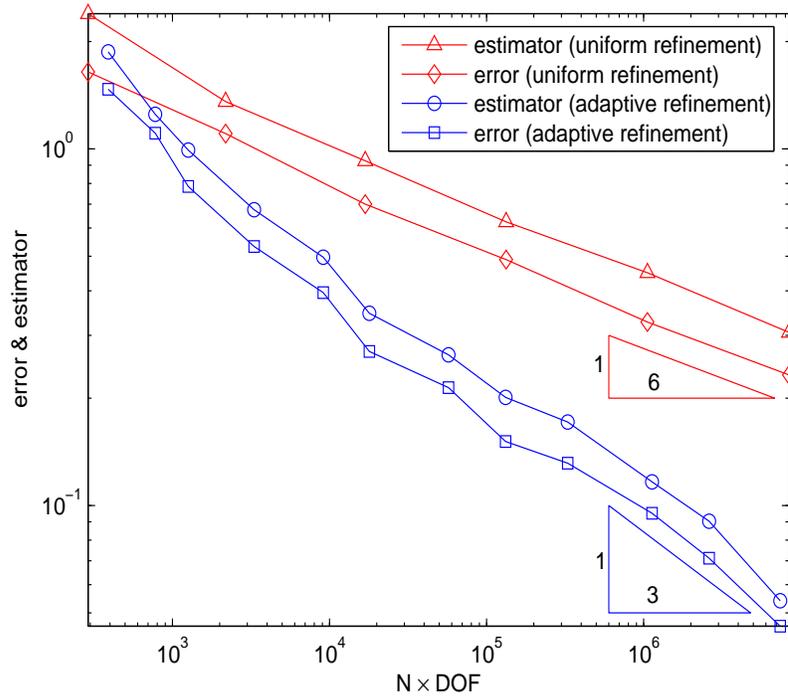


Figure 8.16: 2d Tent Obstacle: error estimator and exact error in $L^2(\tilde{H}^1)$ -norm. For both uniform and adaptive refinements, the a posteriori error estimator converges at the same rate as the exact error asymptotically. Adaptive refinement achieves faster convergence rate, $\mathcal{O}((N \cdot \text{DOF})^{-1/3})$, which is optimal.

In this case, the exact solution is known

$$u_* = \begin{cases} \sqrt{1 - |x|^2} & |x| \leq r_* \\ -r_*^2 \ln(|x|/2) \sqrt{1 - r_*^2} & \text{otherwise,} \end{cases}$$

where $r_* \approx 0.6979651482$. The convergence for a sequence of adaptive meshes are reported in Figure 8.17, 8.18, 8.19, and 8.20.

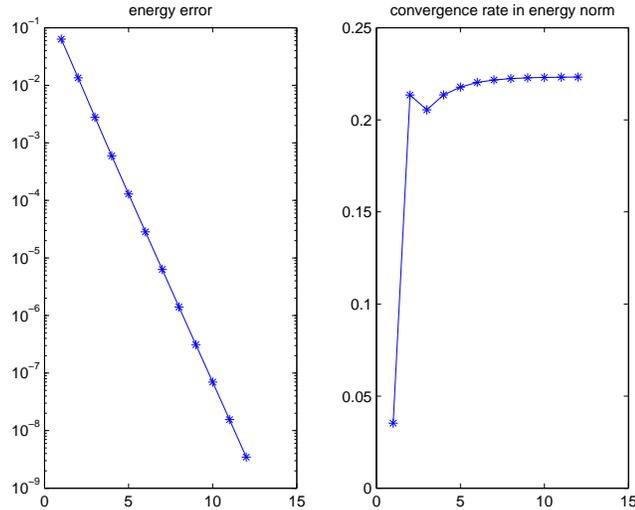


Figure 8.17: Convergence rate of multilevel solver SSC-CDM on a graded mesh with $h_{\min} = 8.839 \times 10^{-2}$. The convergence rate is globally linear as suggested by Theorem 7.12.

8.4.2 Inactive Constraint

Let $\Omega = [-1, 1]^2$, $\chi = 0$ and $f = 1$. In this case, the constraint is inactive and problem is equivalent to a linear equation. We report the reduction rate and h_{\min} in Table 8.15. The reduction rate is still mesh dependent. However, this is not a contradiction because in the theory by Tai and Xu [132] the convergence rate depends on $|\ln(h)|$ also.

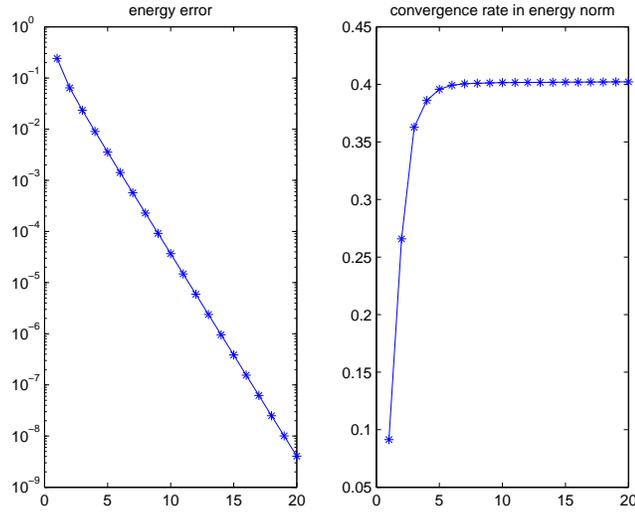


Figure 8.18: Convergence rate of multilevel solver SSC-CDM on a graded mesh with $h_{\min} = 6.250 \times 10^{-2}$. The discrete solver SSC-CDM still converges linearly but with bigger reduction rate.

