

The Numerical Solution of Two-Factor Option Pricing Models

by

Robert Zvan

A thesis

presented to the University of Waterloo

in fulfilment of the

thesis requirement for the degree of

Doctor of Philosophy

in

Computer Science

Waterloo, Ontario, Canada, 2000

©Robert Zvan 2000



**National Library
of Canada**

**Acquisitions and
Bibliographic Services**

395 Wellington Street
Ottawa ON K1A 0N4
Canada

**Bibliothèque nationale
du Canada**

**Acquisitions et
services bibliographiques**

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-52030-7

Canada

The University of Waterloo requires the signatures of all persons using or photocopying this thesis. Please sign below, and give address and date.

Abstract

This work develops a nonconservative finite volume approach for solving two-dimensional partial differential equation option pricing models. The finite volume method is more flexible than finite difference schemes which are often described in the finance literature and frequently used in practice. Moreover, the finite volume method naturally handles cases where the underlying partial differential equation becomes convection dominated or degenerate. This work will demonstrate how a variety of two-dimensional valuation problems can all be solved using the same approach. The generality of the approach is in part due to the fact that changes caused by different model specifications are localized.

For convection dominated pricing problems, a compact positive coefficient scheme is developed. The positive coefficient scheme allows accurate solutions of degenerate problems to be obtained with essentially the same computational cost as nondegenerate problems.

The conditions under which finite volume/element methods, when applied to the two-factor option pricing equation, give rise to discretizations with positive coefficients are also outlined in this work. The importance of positive coefficients in numerical schemes is often stressed in the finance literature. Numerical experiments indicate that constructing a mesh which satisfies the positive coefficient condition may not be necessary, and in some cases appears to even be detrimental. As well, it is shown that schemes with negative coefficients due to the discretization of the diffusion term satisfy approximate local maximum and minimum principles as the mesh spacing approaches zero. This finding is of significance since, for arbitrary diffusion tensors, it may not always be possible to construct a positive coefficient discretization for a given set of nodes.

In addition, it is shown that several lattice methods are equivalent to known finite difference/element schemes.

Acknowledgements

I am indebted to my supervisors, Peter Forsyth and Ken Vetzal, without whom this thesis would not have been possible. I would like to thank Ken and Peter for their patience and for tolerating my stubbornness. I am especially grateful to Peter for the fact that he was always available to answer a question or for an impromptu meeting. I would like to thank (I think) Bruce Simpson for suggesting to Peter that he take me on as a Master's student. I would also like to express my gratitude to Bruce for many productive discussions.

I would like to thank my external examiner, Jeff Dewynne, and my committee, George Labahn, Wei-Pai Tang and Siv Sivaloganathan for helpful suggestions. I am also grateful to Phelim Boyle, Adam Kolkiewicz and Ken Seng Tan for the help that they provided along the way.

Finally, I would like to thank my friends and family, especially my parents, for all the support that they have given me.

Contents

1	Introduction	1
1.1	Contributions	7
1.2	Outline	8
2	Pricing Problems	9
2.1	Some Properties of Option Pricing PDEs	9
2.2	Two Asset Options	16
2.2.1	Puts on the Worst of Two Assets	18
2.2.2	Calls on the Maximum of Two Assets	19
2.2.3	Digital Call Options	20
2.2.4	Discrete Double Barrier Calls on the Maximum of Two Assets	20
2.3	Asian Options	21
2.4	Convertible Bonds	24
3	Discretization	28
3.1	Nonconservative Discretization of Diffusion	34
3.2	Relationship to a Galerkin Discretization	37

3.3	Summary	39
4	Discretization Analysis: Convection	40
4.1	Compact Positive Coefficient Scheme	44
4.2	Summary	51
5	Discretization Analysis: Diffusion	52
5.1	Constant Coefficients	53
5.2	Relation to Previous Work in Finance	57
5.3	Nonconstant Coefficients	60
5.3.1	Boundaries	64
5.4	Approximate Local Maximum and Minimum Principles	65
5.5	Lattices as Finite Differences/Elements	69
5.5.1	The Binomial Method	69
5.5.2	2-D Positive Probability Lattice Schemes	72
5.6	Summary	79
6	Results	81
6.1	Part I	83
6.1.1	Worst of Two Asset Options	83
6.1.2	Asian Options	87
6.1.3	Convertible Bonds	90
6.2	Part II	96
6.3	Summary	109

7 Conclusions	111
7.1 Suggestions for Future Research	115
Bibliography	116

List of Tables

2.1	Symbol definitions.	25
6.1	European put options on the worst of two assets computed on successively finer irregular meshes using the modified van Leer flux limiter.	85
6.2	American put options on the worst of two assets computed on successively finer regular meshes using the modified van Leer flux limiter.	87
6.3	American put options on the worst of two assets computed on irregular meshes using the modified van Leer flux limiter.	88
6.4	American put options on the worst of two assets computed on irregular meshes using central weighting.	88
6.5	European fixed strike Asian call options computed using the van Leer flux limiter on successively finer meshes.	90
6.6	Convertible bond parameter values.	93
6.7	Values of a ten year convertible bond calculated on successively finer irregular meshes using centroid control volumes and central weighting.	93

6.8	Values of a ten year convertible bond calculated on successively finer irregular meshes using centroid control volumes and the modified van Leer flux limiter.	94
6.9	Values of a ten year convertible bond calculated on successively finer irregular meshes using perpendicular bisector control volumes and central weighting.	95
6.10	Values of a ten year convertible bond calculated on successively finer irregular meshes using perpendicular bisector control volumes and the modified van Leer flux limiter.	96
6.11	Values of European call options on the maximum of two assets computed on successively finer regular meshes and positive coefficient meshes constructed using edge swapping.	99

List of Figures

3.1	An example of a control volume constructed using centroids on a triangulated domain.	29
3.2	A boundary control volume.	31
3.3	A self-intersecting boundary control volume.	32
3.4	Finite element/volume triangle with control volume.	34
5.1	A regular triangular mesh.	55
5.2	Edge swapping.	56
5.3	A unique positive coefficient Delaunay triangulation for $\mathbf{D} = \mathbf{I}$	60
5.4	A unique positive coefficient triangulation.	61
5.5	Two regions of nodes with different diffusion tensors.	61
5.6	Boundary triangle with boundary edge parallel to the x axis.	64
5.7	A two period binomial process.	70
5.8	Quadrilateral finite element mesh when correlation is positive. . . .	76
5.9	Transforming a quadrilateral in x, y to a standard square in x', y' . .	76
6.1	Perpendicular bisector and centroid control volumes.	83
6.2	An irregular triangular mesh.	84

6.3	Example of mesh refinement.	84
6.4	A regular triangular mesh.	85
6.5	Values for an American put option on the worst of two assets computed on an irregular mesh using the modified van Leer flux limiter.	89
6.6	European fixed strike Asian call option calculated using central weighting.	91
6.7	European fixed strike Asian call option calculated using the modified van Leer flux limiter.	92
6.8	Ten year convertible bond values computed using centroid control volumes and the modified van Leer limiter.	94
6.9	A positive coefficient mesh constructed using edge swapping.	97
6.10	Level curves of values and deltas of a European call option on the maximum of two assets computed using a regular mesh with 6724 nodes.	100
6.11	Values and deltas of a European call option on the maximum of two assets computed using a positive coefficient edge-swapped mesh with 6724 nodes.	100
6.12	Values and deltas of a European call option on the maximum of two assets computed using a positive coefficient edge-swapped mesh with 26569 nodes.	101
6.13	Level curves of values and values at nodes for a European call option on the maximum of two assets computed using a positive coefficient edge-swapped mesh with 6724 nodes.	101

6.14	A positive coefficient rotated mesh.	103
6.15	Values and deltas of a European call option on the maximum of two assets computed using a positive coefficient rotated mesh with 6614 nodes.	103
6.16	Level curves of deltas and deltas at nodes for a European call option on the maximum of two assets computed using a positive coefficient rotated mesh with 6614 nodes.	104
6.17	Level curves of values and deltas of a digital call option computed using a regular mesh with 26569 nodes.	106
6.18	Values and deltas of a digital call option computed using a positive coefficient edge-swapped mesh with 26569 nodes.	106
6.19	Values and deltas of a digital call option computed using a positive coefficient rotated mesh with 26455 nodes.	107
6.20	Level curves of values and deltas of a European discrete double bar- rier call option on the maximum of two assets computed using a regular mesh with 26569 nodes.	108
6.21	Values and deltas of a European discrete double barrier call option on the maximum of two assets computed using a positive coefficient edge-swapped mesh with 26569 nodes.	108
6.22	Values and deltas of a European discrete double barrier call option on the maximum of two assets computed using a positive coefficient rotated mesh with 26455 nodes.	109

Chapter 1

Introduction

The gross market value of outstanding over-the-counter derivative securities at the end of 1998 has been estimated to be roughly 3.2 trillion U.S. dollars [28]. It goes without saying that the formulation of sound pricing models and the development of accurate numerical techniques to solve pricing models in the absence of closed-form solutions are important.

Options are a type of derivative security. Perhaps the most common numerical methods used to model the value of options are lattice methods. Although it has been argued (see [38, 76]) that lattice methods are in fact finite difference schemes, it appears that it is common in finance to consider finite difference schemes and lattice schemes as qualitatively different. That is, finite difference schemes are discretizations of continuous models while lattice methods are implementations of discrete models. For example, the following passage is taken from [63] (p. 4):

It is well known that the price of the option can be computed by

solving a second-order partial differential equation (PDE). The binomial pricing method can be interpreted as a finite-difference approximation to this PDE, but it is actually more helpful to think of the binomial pricing method as being an *exact* calculation relative to a discrete-time discrete-state Markov process which approximates the log-price process.

The binomial method [19] will be shown to be equivalent to an explicit finite difference scheme in Section 5.5.1.

It is undeniable that lattice methods are simple and efficient techniques for basic pricing problems. One can implement such lattice schemes without knowledge of complex mathematics. However, for more complex option pricing problems, the simplicity of lattice methods breaks down. Evidence of this is provided by the myriad of papers devising lattice schemes for pricing, for example, Asian and barrier options.

In [81] it is shown that certain lattice methods for the pricing of Asian options have a greater computational complexity than standard finite difference/volume approaches. Furthermore, in [31] it is shown that a known lattice scheme for the pricing of Asian options does not converge. It appears that for complicated option pricing problems, viewing lattice schemes as models (even implicitly) and not as numerical techniques is no longer constructive and can be detrimental.

The primary purpose of this work is to develop a general discretization for the solution of two-factor PDE option pricing models. By general, we mean that the discretization can be used to price different types of options with little or no modification. Such an approach differs from lattice techniques, where different lattice schemes are often devised to price different types of options.

Two-factor option pricing problems will typically have one or more of the fol-

lowing characteristics:

- Accurate solutions are required only in a small region.
- Constraints on the solution, such as *barriers*, which can cause irregular boundary geometries (see [57]).
- The underlying convection diffusion equation can become convection dominated or degenerate.
- The treatment of boundary conditions is complicated by the fact that the PDE can degenerate to a first-order hyperbolic equation normal to the boundary.

Consequently, to handle all of the above situations we will use a finite volume approach defined on triangles to discretize the PDE model.