8.4.3 Kink Constraint

Now we consider the following obstacle with a kink on $\Omega = [-1, 1]^2$

$$\chi = \begin{cases} 1 - 2|x| & |x| \leq 0.5 \\ 0 & \text{otherwise.} \end{cases}$$

We take $f = 0$ and report the reduction rates in Table 8.16.

8.4.4 Singular Constraint

We modify the previous kink constraint to the following singular (discontinuous) obstacle constraint

$$\chi = \begin{cases} 1 & |x| \leq \varepsilon \\ 0 & \text{otherwise,} \end{cases}$$

where $\varepsilon \approx 2.220 \times 10^{-16}$ is the machine epsilon. The reduction rates are reported in Table 8.17. Because of the point singularity, the adaptive meshes in this example are strongly graded and we observe the log dependence on h_{\min} ; see Figure 8.21.

It	DOF	h_{\min}	Reduction Rate
6	81	8.839e-1	0.19
7	139	6.250e-2	0.20
8	247	4.419e-2	0.38
9	434	4.419e-2	0.38
10	748	3.125e-2	0.49

Table 8.15: SSC-CDM convergence rate: inactive constraint. In this case, the reduction rate is comparable to \backslash -cycle multigrid method for linear elliptic equations.

It	DOF	h_{\min}	Reduction Rate
6	72	4.419e-2	0.24
7	119	2.210e-2	0.30
8	214	1.563e-2	0.43
9	384	7.813e-3	0.46
10	698	3.906e-3	0.50
11	1276	2.762e-3	0.59

Table 8.16: SSC-CDM convergence rate: kink constraint. For this example with a singular constraint, the reduction rate is closer to 1 than the previous examples with smooth obstacles; but it is still linear.

It	DOF	h_{\min}	Reduction Rate
6	65	6.250e-2	0.48
7	86	3.125e-2	0.62
8	104	2.210e-2	0.72
9	126	1.105e-2	0.80
10	148	5.524e-3	0.82
11	172	2.762e-3	0.85

Table 8.17: SSC-CDM convergence rate: singular constraint. The obstacle is singular and discontinuous in this example, which results in highly graded meshes. The method is still linear with a reduction rate close to 1 when meshsize is small.

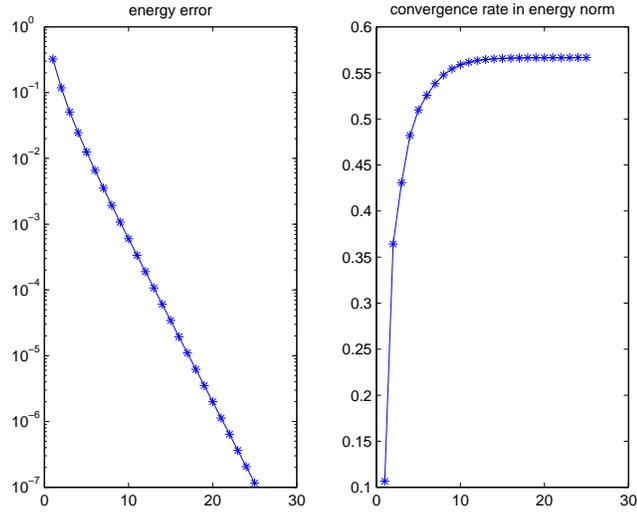


Figure 8.19: Convergence rate of multilevel solver SSC-CDM on a graded mesh with $h_{\min} = 4.419 \times 10^{-2}$.

8.4.5 Unstable Constraint

The last test example is taken from [127]. We take $\Omega = [-1, 1]^2$ and $\chi = 0$. The exact solution is constructed to be

$$u_* = \begin{cases} \frac{1}{2}x_1^4 & x_1 > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Furthermore, the right-hand side is chosen to be

$$f = \begin{cases} -6x_1^2 & x_1 > 0 \\ 0 & \text{otherwise,} \end{cases}$$

such the the contact is unstable. This means the strict complementarity condition is not satisfied in this example. The reduction rate is reported in Table 8.18.

8.5 Conclusions

We have developed a novel a priori and a posteriori error analysis for parabolic integro-differential variational inequalities, including localization features to the non-contact region, and illustrated it with several numerical experiments, some relevant

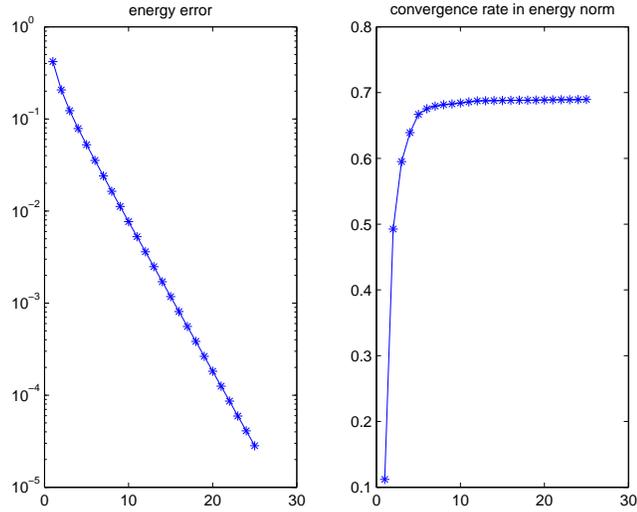


Figure 8.20: Convergence rate of multilevel solver SSC-CDM on a graded mesh with $h_{\min} = 3.125 \times 10^{-2}$.

It	DOF	h_{\min}	Reduction Rate
6	49	8.839e-2	0.38
7	83	6.250e-2	0.36
8	137	4.419e-2	0.42
9	227	3.125e-2	0.35
10	417	2.210e-3	0.40

Table 8.18: SSC-CDM convergence rate: unstable constraint.

in finance. Upon comparing theory and practice we have the following concluding remarks:

- Error Decay:** For problems with smooth data, the energy error in $L^2(0, T; H^1(\Omega))$ decays linearly, namely $\mathcal{O}(h + k)$. This coincides with the a priori theory developed in Chapter 5. If the obstacle χ exhibits a singularity not resolved by the mesh, as in Section 8.1.2, or the initial condition is rough, as in Section 8.3.1, the actual error decays with a suboptimal rate. Suitable mesh refinement in either space or time appears to cure this problem; see again Sections 8.1.2 and 8.3.1.

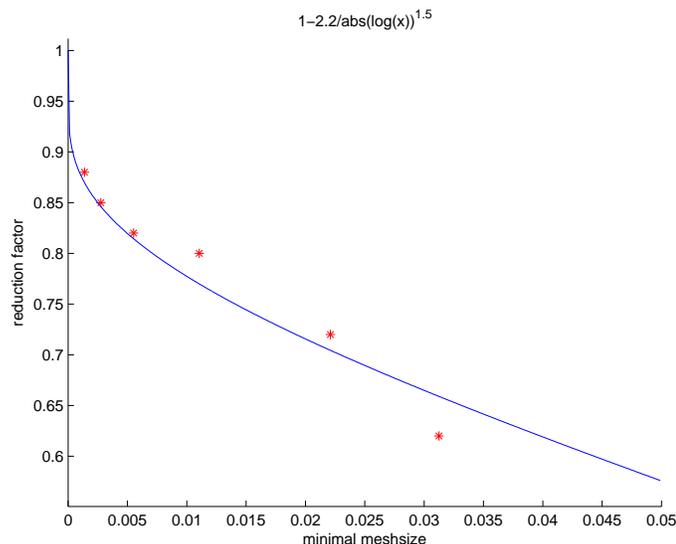


Figure 8.21: Singular constraint example: the reduction rate depends on $|\ln(h_{\min})|$.

- Estimator Decay:** The numerical experiments corroborate that the proposed fully localized error estimator \mathcal{E} decays with the same rate as the actual error e . We have demonstrated experimentally that the components $\mathcal{E}_h, \mathcal{E}_\tau, \mathcal{E}_\chi$ of \mathcal{E} provided valuable a posteriori information of the solution. Experiments with adaptive time-space mesh refinement show effectivity of the error indicators suggested by our a posteriori error estimation.
- Localization of Space Estimator:** Figures 8.3 and 8.4 show that the nodal-based space estimator $\Upsilon_h^n(z)$ vanishes at full-contact nodes $z \in \mathcal{C}_h^n$. Its contribution comes only from the non-contact region where the solution behaves like the solution of a linear parabolic equation. This estimator yields an upper bound also for globally linear parabolic problems and seems to be new in the literature of parabolic PDE.
- Exercise Boundary Approximation:** Accurate approximation of the free (exercise) boundary is an important problem in option pricing. Numerical results in Sections 8.3.2 and 8.3.3, particularly Figures 8.3 and 8.4 as well as Tables 8.3 and 8.10, suggest an excellent agreement between approximate and exact free boundaries. This observation could be made rigorous, upon extending the idea

in [114], provided pointwise a posteriori error estimates were available. This is under further investigation.

- **Multilevel Solver on Bisection Meshes:** The SSC-CDM yields globally linear convergence rate even on highly graded meshes. Unfortunately, the reduction rate of error in energy between two consecutive iterations depends on minimal meshsize due to the unstable decomposition used.