The finite volume discretization supports the use of irregular triangular meshes and the use of perpendicular bisector or centroid control volumes. There is some evidence to suggest that, in some circumstances, centroid control volumes can deteriorate accuracy [6] (which will be investigated in this work). Unlike the orthogonal grids that are typically used in finance, triangular meshes allow one to insert nodes near the region of interest, without introducing nodes elsewhere in the domain. Triangular meshes can also readily support irregular boundary geometries caused by constraints on the solution.

The discretization differs from traditional finite volume approaches because it is nonconservative. Option pricing PDEs are originally in nonconservative form, thus the discretization can be applied without manipulating the original PDEs. The nonconservative discretization often simplifies the treatment of boundaries by

allowing one to solve the original equation on portions of the boundary (typically where the value of an underlying factor approaches zero). In such cases, the need for applying boundary conditions (there are often no explicit boundary conditions for pricing problems) on those portions of the boundary is eliminated.

Option pricing models may become convection dominated or degenerate. For example, some path-dependent option pricing models are degenerate because the diffusion tensor is singular. To ensure the generality of the discretization, a positive coefficient scheme is developed for the convective terms in order to ensure that discrete local maximum and minimum principles hold. The positive coefficient scheme is defined on irregular triangular meshes and uses a modified van Leer flux limiter (a previous work [81] developed a total variation diminishing (TVD) scheme for non-uniform one-dimensional grids using the modified van Leer limiter). The positive coefficient scheme is compact, that is, the Jacobian has the same nonzero structure as would result from using a centrally weighted scheme. Furthermore, diffusion in the underlying PDE can be used to reduce the amount of augmenting diffusion introduced by the scheme.

The PDE for the fair market value of an option is sometimes deduced via dynamic programming principles, and hence this results in a Hamilton-Jacobi-Bellman (HJB) equation. For a rigorous account of existence/uniqueness issues associated with solutions of HJB equations, we refer the reader to [20, 27]. The notion of viscosity solutions is used in the absence of classical solutions for HJB equations. Viscosity solutions of Hamilton-Jacobi equations are analogous to entropy condition satisfying solutions of hyperbolic conservation laws [56]. Note that Hamilton-Jacobi

equations are not in conservative form. In the one-dimensional case, it has been shown that monotone numerical schemes will converge to the viscosity solution for a first-order Hamilton-Jacobi equation [21] and the entropy condition satisfying solution for a hyperbolic conservation law [37]. For detailed accounts of the theory of entropy condition satisfying solutions of hyperbolic conservation laws, we refer the reader to [50], [66] and [48]. Recently the use of positive coefficient schemes for first-order Hamilton-Jacobi equations has been demonstrated in [7]. In view of the success of finite volume methods for conservative problems, it appears that it would be a natural extension to use such approaches for the discretization of the Hamilton-Jacobi-type option pricing equations. Although we are dealing with essentially linear equations (weak nonlinearity is introduced by American-type constraints), we expect that the finite volume method formulated in this work can be generalized in a straightforward fashion to more complex nonlinear PDEs arising in finance.

In the above, the importance of discretizing the convective term in convection dominated option pricing problems so that (ignoring any negative contributions made by discretizing the diffusion term) the coefficients are ensured to be positive was discussed. In this work, the finite volume approach that is developed will also be used to investigate the effect of negative contributions to coefficients from the discretization of the diffusion term in the option pricing PDE. Specifically, we will determine if such contributions result in a reduction in the quality of solutions for several sample pricing problems.

Finance literature stresses that when the underlying pricing equation is dis-

cretized using an explicit finite difference scheme, the coefficients in the resulting difference equations should be nonnegative [13, 41]. The nonnegativity of coefficients is a sufficient condition for the stability of a consistent explicit scheme [33]. For the one-dimensional diffusion equation (with linear basis functions), nonnegativity is also necessary for stability. Although nonnegative coefficients are only sufficient conditions for the stability of explicit schemes in two dimensions, they are also sufficient conditions for schemes of arbitrary temporal weighting to possess discrete local maximum and minimum principles (see [59, 29, 58, 81]). Of course, there exist two-dimensional and higher dimensional explicit schemes that do not satisfy the positive coefficient condition, but are stable nonetheless [11].

Lattices (explicit finite difference/element schemes) are generally constructed in such a way as to ensure that the coefficients are positive (see [10, 1, 43, 34, 44]). This is perhaps primarily due to the fact that the coefficients are viewed as risk-neutral probabilities [26]. The following quote from [34] (p. 14) demonstrates the view of coefficients as probabilities and highlights the emphasis placed on positive coefficients (probabilities):

... the ... model may produce a negative probability which may cause the model to explode, ...

It is not always possible to construct a finite difference or finite element scheme with nonnegative coefficients. Whether such a nonnegative scheme can be constructed depends on the form of the diffusion tensor. That is, the diffusion tensor must be constant so that, effectively, a rotation can be performed to remove any cross-partial terms. If a scheme produces negative coefficients then one cannot generally ensure

that discrete local maximum and minimum principles hold. Hence, the scheme can possibly introduce spurious oscillations into the solution. In Section 5.5.2 it will be shown that the two-dimensional positive probability lattice scheme in [34] has positive coefficients, in part, because it is in fact an explicit finite element scheme which uses *skewed* or nonorthogonal quadrilateral meshes.

1.1 Contributions

The main contributions of this work are:

- The formulation of a nonconservative finite volume discretization defined on irregular triangular meshes that supports both perpendicular bisector and centroid control volumes. The discretization allows us to work directly with the original pricing equation.
- The development of a compact positive coefficient scheme for convection. The scheme maximizes the use of central weighting in order to minimize the amount of augmenting diffusion that is required to obtain positive coefficients. To the best of our knowledge, a high-order positive coefficient scheme that is compact has not been previously developed.
- Demonstrating that several lattice schemes are equivalent to known finite difference/element schemes.
- Establishing that high quality solutions can be obtained when the discretization of diffusion produces negative coefficients. Moreover, demonstrating that

under certain criteria, solutions of poor quality may result when the discretization of diffusion ensures positive coefficients.

- Demonstrating that a wide array of option models, including models of two-asset options, path-dependent options and convertible bonds, can all be solved within the same numerical framework. The discretization makes no assumptions about the underlying factors or the processes which they follow (except that we only consider processes that do not permit negative factor values).

1.2 Outline

The outline of this work is as follows. Chapter 2 presents the underlying option pricing PDE and discusses some of its properties. Chapter 2 also outlines the sample pricing problems which will be used for the numerical experiments in this work. The nonconservative finite volume discretization is derived in Chapter 3. In Chapter 4 a positive coefficient scheme for convection is developed. Chapter 5 contains a discussion of when one can ensure that the finite volume discretization of diffusion will result in positive coefficients. It is shown in Chapter 5 that discretizations of the diffusion term which produce negative coefficients will approximately satisfy discrete local maximum and minimum principles. It is also shown in Chapter 5 that several lattice methods are equivalent to known finite difference/element schemes. The numerical results are contained in Chapter 6. Concluding remarks and suggestions for future research are in Chapter 7.

Chapter 2

Pricing Problems

Before outlining the pricing problems that will be examined in the numerical experiments of this work, this chapter will introduce the form of the underlying PDE for the pricing problems and describes some situations in which the underlying PDE becomes convection dominated or degenerate. In addition, an explanation is provided for the observation that the underlying PDE can often be solved on portions of the boundary.

2.1 Some Properties of Option Pricing PDEs

The value of many financial options is conveniently modelled in terms of two factors [77]. For example, suppose that the value of an option U is a function of two stochastic variables (x_1 and x_2) and time (t^*). The time evolutions of x_1 and x_2

are given by the stochastic differential equations (SDEs):

$$dx_1 = a_1(x_1)dt^* + b_1(x_1)dW_1 \quad (2.1)$$

and

$$dx_2 = a_2(x_2)dt^* + b_2(x_2)dW_2, \quad (2.2)$$

where W_1 and W_2 are Wiener processes (see [55] for a discussion of SDEs).¹ The two Wiener processes are related through their correlation coefficient ρ .

Based on the contingent claims analysis developed by Black and Scholes [8], and Merton [52], a partial differential equation (PDE) for the price of an option which is a function of the above two factors and t^* can be derived. Readers unfamiliar with this theory can find readable accounts in [40, 77]. The final form of the PDE for the option price $U = U(x_1, x_2, t^*)$ is

$$U_{t^*} - \mathbf{V} \cdot \nabla U + (\mathbf{D}\nabla) \cdot \nabla U - rU + P = 0, \quad (2.3)$$

where $\mathbf{D} = \mathbf{D}(x_1, x_2)$ is the diffusion tensor, $\mathbf{V} = \mathbf{V}(x_1, x_2)$ is the velocity tensor, r is an interest rate and P is a penalty term which is used to enforce constraints (see [80]). Some of the possible constraints include early-exercise features (American options), conversion provisions and call provisions. Effectively, P adds or subtracts value in order to ensure that the constraints are met. We will be solving equation

¹Note that a_1 and b_1 can also be dependent on t^* .

(2.3) backwards in time from option maturity ($t^* = T$) to the present ($t^* = 0$). Consequently, by letting $t = T - t^*$ we can convert equation (2.3) into the more familiar form:

$$U_t = -\mathbf{V} \cdot \nabla U + (\mathbf{D}\nabla) \cdot \nabla U - rU + P. \quad (2.4)$$

Equation (2.4) is simply the two-dimensional convection-diffusion equation along with an exponential decay term due to a discounting effect.

The diffusion tensor \mathbf{D} is symmetric positive semidefinite and is usually a function of the space-like coordinates x_1 and x_2 . Typically, x_1 and x_2 represent quantities such as asset value or interest rate (which are usually constrained to be non-negative). In the examples considered in this work $x_1, x_2 \geq 0$ (before any transformations are made). It then follows that we must have (to ensure the nonnegativity of x_1 and x_2)

$$a_1(x_1) \geq 0 \text{ and } b_1(x_1) \rightarrow 0 \text{ as } x_1 \rightarrow 0, \quad (2.5)$$

and

$$a_2(x_2) \geq 0 \text{ and } b_2(x_2) \rightarrow 0 \text{ as } x_2 \rightarrow 0 \quad (2.6)$$

in equations (2.1) and (2.2) [35].² As a consequence of (2.5) and (2.6), the compu-

²If x_1 is an asset price, then we must have $a_1(x_1) \rightarrow 0$ and $b_1(x_1) \rightarrow 0$ as $x_1 \rightarrow 0$ in order to avoid arbitrage opportunities.

tational domain can be restricted to

$$0 \leq x_1 \leq \infty \text{ and } 0 \leq x_2 \leq \infty.$$

The diffusion tensor has the form [40]

$$\mathbf{D}(x_1, x_2) = \begin{pmatrix} b_1(x_1)^2/2 & b_1(x_1)b_2(x_2)\rho/2 \\ b_1(x_1)b_2(x_2)\rho/2 & b_2(x_2)^2/2 \end{pmatrix}.$$

It therefore follows from (2.5) and (2.6) that

$$\lim_{x_1 \rightarrow 0} \mathbf{D}(x_1, x_2) = \begin{pmatrix} 0 & 0 \\ 0 & b_2(x_2)^2/2 \end{pmatrix}$$

and

$$\lim_{x_2 \rightarrow 0} \mathbf{D}(x_1, x_2) = \begin{pmatrix} b_1(x_1)^2/2 & 0 \\ 0 & 0 \end{pmatrix}.$$

That is, there is no diffusion normal to the boundaries at $x_1 = 0$ or $x_2 = 0$. Note that if the diffusion goes to zero sufficiently fast, then the underlying PDE becomes first-order hyperbolic on those portions of the boundary. Another consequence of the fact that x_1 and x_2 cannot become negative is that only outgoing information is required at $x_1 = 0$ or $x_2 = 0$. That is, at $x_1 = 0$ or $x_2 = 0$

$$\mathbf{V} \cdot \vec{n} \geq 0, \tag{2.7}$$

where \vec{n} is the outward pointing unit normal. In general, if $\mathbf{V} \cdot \vec{n} < 0$ or if there is diffusion normal to a boundary, then a Dirichlet condition must be imposed.