BIBLIOGRAPHY

- [1] R. A. Adams. *Sobolev spaces*. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1975. Pure and Applied Mathematics, Vol. 65.
- [2] M. Ainsworth and J. T. Oden. *A posteriori error estimation in finite element analysis*. Pure and Applied Mathematics (New York). Wiley-Interscience [John Wiley & Sons], New York, 2000.
- [3] W. Allegretto, Y. Lin, and H. Yang. Finite element error estimates for a nonlocal problem in American option valuation. *SIAM J. Numer. Anal.*, 39(3):834–857 (electronic), 2001.
- [4] L. Andersen and J. Andreasen. Jump-diffusion processes: Volatility smile fitting and numerical methods for option pricing. *Review of Derivatives Research*, 4(3):231–262, 2000.
- [5] L. Angermann and S. Wang. Convergence of a fitted finite volume method for the penalized blackscholes equation governing european and american option pricing. *Numerische Mathematik*, 2007.
- [6] D. N. Arnold. *A concise introduction to numerical analysis*. 2001.
- [7] I. Babuška and A. K. Aziz. On the angle condition in the finite element method. *SIAM Journal on Numerical Analysis*, 13(2):214–226, 1976.
- [8] I. Babuška and W. C. Rheinboldt. A posteriori error error estimates for the finite element method. *International Journal for Numerical Methods in Engineering*, 12:1597–1615, 1978.
- [9] L. Badea, X.-C. Tai, and J. Wang. Convergence rate analysis of a multiplicative Schwarz method for variational inequalities. *SIAM J. Numer. Anal.*, 41(3):1052–1073 (electronic), 2003.
- [10] C. Baiocchi. Estimations d’erreur dans L^∞ pour les inéquations à obstacle. pages 27–34. Lecture Notes in Math., Vol. 606, 1977.
- [11] C. Baiocchi. Discretization of evolution variational inequalities. In *Partial differential equations and the calculus of variations, Vol. I*, volume 1 of *Progr. Nonlinear Differential Equations Appl.*, pages 59–92, Boston, MA, 1989. Birkhäuser Boston.

- [12] W. Bangerth and R. Rannacher. *Adaptive finite element methods for differential equations*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2003.
- [13] R. E. Bank, P. E. Gill, and R. F. Marcia. *Interior methods for a class of elliptic variational inequalities*, volume 30 of *Lect. Notes Comput. Sci. Eng.*, pages 218–235. Springer, Berlin, 2003.
- [14] E. Bänsch. Local mesh refinement in 2 and 3 dimensions. *Impact of Computing in Science and Engineering*, 3:181–191, 1991.
- [15] S. Bartels and C. Carstensen. Averaging techniques yield reliable a posteriori finite element error control for obstacle problems. *Numer. Math.*, 99(2):225–249, 2004.
- [16] A. Bergam, C. Bernardi, and Z. Mghazli. A posteriori analysis of the finite element discretization of some parabolic equations. *Math. Comp.*, 74(251):1117–1138 (electronic), 2005.
- [17] A. E. Berger and R. S. Falk. An error estimate for the truncation method for the solution of parabolic obstacle variational inequalities. *Math. Comp.*, 31(139):619–628, 1977.
- [18] J. Bergh and J. Löfström. *Interpolation Spaces*. Springer, 1976.
- [19] J. Bey. Simplicial grid refinement: on freudenthal’s algorithm and the optimal number of congruence classes. *Numerische Mathematik*, 85(1):1–29, 2000.
- [20] M. Bieterman and I. Babuška. The finite element method for parabolic equations. I. A posteriori error estimation. *Numer. Math.*, 40(3):339–371, 1982.
- [21] M. Bieterman and I. Babuška. The finite element method for parabolic equations. II. A posteriori error estimation and adaptive approach. *Numer. Math.*, 40(3):339–371, 1982.
- [22] P. Binev, W. Dahmen, and R. DeVore. Adaptive finite element methods with convergence rates. *Numerische Mathematik*, 97(2):219–268, 2004.
- [23] F. Black and M. Scholes. Pricing of options and corporate liabilities. *Journal of Political economy*, 81(3):637–654, 1973.
- [24] S. I. Boyarchenko and S. Z. Levendorskii. Perpetual American options under Lévy processes. *SIAM J. Control Optim.*, 40(6):1663–1696 (electronic), 2002.
- [25] D. Braess. *Finite elements*. Cambridge University Press, Cambridge, second edition, 2001. Theory, fast solvers, and applications in solid mechanics, Translated from the 1992 German edition by Larry L. Schumaker.
- [26] D. Braess. A posteriori error estimators for obstacle problems—another look. *Numer. Math.*, 101(3):415–421, 2005.