In some situations there may be no diffusion in one of the coordinate directions throughout the domain. This is the case for some options where the payoff is a function of the continuously monitored average value of the underlying asset (Asian options). In such situations equation (2.4) is degenerate (with all the usual difficulties).

We can also see that equation (2.4) is in nonconservative form. This is not an artifact of some manipulation of the PDE, but is a direct consequence of the contingent claims analysis. Since \mathbf{V} is generally an arbitrary function of the coordinates, it is advantageous in terms of the discretization to leave the convective term in nonconservative form.

Alternatively, we can write equation (2.4) as

$$U_t = (-\mathbf{V} - (\nabla' \mathbf{D})') \cdot \nabla U + \nabla \cdot \mathbf{D} \nabla U - rU + P, \quad (2.8)$$

where ∇' is the transpose of ∇ . The diffusion term in (2.8) is in standard conservative form, which is convenient for integrating by parts in finite element discretizations. This approach was used in [32] and [80]. However, there are several disadvantages to this approach. The differentiation of the diffusion tensor may introduce singularities into the effective velocity $(-\mathbf{V} - (\nabla' \mathbf{D})')$. Furthermore, for some option models a PDE will be solved at the boundaries. In such cases the original nonconservative equation (2.4) must be discretized at the boundaries in order to ensure the correct flow of information. Discretizing (2.8) for the interior

domain and (2.4) at the boundary complicates software development (see [32]) and mesh construction.

As discussed above, in many models it is common to have no diffusion normal to portions of the boundary. This often leads to misconceptions about what boundary conditions are required at those points on the boundary. In order to clarify this matter, consider the following one-dimensional case. Under the assumption of an interest rate process of the form

$$dr = a(b - r)dt + \sigma_r r^c dW_r \quad (2.9)$$

where a, b, c and σ_r are positive parameters, and dW_r is the increment of a Wiener process (W_r), then, by using standard methods (see [74]), the PDE for the value of a bond $U(r, t)$ where r follows the process (2.9) is given by

$$U_t = \frac{1}{2} \sigma_r^2 r^{2c} U_{rr} + (a(b - r) - \lambda \sigma_r r^c) U_r - rU, \quad (2.10)$$

where $\lambda = \lambda(r, t)$ is the market price of interest rate risk. Equation (2.10) is to be solved on the domain $r \geq 0$.

If U represents the value of a bond paying fixed coupons, then U tends to zero as $r \rightarrow \infty$. Consequently, the boundary condition imposed as $r \rightarrow \infty$ is $\lim_{r \rightarrow \infty} U(r, t) = 0$, $t > 0$. Now, assuming that the diffusion goes to zero sufficiently fast (for example, if $c = \frac{1}{2}$ then we require that $2ab \geq \sigma_r^2$ [18]) and that $\lambda(r, t) \sigma_r r^c \rightarrow$

0 as $r \rightarrow 0$, then taking the limit of equation (2.10) as $r \rightarrow 0$ gives us

$$U_t = abU_r. \quad (2.11)$$

Since $a, b > 0$, equation (2.11) degenerates into a first-order hyperbolic equation, with domain of dependence consisting of points in $r \geq 0$. Hence, no boundary condition is required at $r = 0$. In fact, imposing any type of condition other than equation (2.11) at $r = 0$ would be inappropriate. Consequently, the discrete equations should not require any conditions at $r = 0$. Such a practice is common in computational fluid dynamics and it is known in finance [69].

To solve equation (2.10) numerically, the infinite domain is truncated to produce a finite computational domain. An artificial condition determined by asymptotic analysis or financial reasoning ($U = 0$ in this case as $r \rightarrow \infty$) must then be imposed at the maximum interest rate (r_{max}) on the computational domain. In practice, we have found that if r_{max} is sufficiently large, then the artificial condition has a negligible effect on the solution in the region of interest (see also [4]). The same cannot be said if an inappropriate boundary condition is imposed at $r = 0$ since this may be near the region of interest.

Returning to the general case, for boundaries at $x_1 = 0$ and $x_2 = 0$, equation (2.7) holds and there will be no diffusion normal to the boundaries for the two-factor models considered in this work (before any transformations). Thus, no conditions other than the original PDE need to be imposed along $x_1 = 0$ and $x_2 = 0$.

Options can be broadly classified as European or American. A European option can only be exercised at maturity, while an American option can be exercised at

any time during the life of the option. Thus, the value of an American option can never fall below the payoff (terminal) value. Let $g(x_1, x_2, t)$ denote the payoff function, then the American option pricing problem can be formally stated as the following linear complementarity problem [77]

$$\begin{aligned} \left(\frac{\partial}{\partial t} + \mathbf{V} \cdot \nabla - (\mathbf{D}\nabla) \cdot \nabla + r\right)U(x_1, x_2, t) &\geq 0, \\ U(x_1, x_2, t) - g(x_1, x_2, t) &\geq 0, \\ (U(x_1, x_2, t) - g(x_1, x_2, t))\left(\frac{\partial}{\partial t} + \mathbf{V} \cdot \nabla - (\mathbf{D}\nabla) \cdot \nabla + r\right)U(x_1, x_2, t) &= 0 \end{aligned} \quad (2.12)$$

subject to the conditions that $U(x_1, x_2, t)$ and $\nabla U(x_1, x_2, t)$ are continuous (if we restate the problem as a variational inequality, we require that $U \in H^1$ [23]). Equivalently, we can state the American pricing problem as equation (2.4) where the penalty term is defined such that

$$P \text{ is } \begin{cases} = 0 & \text{if } U(x_1, x_2, t) \geq g(x_1, x_2, t), \\ > 0 & \text{otherwise.} \end{cases}$$

The discrete definition of P will be given in Chapter 3. Other continuously applied constraints on the solution, such as, convertible bond call and conversion provisions, and time-varying barriers can be implemented by using a penalty term.

2.2 Two Asset Options

We will now define several two-asset option pricing problems. A two-asset option such as a put on the minimum of two assets gives the holder the right to sell the

cheaper asset at a specified price (the exercise price). A call on the maximum of two assets allows the holder to buy the more expensive asset for a specified price.

Let the asset price processes be

$$dS_1 = \mu_1 S_1 dt^* + \sigma_{S_1} S_1 dW_{S_1} \quad (2.13)$$

and

$$dS_2 = \mu_2 S_2 dt^* + \sigma_{S_2} S_2 dW_{S_2}, \quad (2.14)$$

where μ_1 and μ_2 are expected rates of return, σ_{S_1} and σ_{S_2} are volatilities, and W_{S_1} and W_{S_2} are Wiener processes. Defining the gradient operator as

$$\nabla = \begin{pmatrix} \frac{\partial}{\partial S_1} \\ \frac{\partial}{\partial S_2} \end{pmatrix},$$

then the price of an option based on two underlying assets, $U(S_1, S_2, t)$, has the form of equation (2.4) with (by the usual no arbitrage arguments [8, 52])

$$\mathbf{D} = \frac{1}{2} \begin{pmatrix} \sigma_{S_1}^2 S_1^2 & \rho \sigma_{S_1} \sigma_{S_2} S_1 S_2 \\ \rho \sigma_{S_1} \sigma_{S_2} S_1 S_2 & \sigma_{S_2}^2 S_2^2 \end{pmatrix}, \quad (2.15)$$

where ρ is the coefficient of correlation, and

$$\mathbf{V} = - \begin{pmatrix} r S_1 \\ r S_2 \end{pmatrix}. \quad (2.16)$$

2.2.1 Puts on the Worst of Two Assets

The payoff function (terminal condition) for a put on the worst of two assets is

$$U(S_1, S_2, 0) = \max(K - \min(S_1, S_2), 0), \quad (2.17)$$

where K is the exercise price. For an American put on the worst of two assets the early-exercise constraint is

$$U(S_1, S_2, t) \geq \max(K - \min(S_1, S_2), 0). \quad (2.18)$$

The boundary conditions are

$$\frac{\partial U}{\partial t} = \frac{1}{2}\sigma_{S_2}^2 S_2^2 \frac{\partial^2 U}{\partial S_2^2} + rS_2 \frac{\partial U}{\partial S_2} - rU + P \text{ as } S_1 \rightarrow 0, \quad (2.19)$$

$$\frac{\partial U}{\partial t} = \frac{1}{2}\sigma_{S_1}^2 S_1^2 \frac{\partial^2 U}{\partial S_1^2} + rS_1 \frac{\partial U}{\partial S_1} - rU + P \text{ as } S_2 \rightarrow 0, \quad (2.20)$$

$$\frac{\partial U}{\partial t} = \frac{1}{2}\sigma_{S_1}^2 S_1^2 \frac{\partial^2 U}{\partial S_1^2} + rS_1 \frac{\partial U}{\partial S_1} - rU + P \text{ as } S_2 \rightarrow \infty, S_1 \neq S_2, \quad (2.21)$$

$$\frac{\partial U}{\partial t} = \frac{1}{2}\sigma_{S_2}^2 S_2^2 \frac{\partial^2 U}{\partial S_2^2} + rS_2 \frac{\partial U}{\partial S_2} - rU + P \text{ as } S_1 \rightarrow \infty, S_1 \neq S_2. \quad (2.22)$$

$$U(S_1, S_2, t) = 0 \text{ as } S_1, S_2 \rightarrow \infty, S_1 = S_2. \quad (2.23)$$

Recalling the discussion in Section 2.1, conditions (2.19) and (2.20) are the limits of the underlying PDE as $S_1 \rightarrow 0$ and $S_2 \rightarrow 0$, respectively. Condition (2.21) is deduced from the fact that for fixed S_1 , the payoff becomes independent of S_2 as $S_2 \rightarrow \infty$. Thus, all derivatives with respect to S_2 vanish as $S_2 \rightarrow \infty$. Condition (2.22) is derived in the same fashion. The Dirichlet condition (2.23) follows from

the payoff function (2.17). Note that the resulting pricing problem is well-posed since conditions (2.19) to (2.22) each have a domain of dependence that falls on the boundary.

2.2.2 Calls on the Maximum of Two Assets

The payoff function for a European call on the maximum of two assets is

$$U(S_1, S_2, 0) = \max(\max(S_1, S_2) - K, 0). \quad (2.24)$$

The boundary conditions are

$$U_t = \frac{1}{2}\sigma_{S_2}^2 S_2^2 \frac{\partial^2 U}{\partial S_2^2} + rS_2 \frac{\partial U}{\partial S_2} - rU \text{ as } S_1 \rightarrow 0, \quad (2.25)$$

$$U_t = \frac{1}{2}\sigma_{S_1}^2 S_1^2 \frac{\partial^2 U}{\partial S_1^2} + rS_1 \frac{\partial U}{\partial S_1} - rU \text{ as } S_2 \rightarrow 0, \quad (2.26)$$

$$U(S_1, S_2, t) = S_1 - Ke^{-rt} \text{ as } S_1 \rightarrow \infty, \quad (2.27)$$

$$U(S_1, S_2, t) = S_2 - Ke^{-rt} \text{ as } S_2 \rightarrow \infty. \quad (2.28)$$

Note that conditions (2.25) and (2.26) are the limits of the underlying equation as $S_1 \rightarrow 0$ and $S_2 \rightarrow 0$, respectively.

2.2.3 Digital Call Options

The payoff function for a European digital call on two assets is

$$U(S_1, S_2, 0) = \begin{cases} 1 & \text{if } S_1 \geq K \text{ and } S_2 \geq K \\ 0 & \text{otherwise.} \end{cases} \quad (2.29)$$

One can impose the following boundary conditions:

$$U(S_1, S_2, t) = 0 \text{ as } S_1 \rightarrow 0, \quad (2.30)$$

$$U(S_1, S_2, t) = 0 \text{ as } S_2 \rightarrow 0, \quad (2.31)$$

$$U_t = \frac{1}{2}\sigma_{S_2}^2 S_2^2 \frac{\partial^2 U}{\partial S_2^2} + rS_2 \frac{\partial U}{\partial S_2} - rU \text{ as } S_1 \rightarrow \infty, S_1 \neq S_2, \quad (2.32)$$

$$U_t = \frac{1}{2}\sigma_{S_1}^2 S_1^2 \frac{\partial^2 U}{\partial S_1^2} + rS_1 \frac{\partial U}{\partial S_1} - rU \text{ as } S_2 \rightarrow \infty, S_1 \neq S_2, \quad (2.33)$$

$$U(S_1, S_2, t) = e^{-rt} \text{ as } S_1, S_2 \rightarrow \infty, S_1 = S_2. \quad (2.34)$$

Condition (2.32) is derived by noting that for fixed S_2 , the payoff (2.29) will be insensitive to small changes in S_1 as $S_1 \rightarrow \infty$. Condition (2.33) is derived using a similar argument. This pricing problem is of interest numerically because of the discontinuous payoff function.

2.2.4 Discrete Double Barrier Calls on the Maximum of Two Assets

The value of a barrier option depends on whether the underlying assets reach a certain price level. For example, the value of a particular European discrete double-

knockout barrier option must satisfy the following constraint

$$U(S_1, S_2, t_+) = \begin{cases} U(S_1, S_2, t_-) & \text{if } H_{lower} < S_1, S_2 < H_{upper} \\ 0 & \text{otherwise} \end{cases} \quad (2.35)$$

where t_+ and t_- are the times just before and after the application (monitoring date) of a barrier, respectively. Also in equation (2.35), H_{lower} and H_{upper} are the lower and upper barriers, respectively. If the option is a call where the payoff depends on the maximum of two assets, then the payoff function is given by equation (2.24) and the boundary conditions are $U(S_1, S_2, t) = 0$ as $S_1 \rightarrow 0$, $S_2 \rightarrow 0$, $S_1 \rightarrow \infty$ or $S_2 \rightarrow \infty$.

The presence of barriers makes such options cheaper than options without barriers. Discrete barrier options are of interest numerically because the discrete application of the barriers will generally introduce discontinuities at each monitoring date.

2.3 Asian Options

Asian options have a payoff which is a function of the average of an underlying asset price. Asian options are often used when the underlying asset is a commodity (for example, oil or aluminum). The use of an average price reduces the option's sensitivity to price changes in the underlying asset near maturity. There continues to be much interest in developing pricing algorithms for both European and American Asian options [47, 42, 62, 5, 81, 15]. Asian option models are of interest numerically because they are degenerate (there is no diffusion in one of the space-like

dimensions).

Assume that the value of an Asian option is a function of the asset price (S), the average price (A) and time. Let the asset price process be

$$dS = \mu S dt + \sigma_S S dW_S, \quad (2.36)$$

where μ is the expected rate of return, σ_S is the volatility and W_S is a Wiener process. The average of the asset price at any time is defined as

$$A = \frac{1}{t^*} \int_0^{t^*} S(\tau) d\tau.$$

By standard arguments, the diffusion and velocity tensors in equation (2.4) are

$$\mathbf{D} = \frac{1}{2} \begin{pmatrix} \sigma_S^2 S^2 & 0 \\ 0 & 0 \end{pmatrix} \quad (2.37)$$

and

$$\mathbf{V} = - \begin{pmatrix} rS \\ \frac{S-A}{T-t} \end{pmatrix}, \quad (2.38)$$

where the gradient operator is defined as

$$\nabla = \begin{pmatrix} \frac{\partial}{\partial S} \\ \frac{\partial}{\partial A} \end{pmatrix}.$$

We have assumed that averaging is continuous, but alternative formulations for

discrete averaging are possible [25, 82].

The payoff function for a fixed-strike Asian call is

$$U(S, A, t = 0) = \max(A - K, 0) \quad (2.39)$$

and the boundary conditions are

$$\frac{\partial U}{\partial t} = -\frac{A}{T-t} \frac{\partial U}{\partial A} - rU \text{ as } S \rightarrow 0, \quad (2.40)$$

$$\frac{\partial U}{\partial t} = \frac{1}{2}\sigma_S^2 S^2 \frac{\partial^2 U}{\partial S^2} + rS \frac{\partial U}{\partial S} + \left(\frac{S}{T-t}\right) \frac{\partial U}{\partial A} - rU \text{ as } A \rightarrow 0, \quad (2.41)$$

$$\frac{\partial U}{\partial t} = \frac{1}{2}\sigma_S^2 S^2 \frac{\partial^2 U}{\partial S^2} + rS \frac{\partial U}{\partial S} + \left(\frac{S-A}{T-t}\right) \frac{\partial U}{\partial A} - rU \text{ as } A \rightarrow \infty, S \neq A, \quad (2.42)$$

$$\frac{\partial U}{\partial t} = \left(\frac{S-A}{T-t}\right) \frac{\partial U}{\partial A} \text{ as } S \rightarrow \infty, S \neq A, \quad (2.43)$$

$$U = A - Ke^{-rt} \text{ as } S, A \rightarrow \infty, S = A. \quad (2.44)$$

With respect to conditions (2.40) to (2.42), note that the domain of dependence of the original PDE is on the interior of the domain and the boundary as $A \rightarrow 0$, $A \rightarrow \infty$ ($A \geq S$) or $S \rightarrow 0$. Condition (2.43) is deduced from the no-arbitrage jump condition for discretely observed Asian options [82], taking the limit as the observation interval tends to zero (the continuous observation limit). Equation (2.44) follows from the fact that if $A = S \rightarrow \infty$, then the option will surely be exercised.

2.4 Convertible Bonds

A convertible bond is a hybrid bond that allows the holder to convert the bond into a specified number of shares, thereby allowing the holder to gain from increases in the stock's value. There will often be a call provision that allows the issuer to force holders to redeem the bonds at a specified price. Firms issue convertible bonds primarily to serve as a form of delayed equity or in order to *sweeten* the debt [54].

The first PDE model for pricing convertible bonds was formulated in [45] and [12]. This model was a one factor model where the interest rate was taken to be constant. A two-factor model which incorporated a stochastic interest rate factor was first proposed in [14]. In these early models the value of the firm was taken to be an underlying factor. Dilution and bankruptcy effects were also incorporated into the model in [14]. In more recent work the value of the stock itself is taken to be an underlying factor and dilution and bankruptcy effects are typically ignored [51, 16, 39].

Consider a model with interest rate process

$$dr = a(b - r)dt + \sigma_r r^c dW_r$$

and stock price process

$$dS = \mu S dt + \sigma_S S dW_S.$$

By the usual no arbitrage arguments we obtain an equation of the form (2.4) with

the gradient operator defined as

$$\nabla = \begin{pmatrix} \frac{\partial}{\partial S} \\ \frac{\partial}{\partial r} \end{pmatrix},$$

and

$$\mathbf{D} = \frac{1}{2} \begin{pmatrix} \sigma_S S^2 & \rho \sigma_S \sigma_r S r^c \\ \rho \sigma_S \sigma_r S r^c & \sigma_r^2 r^{2c} \end{pmatrix} \quad (2.45)$$

and

$$\mathbf{V} = - \begin{pmatrix} rS \\ a(b-r) - \lambda \sigma_r r^c \end{pmatrix}, \quad (2.46)$$

(see [77]) where the symbols are defined in Table 2.1. Note that a, b are assumed to be non-negative.

Symbol	Definition
σ_r	interest rate volatility
a	reversion rate of r
b	reversion level of r
c	distributional parameter for r
λ	market price of interest rate risk
ρ	coefficient of correlation

Table 2.1: *Symbol definitions.*

It is assumed that the holder of the bond can convert the bond at any time and receive ω (the conversion ratio) shares of the stock. We also assume that the bond

is continuously callable prior to maturity. That is, the issuer can buy back the bond at any time before maturity for the call price (C_p). Dilution and bankruptcy effects will not be considered.

Equation (2.4) is solved subject to the terminal condition

$$U(S, r, t = 0) = \max(F, \omega S), \quad (2.47)$$

where F is the face value of the bond. Equation (2.47) simply states that the holder will elect to receive the face value of the bond or ω shares, whichever is worth more.

The conversion and call provisions introduce two constraints on the value of the convertible. First, if the holder can convert the bond into shares at any time, then the value of the bond cannot be less than the conversion value. Hence we have

$$U(S, r, t) \geq \omega S. \quad (2.48)$$

Secondly, the call provision prevents the value of the bond from exceeding the call price since it is optimal for the issuer to call the bonds as soon as their value equals the call price. Thus, the constraint introduced by the call provision is

$$U(S, r, t) \leq C_p. \quad (2.49)$$

The boundary conditions for the convertible bond are

$$\frac{\partial U}{\partial t} = \frac{1}{2} \sigma_r^2 r^{2c} \frac{\partial^2 U}{\partial r^2} + (a(b - r) - \lambda \sigma_r r^c) \frac{\partial U}{\partial r} - rU + P \text{ as } S \rightarrow 0, \quad (2.50)$$

$$\frac{\partial U}{\partial t} = \frac{1}{2}\sigma_S^2 S^2 \frac{\partial^2 U}{\partial S^2} + ab \frac{\partial U}{\partial r} + P \text{ as } r \rightarrow 0, \quad (2.51)$$

$$\frac{\partial U}{\partial t} = \frac{1}{2}\sigma_S^2 S^2 \frac{\partial^2 U}{\partial S^2} + rS \frac{\partial U}{\partial S} - rU + P \text{ as } r \rightarrow \infty, \quad (2.52)$$

$$U(S, r, t) = C_p \text{ as } S \rightarrow \infty. \quad (2.53)$$

Note that equations (2.50) and (2.51) are simply the limiting forms of the PDE as $S \rightarrow 0$ and $r \rightarrow 0$, respectively. Equation (2.52) is deduced from the fact that the value of a straight bond tends to zero as $r \rightarrow \infty$. Thus, the convertible bond derives value (which is capped by the call price) only from the conversion provision (which depends upon S) as $r \rightarrow \infty$. Equation (2.53) follows from the call provision. That is, $\omega S \rightarrow \infty$ (the conversion value) as $S \rightarrow \infty$, but the value of the bond cannot exceed the call price.