- [27] J. H. Bramble, J. E. Pasciak, and A. H. Schatz. The construction of preconditioners for elliptic problems by substructuring, I. *Mathematics of Computation*, 47:103–134, 1986.
- [28] A. Brandt and C. W. Cryer. Multigrid algorithms for the solution of linear complementarity problems arising from free boundary problems. *SIAM J. Sci. Statist. Comput.*, 4(4):655–684, 1983.
- [29] S. C. Brenner and L. R. Scott. *The mathematical theory of finite element methods*, volume 15 of *Texts in Applied Mathematics*. Springer-Verlag, New York, second edition, 2002.
- [30] H. Brézis. Problèmes unilatéraux. *J. Math. Pures Appl. (9)*, 51:1–168, 1972.
- [31] H. Brézis. *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*. North-Holland Publishing Co., Amsterdam, 1973. North-Holland Mathematics Studies, No. 5. Notas de Matemática (50).
- [32] H. Brézis and F. E. Browder. Nonlinear integral equations and systems of Hammerstein type. *Advances in Math.*, 18(2):115–147, 1975.
- [33] H. Brézis and M. Sibony. Équivalence de deux inéquations variationnelles et applications. *Arch. Rational Mech. Anal.*, 41:254–265, 1971.
- [34] H. R. Brezis and G. Stampacchia. Sur la régularité de la solution d'inéquations elliptiques. *Bull. Soc. Math. France*, 96:153–180, 1968.
- [35] F. Brezzi, W. W. Hager, and P.-A. Raviart. Error estimates for the finite element solution of variational inequalities. *Numer. Math.*, 28(4):431–443, 1977.
- [36] F. Brezzi, W. W. Hager, and P.-A. Raviart. Error estimates for the finite element solution of variational inequalities. II. Mixed methods. *Numer. Math.*, 31(1):1–16, 1978/79.
- [37] M. Broadie and J. Detemple. Recent advances in numerical methods for pricing derivative securities. pages 43–66, 1997.
- [38] L. A. Caffarelli. The regularity of monotone maps of finite compression. *Comm. Pure Appl. Math.*, 50(6):563–591, 1997.
- [39] L. A. Caffarelli. The obstacle problem revisited. *J. Fourier Anal. Appl.*, 4(4-5):383–402, 1998.
- [40] P. Carr, H. Geman, D. B. Madan, and M. Yor. The fine structure of asset returns: An empirical investigation. *JOURNAL OF BUSINESS*, 75:305–332, 2002.

- [41] C. Carstensen. Efficiency of a posteriori BEM-error estimates for first-kind integral equations on quasi-uniform meshes. *Math. Comp.*, 65(213):69–84, 1996.
- [42] C. Carstensen. An a posteriori error estimate for a first-kind integral equation. *Math. Comp.*, 66(217):139–155, 1997.
- [43] C. Carstensen and E. P. Stephan. A posteriori error estimates for boundary element methods. *Math. Comp.*, 64(210):483–500, 1995.
- [44] J. M. Cascon, C. Kreuzer, R. H. Nochetto, and K. G. Siebert. Quasi-optimal convergence rate for an adaptive finite element method. (submitted).
- [45] L. Chen, R. Nochetto, and J. Xu. Multilevel methods on bisection grids. Technical report, University of Maryland, 2007.
- [46] L. Chen and C.-S. Zhang. Afem@matlab: a matlab package of adaptive finite element methods. 2006.
- [47] L. Chen and C.-S. Zhang. A coarsening algorithm and multilevel methods on adaptive grids by newest vertex bisection. (in preparation).
- [48] Z. Chen and J. Feng. An adaptive finite element algorithm with reliable and efficient error control for linear parabolic problems. *Math. Comp.*, 73(247):1167–1193 (electronic), 2004.
- [49] Z. Chen and R. H. Nochetto. Residual type a posteriori error estimates for elliptic obstacle problems. *Numer. Math.*, 84(4):527–548, 2000.
- [50] P. G. Ciarlet. *The Finite Element Method for Elliptic Problems*, volume 4 of *Studies in Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam-New York-Oxford, 1978.
- [51] P. Clément. Approximation by finite element functions using local regularization. *RAIRO Anal. Numér.*, 2:77–84, 1975.
- [52] C. W. Cryer. *Successive overrelaxation methods for solving linear complementarity problems arising from free boundary problems*, pages 109–131. Ist. Naz. Alta Mat. Francesco Severi, Rome, 1980.
- [53] R. A. DeVore. Nonlinear approximation. *Acta Numerica*, pages 51–150, 1998.
- [54] W. Dörfler. A convergent adaptive algorithm for Poisson’s equation. *SIAM Journal on Numerical Analysis*, 33:1106–1124, 1996.
- [55] J. Duoandikoetxea. *Fourier Analysis*. Graduate Studies in Mathematics, vol. 29. American Math. Soc., Providence, RI, 2001.
- [56] T. F. Dupont. Mesh modification for evolution equations. *Mathematics of Computation*, 39(159):85–107, 1982.