Convertible bonds with discrete coupon payments can be modelled using

$$U(S, r, t^+) = U(S, r, t^-) + C, \quad (2.54)$$

where C is the dollar amount of the coupon. In equation (2.54), t^+ and t^- are the times the instant before and after the coupon payment (recall that $t = T - t^*$ is moving backwards in real time).

The above model is of interest numerically for several reasons. First, we have Dirichlet conditions and differential equations on the boundaries. Secondly, there are two algebraic constraints on the solution. Since these constraints are continuously applied, they are mathematically a type of American constraint. Finally, the underlying PDE becomes degenerate on portions of the boundary.

Chapter 3

Discretization

In this chapter we will derive a nonconservative finite volume discretization for PDE (2.4). We will first consider the treatment of interior nodes before incorporating nodes on the boundary into the discretization. The discretization is defined on a two-dimensional computational domain Ω which is tiled by triangles. The nonconservative discretization will allow us to work with the original PDE directly, that is, without converting it into conservative form. Working with the original PDE will permit us to often solve the PDE on portions of the boundary (such boundaries can be handled by the nonconservative discretization without requiring special treatment).

It will also be shown that the nonconservative finite volume discretization of the diffusion term in PDE (2.4) is equivalent to a nonconservative Galerkin finite element approach which uses a low order quadrature rule. In the case of constant D and V , the finite volume discretization using central weighting is equivalent to a standard finite element discretization with mass lumping.

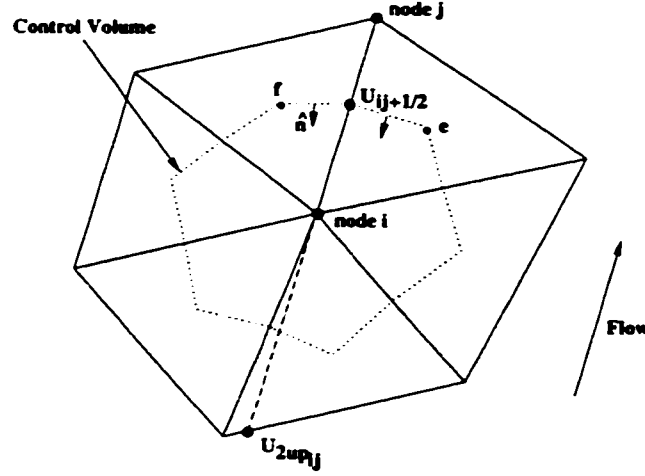


Figure 3.1: An example of a control volume constructed using centroids on a triangulated domain. Points e and f are the centroids of their respective triangles. The face (line segments from e to f) passes through the middle of the edge connecting nodes i and j . U_{2upij} is the second upstream value.

For a node i in the interior of the computational domain a control volume can be constructed by connecting the midpoints of triangle edges to the triangle centroids (refer to Figure 3.1), or using perpendicular bisectors (see [53]). A centroid construction can always be carried out, but the perpendicular bisector construction requires a Delaunay triangulation [6]. Integrating equation (2.4) over the finite volume FV_i gives

$$\int_{FV_i} U_t d\Omega = - \int_{FV_i} \mathbf{V} \cdot \nabla U d\Omega + \int_{FV_i} (\mathbf{D}\nabla) \cdot \nabla U d\Omega - \int_{FV_i} \tau U d\Omega + \int_{FV_i} P d\Omega. \quad (3.1)$$

Let $U_i^{n+1} = U(x_i, y_i, t^{n+1})$. Using fully implicit time stepping for ease of exposition, the following approximations are used for the terms in equation (3.1)

$$\int_{FV_i} U_t d\Omega \approx A_i \left(\frac{U_i^{n+1} - U_i^n}{\Delta t} \right), \quad (3.2)$$

$$- \left(\int_{FV_i} \mathbf{V} \cdot \nabla U d\Omega \right)^{n+1} \approx -\mathbf{V}_i \cdot \oint_{\partial FV_i} U^{n+1} \bar{\mathbf{n}} d\Gamma, \quad (3.3)$$

$$\left(\int_{FV_i} (\mathbf{D} \nabla) \cdot \nabla U d\Omega \right)^{n+1} \approx \oint_{\partial FV_i} (\mathbf{D}_i \nabla U^{n+1}) \cdot \bar{\mathbf{n}} d\Gamma, \quad (3.4)$$

$$\left(\int_{FV_i} \tau U d\Omega \right)^{n+1} \approx A_i \tau_i U_i^{n+1}, \quad (3.5)$$

$$\left(\int_{FV_i} P d\Omega \right)^{n+1} \approx p_i^{n+1}, \quad (3.6)$$

where A_i denotes the area of the finite volume FV_i , Δt is the time step size, Ω_i is the set of nodes that neighbour node i , $\bar{\mathbf{n}}$ is the outward pointing unit normal and p_i is the discrete form of the penalty function used to enforce Dirichlet conditions and constraints such as those introduced by American early-exercise features, and by call and conversion provisions. Note that in equations (3.3) to (3.5) the integrals have been approximated by evaluating terms which depend on the space-like variables at node i . This will be advantageous when dealing with boundary nodes.

Equation (3.3) is further discretized in the following manner. Let

$$\bar{L}_{ij} = \int_e^f \hat{\mathbf{n}} d\Gamma$$

where e and f are the endpoints of the face, and $\hat{\mathbf{n}}$ is an inward pointing unit normal to the face (refer to Figure 3.1) between nodes i and j . If $U_{ij+\frac{1}{2}}^{n+1}$ is the value at the control volume face separating nodes i and j , then equation (3.3) becomes

$$-\mathbf{V}_i \cdot \oint_{\partial FV_i} U^{n+1} \bar{\mathbf{n}} d\Gamma \approx \mathbf{V}_i \cdot \sum_{j \in \Omega_i} \bar{L}_{ij} U_{ij+\frac{1}{2}}^{n+1}. \quad (3.7)$$

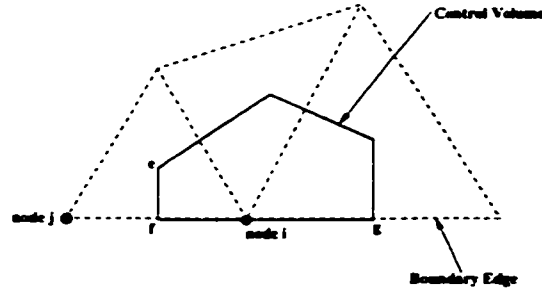


Figure 3.2: A boundary control volume. The segment (e, f) is the face between nodes i and j . The segment (f, g) is the portion of the boundary edge included in order to construct a closed control volume.

The calculation of $U_{ij+\frac{1}{2}}^{n+1}$ will be discussed in Chapter 4.

Let N_i be the usual C^0 Lagrange basis functions defined on triangles where

$$\begin{aligned} N_i &= 1 \text{ at node } i, \\ &= 0 \text{ at all other nodes,} \\ \sum_j N_j &= 1 \text{ everywhere in the solution domain.} \end{aligned}$$

Also, let $U^{n+1} = \sum_j N_j U_j^{n+1}$. Then

$$\oint_{\partial FV_i} (\mathbf{D}_i \nabla U^{n+1}) \cdot \vec{n} d\Gamma \approx \sum_{j \in \Omega_i} - \int_{\Omega} \nabla' N_i \mathbf{D}_i \nabla N_j d\Omega (U_j^{n+1} - U_i^{n+1}) \quad (3.8)$$

in equation (3.4). The derivation of approximation (3.8) is contained in Section 3.1.

We have assumed that node i is an interior node. For the case where node i is on the boundary, consider the boundary control volume represented in Figure 3.2. The construction of boundary control volumes is similar to that of control volumes

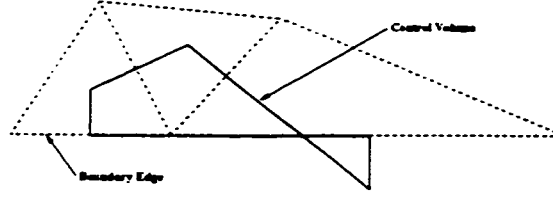


Figure 3.3: A self-intersecting boundary control volume.

for interior nodes, except that a portion of the boundary edge must be included in order to construct a closed volume. That is, in Figure 3.2 the face denoted by the boundary segment (f, g) must be included. The boundary node has no neighbour for the face (f, g) . Without loss of generality, we can assume that $\mathbf{V} \cdot \tilde{\mathbf{n}} \geq 0$ and that there is no diffusion normal to the face (f, g) . If either condition is violated, then i must be a Dirichlet node. Hence, an appropriate discretization method will require knowledge of the solution only at points within the computational domain.

We can then define the discrete outflow term

$$w_i^{n+1} = \begin{cases} \int_f^g \hat{\mathbf{n}} ds \cdot \mathbf{V}_i U_i^{n+1} & \text{if } i \text{ is a boundary node and } \int_f^g \hat{\mathbf{n}} ds \cdot \mathbf{V}_i < 0 \\ 0 & \text{otherwise.} \end{cases} \quad (3.9)$$

Note that for perpendicular bisector control volumes, triangles that contain a boundary edge should contain angles no greater than $\frac{\pi}{2}$ in order to ensure that the control volumes fall within the computational domain (which also prevents self-intersecting control volumes, refer to Figure 3.3. Combining approximations (3.2) to (3.9) and incorporating a temporal weighting factor, θ , gives us the following nonconservative finite volume discretization of equation (2.4)

$$A_i \left(\frac{U_i^{n+1} - U_i^n}{\Delta t} \right) = \theta \left(\sum_{j \in \Omega_i} \eta_{ij} (U_j^{n+1} - U_i^{n+1}) + \mathbf{V}_i \cdot \sum_{j \in \Omega_i} \tilde{L}_{ij} U_{ij+\frac{1}{2}}^{n+1} - A_i \mathbf{r}_i U_i^{n+1} \right)$$

$$\begin{aligned}
& + (1 - \theta) \left(\sum_{j \in \Omega_i} \eta_{ij} (U_j^n - U_i^n) + \mathbf{V}_i \cdot \sum_{j \in \Omega_i} \tilde{L}_{ij} U_{ij+\frac{1}{2}}^n - A_i r_i U_i^n \right) \\
& + p_i^{n+1} + \theta w_i^{n+1} + (1 - \theta) w_i^n,
\end{aligned} \tag{3.10}$$

where $\eta_{ij} = - \int_{\Omega} \nabla' N_i \mathbf{D}_i \nabla N_j d\Omega$. When $\theta = 1$ scheme (3.10) is fully implicit, $\theta = \frac{1}{2}$ gives us the Crank-Nicolson method and $\theta = 0$ produces a fully explicit scheme. Note that since discretization (3.10) is nonconservative, it is valid for both interior and boundary nodes. If the discretization were conservative we could no longer solve the underlying PDE on portions of the boundary. That is, we could no longer assume that there is no diffusion normal to the boundary, or that the convective flow is outward.