- [57] K. Erickson and C. Johnson. Adaptive finite element methods for parabolic problems. i. a linear model problem. *SIAM Journal on Numerical Analysis*, 28(1):43–77, 1991.
- [58] K. Eriksson and C. Johnson. Adaptive finite element methods for parabolic problems II: Optimal error estimates in $l_\infty l_2$ and $l_\infty l_\infty$. *SIAM Journal on Numerical Analysis*, 32(3):706–740, 1995.
- [59] K. Eriksson and C. Johnson. Adaptive finite element methods for parabolic problems IV: Nonlinear problems. *SIAM Journal on Numerical Analysis*, 32:1729–1749, 1995.
- [60] K. Eriksson and C. Johnson. Adaptive finite element methods for parabolic problems V: Long-time integration. *SIAM Journal on Numerical Analysis*, 32(6):1750–1763, 1995.
- [61] K. Eriksson, C. Johnson, and S. Larsson. Adaptive finite element methods for parabolic problems VI: Analytic semigroups. *SIAM Journal on Numerical Analysis*, 35(4):1315–1325, 1998.
- [62] L. C. Evans. *Partial Differential Equations*. American Mathematical Society, 1998.
- [63] F. Facchinei and J.-S. Pang. *Finite-dimensional variational inequalities and complementarity problems. Vol. I*. Springer Series in Operations Research. Springer-Verlag, New York, 2003.
- [64] F. Facchinei and J.-S. Pang. *Finite-dimensional variational inequalities and complementarity problems, Vol. II*. Springer Series in Operations Research. Springer-Verlag, New York, 2003.
- [65] B. Faermann. *Lokale a-posteriori Fehlerschätzer bei der Diskretisierung von Randintegralgleichungen*. PhD thesis, University of Kiel, 1993.
- [66] B. Faermann. Efficient and reliable a-posteriori error estimates for boundary element methods. 379:87–91, 1998.
- [67] B. Faermann. Efficient and reliable a posteriori error estimates for boundary integral operators of positive and negative order. pages 303–310, 1998.
- [68] B. Faermann. Local a-posteriori error indicators for the Galerkin discretization of boundary integral equations. *Numer. Math.*, 79(1):43–76, 1998.
- [69] R. S. Falk. Error estimates for the approximation of a class of variational inequalities. *Mathematics of Computation*, 28:963–971, 1974.
- [70] M. C. Ferris and J. S. Pang. Engineering and economic applications of complementarity problems. *SIAM Rev.*, 39(4):669–713, 1997.

- [71] A. Fetter. L^∞ -error estimate for an approximation of a parabolic variational inequality. *Numer. Math.*, 50(5):557–565, 1987.
- [72] F. Fierro and A. Veaser. A posteriori error estimators for regularized total variation of characteristic functions. *SIAM J. Numer. Anal.*, 41(6):2032–2055 (electronic), 2003.
- [73] A. Friedman. *Variational principles and free-boundary problems*. Robert E. Krieger Publishing Co. Inc., Malabar, FL, second edition, 1988.
- [74] R. Glowinski. *Numerical methods for nonlinear variational problems*. Springer-Verlag, New York, 1984.
- [75] R. Glowinski, J. Lions, and R. Trémolières. *Numerical analysis of variational inequalities*. North-Holland New York, 1981.
- [76] H. Han and X. Wu. A fast numerical method for the Black-Scholes equation of American options. *SIAM J. Numer. Anal.*, 41(6):2081–2095 (electronic), 2003.
- [77] A. Hirta and D. B. Madan. Pricing american options under variance gamma. *Journal of Computational Finance*, 7(2):63–80, 2003.
- [78] J. Hull. *Options, Futures, and Other Derivatives*. Prentice Hall, 2005.
- [79] K. Ito and K. Kunisch. Parabolic variational inequalities: the Lagrange multiplier approach. *J. Math. Pures Appl. (9)*, 85(3):415–449, 2006.
- [80] P. Jaillet, D. Lamberton, and B. Lapeyre. Inéquations variationnelles et théorie des options. *C. R. Acad. Sci. Paris Sér. I Math.*, 307(19):961–965, 1988.
- [81] P. Jaillet, D. Lamberton, and B. Lapeyre. Variational inequalities and the pricing of American options. *Acta Appl. Math.*, 21(3):263–289, 1990.
- [82] C. Johnson. A convergence estimate for an approximation of a parabolic variational inequality. *SIAM J. Numer. Anal.*, 13(4):599–606, 1976.
- [83] C. Johnson. *Numerical Solution of Partial Differential Equations by the Finite Element Method*. Cambridge University Press, Cambridge, 1987.
- [84] D. Kinderlehrer and G. Stampacchia. *An introduction to variational inequalities and their applications*, volume 88 of *Pure and Applied Mathematics*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1980.
- [85] R. Kornhuber. Monotone multigrid methods for elliptic variational inequalities. I. *Numer. Math.*, 69(2):167–184, 1994.
- [86] R. Kornhuber. Monotone multigrid methods for elliptic variational inequalities. II. *Numer. Math.*, 72(4):481–499, 1996.

- [87] R. Kornhuber. *Adaptive monotone multigrid methods for nonlinear variational problems*. 1997.
- [88] I. Kossaczky. A recursive approach to local mesh refinement in two and three dimensions. *Journal of Computational and Applied Mathematics*, 55:275–288, 1994.
- [89] H. W. Kuhn and A. W. Tucker. Nonlinear programming. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1950*, pages 481–492, Berkeley and Los Angeles, 1951. University of California Press.
- [90] O. Lakkis and C. Makridakis. Elliptic reconstruction and a posteriori error estimates for fully discrete linear parabolic problems. *Math. Comp.*, 75(256):1627–1658 (electronic), 2006.
- [91] B. Leblanc and M. Yor. Lvy processes in finance: a remedy to the non-stationarity of continuous martingales. *Finance and Stochastics*, 2(4):399–408, August 1998.
- [92] J. Lions and E. Magenes. *Non-Homogeneous Boundary Value Problems and Applications I*. Springer-Verlag Berlin Heidelberg New York, 1973.
- [93] J.-L. Lions and G. Stampacchia. Variational inequalities. *Comm. Pure Appl. Math.*, 20:493–519, 1967.
- [94] D. B. Madan, P. P. Carr, and E. C. Chang. The variance gamma process and option pricing. *Europ. Finance Rev.*, 2:79–105, 1998.
- [95] D. B. Madan and E. Seneta. The variance-gamma (v. g.) model for share market returns. *J. Business*, 63:511–524, 1990.
- [96] C. Makridakis and R. H. Nochetto. Elliptic reconstruction and a posteriori error estimates for parabolic problems. *SIAM J. Numer. Anal.*, 41(4):1585–1594 (electronic), 2003.
- [97] J. Mandel. A multilevel iterative method for symmetric, positive definite linear complementarity problems. *Appl. Math. Optim.*, 11(1):77–95, 1984.
- [98] A.-M. Matache, P.-A. Nitsche, and C. Schwab. Wavelet Galerkin pricing of American options on Lévy driven assets. *Quant. Finance*, 5(4):403–424, 2005.
- [99] A.-M. Matache, C. Schwab, and T. P. Wihler. Fast numerical solution of parabolic integrodifferential equations with applications in finance. *SIAM J. Sci. Comput.*, 27(2):369–393 (electronic), 2005.
- [100] A.-M. Matache, C. Schwab, and T. P. Wihler. Linear complexity solution of parabolic integro-differential equations. *Numer. Math.*, 104(1):69–102, 2006.

- [101] K. Mekchay and R. Nochetto. Convergence of adaptive finite element methods for general second order linear elliptic PDE. *SIAM Journal on Numerical Analysis*, 43(5):1803–1827, 2005.
- [102] R. C. Merton. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3(1-2):125–144, 1976. available at <http://ideas.repec.org/a/eee/jfinec/v3y1976i1-2p125-144.html>.
- [103] W. F. Mitchell. A comparison of adaptive refinement techniques for elliptic problems. *ACM Transactions on Mathematical Software (TOMS) archive*, 15(4):326 – 347, 1989.
- [104] K.-S. Moon, R. H. Nochetto, T. von Petersdorff, and C.-S. Zhang. A posteriori error analysis for parabolic variational inequalities. *Mathematical Modelling and Numerical Analysis (M2AN)*, (to appear).
- [105] K.-S. Moon, E. Schwerin, A. Szepessy, and R. Tempone. Convergence rates for adaptive finite element methods. *Talk*, 2003.
- [106] P. Morin, R. Nochetto, and K. Siebert. Data oscillation and convergence of adaptive FEM. *SIAM Journal on Numerical Analysis*, 38(2):466–488, 2000.
- [107] P. Morin, R. H. Nochetto, and K. G. Siebert. Convergence of adaptive finite element methods. *SIAM Review*, 44(4):631–658, 2002.
- [108] P. Morin, R. H. Nochetto, and K. G. Siebert. Local problems on stars: A posteriori error estimators, convergence, and performance. *Mathematics of Computation*, 72:1067–1097, 2003.
- [109] R. H. Nochetto, G. Savaré, and C. Verdi. Error control of nonlinear evolution equations. *C. R. Acad. Sci. Paris Sér. I Math.*, 326(12):1437–1442, 1998.
- [110] R. H. Nochetto, G. Savaré, and C. Verdi. A posteriori error estimates for variable time-step discretizations of nonlinear evolution equations. *Comm. Pure Appl. Math.*, 53(5):525–589, 2000.
- [111] R. H. Nochetto, A. Schmidt, K. G. Siebert, and A. Veiser. Pointwise a posteriori error estimates for monotone semi-linear equations. *Numer. Math.*, 104(4):515–538, 2006.
- [112] R. H. Nochetto, A. Schmidt, and C. Verdi. A posteriori error estimation and adaptivity for degenerate parabolic problems. *Mathematics of Computation*, 229(220):1–24, 1999.
- [113] R. H. Nochetto, K. G. Siebert, and A. Veiser. Pointwise a posteriori error control for elliptic obstacle problems. *Numer. Math.*, 95(1):163–195, 2003.
- [114] R. H. Nochetto, K. G. Siebert, and A. Veiser. Fully localized a posteriori error estimators and barrier sets for contact problems. *SIAM J. Numer. Anal.*, 42(5):2118–2135 (electronic), 2005.