Let L be a large number and U_i^* be the value that must be met by U_i^{n+1} in order to satisfy a constraint on the solution. For example, in the case of an American option, U_i^* is equal to the payoff function. For an American option, the penalty term in equation (3.10) is defined to be

$$p_i^{n+1} = \begin{cases} \frac{A_i L}{\Delta t} (U_i^* - U_i^{n+1}) & \text{if } U_i^{n+1} < U_i^* \\ 0 & \text{otherwise.} \end{cases}$$

For a sufficiently large L , equation (3.10) becomes¹ $U_i^{n+1} = \frac{L}{L+1} U_i^*$ or, equivalently, $U_i^{n+1} = U_i^* - \epsilon$ where $\epsilon = \frac{U_i^*}{L+1}$.

The selection of an appropriately sized L is related to the convergence tolerance [80]. In [80] it is shown that the above penalty method approach for American

¹Assuming that the terms $\sum_{j \in \Omega_i} \eta_{ij} (U_j^{n+1} - U_i^{n+1})$ and $\mathbf{V}_i \cdot \sum_{j \in \Omega_i} \tilde{L}_{ij} U_{ij+\frac{1}{2}}^{n+1}$ in equation (3.10) are bounded. This should be the case in the presence of the penalty term for $t > 0$ because of parabolic smoothing, although we have not proven it.



3.1 Nonconservative Discretization of Diffusion

$$\oint_{\partial FV_i} (\mathbf{D}_i \nabla U) \cdot \vec{n} d\Gamma = \sum_{\Delta_m \in \mathcal{T}_i} \int_{\partial FV_i^m} (\mathbf{D}_i \nabla U) \cdot \vec{n} d\Gamma. \quad (3.11)$$

With reference to Figure 3.4, $\nabla U = \sum_{j \neq i} \nabla N_j^m (U_j - U_i)$ since $U = \sum_j N_j U_j$, where the N_j^m are the usual C^0 Lagrange basis functions on Δ_m (note that $\nabla N_i = -\nabla N_j - \nabla N_k$). Then the R.H.S. of equation (3.11) can be written as

$$\sum_{\Delta_m \in \mathcal{T}_i} \left(\int_{\partial FV_i^m} (\mathbf{D}_i(\nabla N_j^m (U_j - U_i) + \nabla N_k^m (U_k - U_i))) \cdot \tilde{\mathbf{n}} d\Gamma \right). \quad (3.12)$$

We can rewrite expression (3.12) as

$$\sum_{\Delta_m \in \mathcal{T}_i} \left((\mathbf{D}_i(\nabla N_j^m (U_j - U_i) + \nabla N_k^m (U_k - U_i))) \cdot \int_{\partial FV_i^m} \tilde{\mathbf{n}} d\Gamma \right)$$

which is equal to

$$\sum_{\Delta_m \in \mathcal{T}_i} \left((\mathbf{D}_i(\nabla N_j^m (U_j - U_i) + \nabla N_k^m (U_k - U_i))) \times \left(\int_{\partial \Gamma_{face_j}^m} \tilde{\mathbf{n}} d\Gamma_{face_j}^m + \int_{\partial \Gamma_{face_k}^m} \tilde{\mathbf{n}} d\Gamma_{face_k}^m \right) \right), \quad (3.13)$$

where $\Gamma_{face_j}^m$ and $\Gamma_{face_k}^m$ are the portions of the control volume in Δ_m (see Figure 3.4). Note that the integrals in expression (3.13) are independent of path. Hence, without loss of generality, we can take $\Gamma_{face_j}^m$ and $\Gamma_{face_k}^m$ to be portions of a centroid control volume. In [65] it is shown that for centroid control volumes

$$\int_{\partial \Gamma_{face_j}^m} \tilde{\mathbf{n}} d\Gamma_{face_j}^m + \int_{\partial \Gamma_{face_k}^m} \tilde{\mathbf{n}} d\Gamma_{face_k}^m = \frac{1}{6} l_i^m \tilde{\mathbf{n}}_i^m - \frac{1}{6} l_j^m \tilde{\mathbf{n}}_j^m + \frac{1}{6} l_i^m \tilde{\mathbf{n}}_i^m - \frac{1}{6} l_k^m \tilde{\mathbf{n}}_k^m,$$

where l_i^m is the length of triangle edge i for Δ_m and \bar{n}_i^m is the outward pointing unit normal to edge i (refer to Figure 3.4). Thus, expression (3.13) equals

$$\sum_{\Delta_m \in \mathcal{T}_i} \left((\mathbf{D}_i(\nabla N_j^m(U_j - U_i) + \nabla N_k^m(U_k - U_i))) \cdot \left(\frac{1}{6} l_i^m \bar{n}_i^m - \frac{1}{6} l_j^m \bar{n}_j^m + \frac{1}{6} l_i^m \bar{n}_i^m - \frac{1}{6} l_k^m \bar{n}_k^m \right) \right). \quad (3.14)$$

After noting that $\nabla N_i^m = -\frac{l_i^m \bar{n}_i^m}{2|\Delta_m|}$ (see [6]), where $|\Delta_m|$ is the area of Δ_m , expression (3.14) becomes

$$\sum_{\Delta_m \in \mathcal{T}_i} \left((\mathbf{D}_i(\nabla N_j^m(U_j - U_i) + \nabla N_k^m(U_k - U_i))) \cdot \left(-\frac{1}{3} \nabla N_i^m + \frac{1}{3} \nabla N_j^m - \frac{1}{3} \nabla N_i^m + \frac{1}{3} \nabla N_k^m \right) |\Delta_m| \right)$$

which simplifies to

$$\sum_{\Delta_m \in \mathcal{T}_i} \left((\mathbf{D}_i(\nabla N_j^m(U_j - U_i) + \nabla N_k^m(U_k - U_i))) \cdot (-\nabla N_i^m) |\Delta_m| \right) \quad (3.15)$$

since $\nabla N_i = -\nabla N_j - \nabla N_k$. Expression (3.15) equals

$$\sum_{\Delta_m \in \mathcal{T}_i} \left(-\int_{\Delta_m} \nabla' N_i^m \mathbf{D}_i \nabla N_j^m d\Omega (U_j - U_i) - \int_{\Delta_m} \nabla' N_i^m \mathbf{D}_i \nabla N_k^m d\Omega (U_k - U_i) \right)$$

which simplifies to

$$\sum_{j \in \Omega_i} \int_{\Omega} \nabla' N_i \mathbf{D}_i \nabla N_j d\Omega (U_j - U_i). \quad (3.16)$$

Expression (3.16) is the finite volume discretization of the nonconservative diffusion term in equation (2.4).

3.2 Relationship to a Galerkin Discretization

Equation (3.10) can be viewed as a finite element discretization. We will now show that the $\eta_{ij} = - \int_{\Omega} \nabla' N_i \mathbf{D}_i \nabla N_j d\Omega$ in discretization (3.10) are equivalent to the η_{ij} which result from using a Galerkin approach with a low order quadrature rule.

After performing integration by parts, we have, for the nonconservative diffusion in (2.4)

$$\eta_{ij} = - \int_{\Omega} \nabla' (N_i \mathbf{D}) \nabla N_j d\Omega. \quad (3.17)$$

By Green's theorem equation (3.17) becomes

$$\eta_{ij} = \sum_{\Delta_m \in \mathcal{T}_{ij}} \left(- \oint_{\partial \Delta_m} N_i^m \mathbf{D} \nabla N_j^m \cdot \tilde{\mathbf{n}} d\Gamma \right), \quad (3.18)$$

where \mathcal{T}_{ij} is the set of two triangles which share the common edge k (refer to Figure 3.4) and N_i^m is basis function N_i on triangle m (Δ_m). Equation (3.18) is equivalent to

$$\eta_{ij} = \sum_{\Delta_m \in \mathcal{T}_{ij}} \left(- \int_{\Gamma_j^m} N_i^m \mathbf{D} \nabla N_j^m \cdot \tilde{\mathbf{n}}_j^m d\Gamma_j^m - \int_{\Gamma_k^m} N_i^m \mathbf{D} \nabla N_j^m \cdot \tilde{\mathbf{n}}_k^m d\Gamma_k^m \right) \quad (3.19)$$

where Γ_j^m is triangle edge j and $\tilde{\mathbf{n}}_j^m$ is the outward pointing unit normal to edge j (see Figure 3.4) for Δ_m . Note that edge i does not appear in equation (3.19)

because N_i is zero along edge i . After integrating using the trapezoidal rule, (3.19) becomes

$$\eta_{ij} = \sum_{\Delta_m \in \mathcal{T}_{ij}} \left(-\frac{1}{2} l_j^m (N_i^m \mathbf{D}_i + N_j^m \mathbf{D}_j) \nabla N_j^m \cdot \bar{\mathbf{n}}_j^m - \frac{1}{2} l_k^m (N_i^m \mathbf{D}_i + N_k^m \mathbf{D}_k) \nabla N_j^m \cdot \bar{\mathbf{n}}_k^m \right) + \mathcal{O}(h) \quad (3.20)$$

where l_j^m is the length of triangle edge j for Δ_m , N_i^m is the value of N_i^m at node j , \mathbf{D}_j denotes \mathbf{D} evaluated at node j and h is the mesh size parameter. After noting that N_i^m is zero at nodes j and k , $\nabla N_j^m = -\frac{l_j^m \bar{\mathbf{n}}_j^m}{2|\Delta_m|}$ where $|\Delta_m|$ is the area of triangle m and N_i^m is one at node i , equation (3.21) becomes

$$\eta_{ij} = \sum_{\Delta_m \in \mathcal{T}_{ij}} \left(-\mathbf{D}_i \nabla N_j^m \cdot (-\nabla N_j^m) |\Delta_m| - \mathbf{D}_i \nabla N_j^m \cdot (-\nabla N_k^m) |\Delta_m| \right) + \mathcal{O}(h),$$

which is equivalent to

$$\eta_{ij} = - \int_{\Omega} (-\nabla' N_j - \nabla' N_k) \mathbf{D}_i \nabla N_j d\Omega + \mathcal{O}(h). \quad (3.21)$$

After noting that $\nabla N_i = -\nabla N_i - \nabla N_j$, equation (3.21) becomes

$$\eta_{ij} = - \int_{\Omega} \nabla' N_i \mathbf{D}_i \nabla N_j d\Omega + \mathcal{O}(h). \quad (3.22)$$

Note that in general $\eta_{ij} \neq \eta_{ji}$, which is a consequence of the fact that equation (2.4) is in nonconservative form.

Although the diffusion term in equation (2.4) is not self-adjoint, using a quadra-

ture rule of the above accuracy is sufficient for achieving first-order convergence in the H^1 norm for self-adjoint elliptic problems using linear elements [17]. For the sample pricing problems considered in this work, the numerical results in Chapter 6 indicate second-order convergence in L^∞ for discretization (3.10).

In addition, the finite volume approximation of the convective term (3.3) can be viewed as a Galerkin finite element method with a special quadrature rule [65]. Approximations (3.2) and (3.5) can also be derived by using a Galerkin approach with mass lumping (see [78]) since $A_i = \int_\Omega N_i d\Omega$, if the finite volumes are constructed using triangle centroids.

3.3 Summary

In this chapter we derived a nonconservative FVM, where the discretization of diffusion is equivalent to a nonconservative Galerkin finite element approach. Unlike a conservative discretization, the nonconservative approach allows one to directly (without converting the PDE into conservative form) discretize equation (2.4). Converting the underlying PDE into conservative form may introduce several complications. First, the new velocity tensor may become singular. Secondly, by converting the PDE into conservative form, we may alter the direction of flow at the boundary. Recalling the discussion in Chapter 2, we often can solve the underlying PDE on portions of the boundary because information is outward flowing (typically where the value of an underlying factor approaches zero). If the flow at those portions of the boundary is altered to be inward flowing, then we can no longer simply discretize the equation at those portions of the boundary.

Chapter 4

Discretization Analysis: Convection

In Chapter 2 we described instances in which the underlying pricing equation (2.4) becomes convection dominated or degenerate. A discretization of equation (2.4) must cope with such cases in order to be considered a general method. In this chapter, we will develop a positive coefficient scheme in order to handle convection dominated problems. The positive coefficient scheme possesses discrete local maximum and minimum principles (thus the scheme is stable). Furthermore, owing to the fact that the discretization is nonconservative, the positive coefficient scheme is compact. That is, the positive coefficient scheme produces a Jacobian with the same nonzero structure as a Jacobian that results from using central weighting or first-order upwinding. To the best of our knowledge, a compact positive coefficient scheme has not been previously developed.

To see how a positive coefficient scheme can bound local extrema, consider the

following discrete equation

$$(1 + \sum_{j \in \Omega_i} c_{ij}) U_i^{n+1} = U_i^n + \sum_{j \in \Omega_i} c_{ij} U_j^{n+1}. \quad (4.1)$$

Define $U_i^{\max} = \max(U_i^n, U_{j \in \Omega_i}^{n+1})$. If all the c_{ij} are nonnegative in equation (4.1), we can write equation (4.1) as

$$(1 + \sum_{j \in \Omega_i} c_{ij}) U_i^{n+1} \leq (1 + \sum_{j \in \Omega_i} c_{ij}) U_i^{\max}, \quad (4.2)$$

which gives us $U_i^{n+1} \leq U_i^{\max}$. Hence, U_i^{n+1} cannot exceed the value at any node to which it is connected, nor its value at the previous time step. Similarly, by defining $U_i^{\min} = \min(U_i^n, U_{j \in \Omega_i}^{n+1})$ it can also be shown that $U_i^{n+1} \geq U_i^{\min}$.

Positive coefficient schemes are sometimes called local extremum diminishing schemes (LED). However, if a scheme only possesses the property that $U_i^{\min} \leq U_i^{n+1} \leq U_i^{\max}$, this does not imply that local extrema will diminish, but rather only that local extrema will be bounded. Hence, calling such a scheme LED is a misnomer. Nonetheless, positive coefficient schemes have been applied to Hamilton-Jacobi equations (see [7]), which are similar in form to equation (2.4). It should be noted that in one dimension, positive coefficient schemes are also total variation diminishing (TVD) [72].

A standard approach for computing the value $U_{ij+\frac{1}{2}}^{n+1}$ in (3.10) is to use central weighting, that is

$$U_{ij+\frac{1}{2}}^{n+1} = \frac{U_i^{n+1} + U_j^{n+1}}{2}. \quad (4.3)$$

Such a scheme is second-order accurate, but it can introduce spurious oscillations into the solution if the problem is convection dominated. For example, one cannot ensure that the solution for the hyperbolic equation

$$\frac{\partial U}{\partial t} = -\mathbf{V}(x, y) \cdot \nabla U \quad (4.4)$$

will be free of spurious oscillations if central weighting is used. An alternative to using (4.3) is to use first-order upstream weighting, where

$$U_{ij+\frac{1}{2}}^{n+1} = U_{upij}^{n+1} = \begin{cases} U_i^{n+1} & \text{if } \vec{L}_{ij} \cdot \mathbf{V} < 0 \\ U_j^{n+1} & \text{otherwise,} \end{cases} \quad (4.5)$$

which generally produces solutions of poor quality because of excessive numerical diffusion. If U_i^{n+1} is the upstream point (U_{upij}^{n+1}) in (4.5), then U_j^{n+1} is referred to as the downstream point (U_{downij}^{n+1}) [61].

In order to avoid the excessive diffusion of first-order upstream weighting, the following flux limiting scheme can be used

$$U_{ij+\frac{1}{2}}^{n+1} = U_{upij}^{n+1} + \frac{\phi(q_{ij+\frac{1}{2}}^{n+1})}{2} (U_{downij}^{n+1} - U_{upij}^{n+1}) \quad (4.6)$$

where

$$q_{ij+\frac{1}{2}}^{n+1} = \frac{U_{upij}^{n+1} - U_{2upij}^{n+1}}{U_{downij}^{n+1} - U_{upij}^{n+1}}, \quad (4.7)$$

$$\phi(q_{ij+\frac{1}{2}}^{n+1}) = \frac{|q_{ij+\frac{1}{2}}^{n+1}| + q_{ij+\frac{1}{2}}^{n+1}}{1/\gamma_{ij+\frac{1}{2}} + |q_{ij+\frac{1}{2}}^{n+1}|} \quad (4.8)$$

and

$$\gamma_{ij+\frac{1}{2}} = \frac{\|(x_{down_{ij}}, y_{down_{ij}}) - (x_{up_{ij}}, y_{up_{ij}})\|}{\|(x_{up_{ij}}, y_{up_{ij}}) - (x_{2up_{ij}}, y_{2up_{ij}})\|},$$

where $\|\cdot\|$ denotes Euclidean length. $U_{2up_{ij}}^{n+1}$ in equation (4.7) is a convex weighting (i.e., linear interpolation) of U^{n+1} from the two adjacent nodes in the upstream triangle (see Figure 3.1). Equation (4.8) is known as the van Leer limiter [73, 70]. Since the limiter was originally designed for uniform grids, it has been modified to account for nonuniform (irregular) meshes [81]. The modified van Leer limiter (4.8) possesses the following properties

$$0 \leq \phi(q_{ij+\frac{1}{2}}^{n+1}) \leq 2$$

and

$$0 \leq \frac{\phi(q_{ij+\frac{1}{2}}^{n+1})}{q_{ij+\frac{1}{2}}^{n+1}} \leq 2\gamma_{ij+\frac{1}{2}}. \quad (4.9)$$

Equation (4.6) reverts to central weighting when $\phi(q_{ij+\frac{1}{2}}^{n+1}) = 1$ and becomes

$$U_{ij+\frac{1}{2}}^{n+1} = U_{up_{ij}}^{n+1} + \frac{\gamma_{ij+\frac{1}{2}}}{2}(U_{up_{ij}}^{n+1} - U_{2up_{ij}}^{n+1}) \quad (4.10)$$

when $\phi(q_{ij+\frac{1}{2}}^{n+1}) = q_{ij+\frac{1}{2}}^{n+1}\gamma_{ij+\frac{1}{2}}$. Equation (4.10) is equivalent to the two-point up-

stream scheme

$$U_{ij+\frac{1}{2}}^{n+1} = U_{upij}^{n+1} + \left(\frac{\|(\mathbf{x}_{downij}, \mathbf{y}_{downij}) - (\mathbf{x}_{upij}, \mathbf{y}_{upij})\|}{2} \right) \left(\frac{U_{upij}^{n+1} - U_{2upij}^{n+1}}{\|(\mathbf{x}_{upij}, \mathbf{y}_{upij}) - (\mathbf{x}_{2upij}, \mathbf{y}_{2upij})\|} \right).$$

Both central weighting and the two-point upstream scheme linearly reconstruct U^{n+1} at the cell interface.

For one-dimensional problems the flux limiting scheme (4.6) possesses local maximum and minimum principles (see [72]) and is TVD for arbitrary temporal weightings if a Courant, Friedrichs, Levy (CFL)-like condition is met [9, 81]. Equation (4.6) is a convex weighting of a second-order accurate two-point upstream scheme and central weighting (4.3). Hence, scheme (4.6) is second-order accurate except at local extrema where it reverts to a first-order upstream scheme. Unfortunately, TVD schemes can be no more than first-order accurate in two-dimensions [36].

4.1 Compact Positive Coefficient Scheme

To maintain high order accuracy for convection dominated problems, positive coefficient schemes which bound local extrema have been developed [67, 2, 6, 46]. We will now derive the conditions under which the following scheme,

$$U_{ij+\frac{1}{2}}^{n+1} = \begin{cases} \frac{1}{2}(U_i^{n+1} + U_j^{n+1}) & \text{if } \tilde{L}_{ij} \cdot \mathbf{V}(\mathbf{x}_i, \mathbf{y}_i) \geq 0 \\ U_{upij}^{n+1} + \frac{\phi(q_{ij+\frac{1}{2}}^{n+1})}{2}(U_{downij}^{n+1} - U_{upij}^{n+1}) & \text{if } \tilde{L}_{ij} \cdot \mathbf{V}(\mathbf{x}_i, \mathbf{y}_i) < 0, \end{cases} \quad (4.11)$$

with arbitrary temporal weighting and on an arbitrary triangular mesh produces a positive coefficient scheme for nonconservative hyperbolic equations such as (4.4). We will also show that the positivity of coefficients in scheme (4.11) implies that local extrema are bounded.

For ease of exposition, we will consider discretizing the first-order hyperbolic equation (4.4). The following results can be easily, although tediously, generalized to equation (2.4). We will simply provide the final result for equation (2.4).

Discretizing equation (4.4), along the lines used to derive (3.10), using scheme (4.11) gives us

$$\begin{aligned}
A_i \left(\frac{U_i^{n+1} - U_i^n}{\Delta t} \right) &= \theta \sum_{j \in \Omega_i} \alpha_{ij} [U_i^{n+1} + \frac{\phi(q_{ij+\frac{1}{2}}^{n+1})}{2} (U_j^{n+1} - U_i^{n+1})] \\
&+ \theta \sum_{j \in \Omega_i} \beta_{ij} [\frac{1}{2} (U_i^{n+1} + U_j^{n+1})] \\
&+ (1 - \theta) \sum_{j \in \Omega_i} \alpha_{ij} [U_i^n + \frac{\phi(q_{ij+\frac{1}{2}}^n)}{2} (U_j^n - U_i^n)] \\
&+ (1 - \theta) \sum_{j \in \Omega_i} \beta_{ij} [\frac{1}{2} (U_i^n + U_j^n)], \tag{4.12}
\end{aligned}$$

where A_i is the area of the control volume, $\alpha_{ij} = \min(\vec{L}_{ij} \cdot \mathbf{V}(x_i, y_i), 0)$ and $\beta_{ij} = \max(\vec{L}_{ij} \cdot \mathbf{V}(x_i, y_i), 0)$. After noting that $q_{ij+\frac{1}{2}}^{n+1} = \frac{U_i^{n+1} - U_{2up_{ij}}^{n+1}}{U_j^{n+1} - U_i^{n+1}}$, equation (4.12) becomes

$$\begin{aligned}
A_i \left(\frac{U_i^{n+1} - U_i^n}{\Delta t} \right) &= \theta \sum_{j \in \Omega_i} \alpha_{ij} [U_i^{n+1} + \frac{\phi(q_{ij+\frac{1}{2}}^{n+1})}{2q_{ij+\frac{1}{2}}^{n+1}} (U_i^{n+1} - U_{2up_{ij}}^{n+1})] \\
&+ \theta \sum_{j \in \Omega_i} \beta_{ij} [\frac{1}{2} (U_i^{n+1} + U_j^{n+1})]
\end{aligned}$$