- [115] R. H. Nochetto, T. von Petersdorff, and C.-S. Zhang. A posteriori error estimates for a class of variational inequalities with integro-differential operators. (in preparation).
- [116] R. H. Nochetto and L. B. Wahlbin. Positivity preserving finite element approximation. *Math. Comp.*, 71(240):1405–1419 (electronic), 2002.
- [117] R. H. Nochetto and C.-S. Zhang. Adaptive mesh refinement for evolution obstacle problems. (in preparation).
- [118] D. Nualart and W. Schoutens. Backward stochastic differential equations and Feynman-Kac formula for Lévy processes, with applications in finance. *Bernoulli*, 7(5):761–776, 2001.
- [119] M. Picasso. Adaptive finite elements for a linear parabolic problem. *Comput. Methods Appl. Mech. Engrg.*, 167(3-4):223–237, 1998.
- [120] J.-F. Rodrigues. *Obstacle problems in mathematical physics*, volume 134 of *North-Holland Mathematics Studies*. North-Holland Publishing Co., Amsterdam, 1987. Notas de Matemática [Mathematical Notes], 114.
- [121] K.-I. Sato. *Lévy processes and infinitely divisible distributions*, volume 68 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999. Translated from the 1990 Japanese original, Revised by the author.
- [122] G. Savaré. Weak solutions and maximal regularity for abstract evolution inequalities. *Adv. Math. Sci. Appl.*, 6(2):377–418, 1996.
- [123] A. Schmidt and K. G. Siebert. *Design of adaptive finite element software*, volume 42 of *Lecture Notes in Computational Science and Engineering*. Springer-Verlag, Berlin, 2005. The finite element toolbox ALBERTA, With 1 CD-ROM (Unix/Linux).
- [124] C. Schwab. Variable order composite quadrature of singular and nearly singular integrals. *Computing*, 53(2):173–194, 1994.
- [125] R. Scott and S. Zhang. Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Mathematics of Computation*, 54:483–493, 1990.
- [126] E. G. Sewell. Automatic generation of triangulations for piecewise polynomial approximation. In *Ph. D. dissertation*. Purdue Univ., West Lafayette, Ind., 1972.
- [127] K. G. Siebert and A. Veiser. A constrained quadratic minimization with adaptive finite elements. *Quaderno n. 13/2005, Dipartimento di Matematica "F. Enriques", Universit degli Studi di Milano*.

- [128] L. Silvestre. Regularity of the obstacle problem for a fractional power of the Laplace operator. *Comm. Pure Appl. Math.*, 60(1):67–112, 2007.
- [129] R. Stevenson. Optimality of a standard adaptive finite element method. *Department of Mathematics*, 2005.
- [130] X.-C. Tai. Rate of convergence for some constraint decomposition methods for nonlinear variational inequalities. *Numer. Math.*, 93(4):755–786, 2003.
- [131] X.-C. Tai, B. Heimsund, and J. Xu. Rate of convergence for parallel subspace correction methods for nonlinear variational inequalities. In *Domain decomposition methods in science and engineering (Lyon, 2000)*, Theory Eng. Appl. Comput. Methods, pages 127–138. Internat. Center Numer. Methods Eng. (CIMNE), Barcelona, 2002.
- [132] X.-C. Tai and J. Xu. Global convergence of subspace correction methods for convex optimization problems. *Mathematics of Computation*, 71(237):105–124, 2002.
- [133] M. E. Taylor. *Pseudodifferential operators*, volume 34 of *Princeton Mathematical Series*. Princeton University Press, Princeton, N.J., 1981.
- [134] A. Veerer. Efficient and reliable a posteriori error estimators for elliptic obstacle problems. *SIAM J. Numer. Anal.*, 39(1):146–167 (electronic), 2001.
- [135] R. Verfürth. *A review of a posteriori error estimation and adaptive mesh refinement techniques*. Wiley and Teubner, 1996.
- [136] R. Verfürth. A posteriori error estimates for finite element discretizations of the heat equation. *Calcolo*, 40(3):195–212, 2003.
- [137] T. von Petersdorff and C. Schwab. Numerical solution of parabolic equations in high dimensions. *M2AN Math. Model. Numer. Anal.*, 38(1):93–127, 2004.
- [138] C. Vuik. An L^2 -error estimate for an approximation of the solution of a parabolic variational inequality. *Numer. Math.*, 57(5):453–471, 1990.
- [139] W. L. Wendland and D. H. Yu. Adaptive boundary element methods for strongly elliptic integral equations. *Numer. Math.*, 53(5):539–558, 1988.
- [140] W. L. Wendland and D. H. Yu. A posteriori local error estimates of boundary element methods with some pseudo-differential equations on closed curves. *J. Comput. Math.*, 10(3):273–289, 1992.
- [141] P. Wilmott. *Derivatives*. John Wiley and Sons Ltd, Chichester, 1998.
- [142] P. Wilmott, J. Dewynne, and S. Howison. *Option pricing: mathematical models and computation*. Oxford Financial Press, Oxford, UK, 1993.

- [143] H. Wu and Z. Chen. Uniform convergence of multigrid v-cycle on adaptively refined finite element meshes for second order elliptic problems. *Preprint*, 2003.
- [144] J. Xu. Iterative methods by space decomposition and subspace correction. *SIAM Review*, 34:581–613, 1992.
- [145] J. Xu. An introduction to multigrid convergence theory. In R. Chan, T. Chan, and G. Golub, editors, *Iterative Methods in Scientific Computing*. Springer-Verlag, 1997.
- [146] J. Xu and L. Zikatanov. The method of alternating projections and the method of subspace corrections in Hilbert space. *Journal of The American Mathematical Society*, 15:573–597, 2002.
- [147] D. H. Yu. A posteriori error estimates and adaptive approaches for some boundary element methods. pages 241–256, 1987.