$$\begin{aligned}
& + (1 - \theta) \sum_{j \in \Omega_i} \alpha_{ij} \left[U_i^n + \frac{\phi(q_{ij+\frac{1}{2}}^n)}{2q_{ij+\frac{1}{2}}^n} (U_i^n - U_{2up,ij}^n) \right] \\
& + (1 - \theta) \sum_{j \in \Omega_i} \beta_{ij} \left[\frac{1}{2} (U_i^n + U_j^n) \right],
\end{aligned}$$

which after collecting terms becomes

$$\begin{aligned}
A_i \left(\frac{U_i^{n+1} - U_i^n}{\Delta t} \right) & = \theta \sum_{j \in \Omega_i} \left(\alpha_{ij} + \alpha_{ij} \frac{\phi(q_{ij+\frac{1}{2}}^{n+1})}{2q_{ij+\frac{1}{2}}^{n+1}} + \frac{\beta_{ij}}{2} \right) U_i^{n+1} \\
& + \theta \sum_{j \in \Omega_i} \alpha_{ij} \frac{\phi(q_{ij+\frac{1}{2}}^{n+1})}{2q_{ij+\frac{1}{2}}^{n+1}} (-U_{2up,ij}^{n+1}) + \theta \sum_{j \in \Omega_i} \frac{\beta_{ij}}{2} U_j^{n+1} \\
& + (1 - \theta) \sum_{j \in \Omega_i} \left(\alpha_{ij} + \alpha_{ij} \frac{\phi(q_{ij+\frac{1}{2}}^n)}{2q_{ij+\frac{1}{2}}^n} + \frac{\beta_{ij}}{2} \right) U_i^n \\
& + (1 - \theta) \sum_{j \in \Omega_i} \alpha_{ij} \frac{\phi(q_{ij+\frac{1}{2}}^n)}{2q_{ij+\frac{1}{2}}^n} (-U_{2up,ij}^n) \\
& + (1 - \theta) \sum_{j \in \Omega_i} \frac{\beta_{ij}}{2} U_j^n. \tag{4.13}
\end{aligned}$$

After letting $\bar{\alpha}_{ij} = \frac{\Delta t}{A_i} \alpha_{ij}$ and $\bar{\beta}_{ij} = \frac{\Delta t}{A_i} \beta_{ij}$, and regrouping terms, equation (4.13) becomes

$$\begin{aligned}
& \left[1 + \theta \sum_{j \in \Omega_i} \left(-\bar{\alpha}_{ij} - \bar{\alpha}_{ij} \frac{\phi(q_{ij+\frac{1}{2}}^{n+1})}{2q_{ij+\frac{1}{2}}^{n+1}} - \frac{\bar{\beta}_{ij}}{2} \right) \right] U_i^{n+1} \\
& = \theta \sum_{j \in \Omega_i} \bar{\alpha}_{ij} \frac{\phi(q_{ij+\frac{1}{2}}^{n+1})}{2q_{ij+\frac{1}{2}}^{n+1}} (-U_{2up,ij}^{n+1}) + \theta \sum_{j \in \Omega_i} \frac{\bar{\beta}_{ij}}{2} U_j^{n+1} \\
& + \left[1 + (1 - \theta) \sum_{j \in \Omega_i} \left(\bar{\alpha}_{ij} + \bar{\alpha}_{ij} \frac{\phi(q_{ij+\frac{1}{2}}^n)}{2q_{ij+\frac{1}{2}}^n} + \frac{\bar{\beta}_{ij}}{2} \right) \right] U_i^n \\
& + (1 - \theta) \sum_{j \in \Omega_i} \bar{\alpha}_{ij} \frac{\phi(q_{ij+\frac{1}{2}}^n)}{2q_{ij+\frac{1}{2}}^n} (-U_{2up,ij}^n) + (1 - \theta) \sum_{j \in \Omega_i} \frac{\bar{\beta}_{ij}}{2} U_j^n. \tag{4.14}
\end{aligned}$$

Note that $\sum_{j \in \Omega_i} (\bar{\alpha}_{ij} + \bar{\beta}_{ij}) = 0$ since \mathbf{V} is always evaluated at node i because equation (4.4) was discretized in nonconservative form and $\oint \bar{n} d\Gamma = \int \nabla(1) d\Omega = 0$.

Thus, equation (4.14) simplifies to

$$\begin{aligned}
& \left[1 + \theta \sum_{j \in \Omega_i} \left(-\bar{\alpha}_{ij} \frac{\phi(q_{ij+\frac{1}{2}}^{n+1})}{2q_{ij+\frac{1}{2}}^{n+1}} + \frac{\bar{\beta}_{ij}}{2} \right) \right] U_i^{n+1} \\
&= \theta \sum_{j \in \Omega_i} \left(-\bar{\alpha}_{ij} \frac{\phi(q_{ij+\frac{1}{2}}^{n+1})}{2q_{ij+\frac{1}{2}}^{n+1}} U_{2up_{ij}}^{n+1} + \frac{\bar{\beta}_{ij}}{2} U_j^{n+1} \right) \\
&+ \left[1 + (1 - \theta) \sum_{j \in \Omega_i} \left(\bar{\alpha}_{ij} \frac{\phi(q_{ij+\frac{1}{2}}^n)}{2q_{ij+\frac{1}{2}}^n} - \frac{\bar{\beta}_{ij}}{2} \right) \right] U_i^n \\
&+ (1 - \theta) \sum_{j \in \Omega_i} \left(-\bar{\alpha}_{ij} \frac{\phi(q_{ij+\frac{1}{2}}^n)}{2q_{ij+\frac{1}{2}}^n} U_{2up_{ij}}^n + \frac{\bar{\beta}_{ij}}{2} U_j^n \right). \tag{4.15}
\end{aligned}$$

Note that $\bar{\alpha}_{ij} \leq 0$, $\bar{\beta}_{ij} \geq 0$ and $\frac{\phi(q_{ij+\frac{1}{2}}^n)}{q_{ij+\frac{1}{2}}^n} \geq 0$ (recall property (4.9)). Hence, in order that all the coefficients in (4.15) be positive we need to ensure that

$$(1 - \theta) \sum_{j \in \Omega_i} \left(\bar{\alpha}_{ij} \frac{\phi(q_{ij+\frac{1}{2}}^n)}{2q_{ij+\frac{1}{2}}^n} - \frac{\bar{\beta}_{ij}}{2} \right) \geq -1$$

which, by property (4.9), simplifies to

$$\sum_{j \in \Omega_i} \left(-\bar{\alpha}_{ij} \gamma_{ij+\frac{1}{2}} + \frac{\bar{\beta}_{ij}}{2} \right) \leq \frac{1}{(1 - \theta)}. \tag{4.16}$$

If condition (4.16) is met, then all the coefficients in (4.15) are positive. By defining

$$U_i^{\max} = \max(U_{k \in \Omega_i}^{n+1}, U_{k \in \Omega_i}^n, U_i^n)$$

(noting that $U_{2up_i}^{n+1}$ is the convex weighting of two $U_{k, k \in \Omega_i}^{n+1}$) we can write equation (4.15) as

$$\begin{aligned}
& \left[1 + \theta \sum_{j \in \Omega_i} \left(-\bar{\alpha}_{ij} \frac{\phi(q_{ij+\frac{1}{2}}^{n+1})}{2q_{ij+\frac{1}{2}}^{n+1}} + \frac{\bar{\beta}_{ij}}{2} \right) \right] U_i^{n+1} \\
& \leq \theta \sum_{j \in \Omega_i} \left(-\bar{\alpha}_{ij} \frac{\phi(q_{ij+\frac{1}{2}}^{n+1})}{2q_{ij+\frac{1}{2}}^{n+1}} + \frac{\bar{\beta}_{ij}}{2} \right) U_i^{\max} \\
& + \left[1 + (1 - \theta) \sum_{j \in \Omega_i} \left(\bar{\alpha}_{ij} \frac{\phi(q_{ij+\frac{1}{2}}^n)}{2q_{ij+\frac{1}{2}}^n} - \frac{\bar{\beta}_{ij}}{2} \right) \right] U_i^{\max} \\
& + (1 - \theta) \sum_{j \in \Omega_i} \left(-\bar{\alpha}_{ij} \frac{\phi(q_{ij+\frac{1}{2}}^n)}{2q_{ij+\frac{1}{2}}^n} + \frac{\bar{\beta}_{ij}}{2} \right) U_i^{\max}
\end{aligned}$$

which simplifies to

$$\begin{aligned}
& \left[1 + \theta \sum_{j \in \Omega_i} \left(-\bar{\alpha}_{ij} \frac{\phi(q_{ij+\frac{1}{2}}^{n+1})}{2q_{ij+\frac{1}{2}}^{n+1}} + \frac{\bar{\beta}_{ij}}{2} \right) \right] U_i^{n+1} \\
& \leq \left[1 + \theta \sum_{j \in \Omega_i} \left(-\bar{\alpha}_{ij} \frac{\phi(q_{ij+\frac{1}{2}}^{n+1})}{2q_{ij+\frac{1}{2}}^{n+1}} + \frac{\bar{\beta}_{ij}}{2} \right) \right] U_i^{\max}.
\end{aligned}$$

Thus,

$$U_i^{n+1} \leq U_i^{\max}.$$

Similarly, if

$$U_i^{\min} = \min(U_{k \in \Omega_i}^{n+1}, U_{k \in \Omega_i}^n, U_i^n)$$

then

$$U_i^{n+1} \geq U_i^{\min}.$$

Hence, if condition (4.16) is satisfied then scheme (4.11) possesses local maximum and minimum principles. It is interesting to note that this result holds only when a nonconservative discretization is used, since the velocity (\mathbf{V}) is an arbitrary function of the coordinates.

Although similar results can be obtained if the flux limiting scheme (4.6) is used when either $\tilde{L}_{ij} \cdot \mathbf{V}(x_i, y_i) < 0$ or $\tilde{L}_{ij} \cdot \mathbf{V}(x_i, y_i) \geq 0$, there are two advantages to using scheme (4.6) only when $\tilde{L}_{ij} \cdot \mathbf{V}(x_i, y_i) < 0$. First, the amount of additional numerical diffusion is reduced when we decrease the use of the flux limiter. Secondly, the value of U_i^{n+1} depends only on values at neighbouring nodes. That is, the value of U_i^{n+1} depends only on the $U_{j \in \Omega_i}^n$ and the $U_{j \in \Omega_i}^{n+1}$. Hence, the nonzero structure of the Jacobian matrix is no different from that resulting when central weighting or first-order upstream weighting is used.

When diffusion is present¹, scheme (4.11) can be used. In such cases condition (4.16) becomes (using similar arguments)

$$\sum_{j \in \Omega_i} \left(-\bar{\alpha}_{ij} \gamma_{ij+\frac{1}{2}} + \frac{\bar{\beta}_{ij}}{2} \right) - \bar{\eta}_{ii} \leq \frac{1}{(1-\theta)}, \quad (4.17)$$

where $\bar{\eta}_{ii}$ is $\frac{\Delta t}{A_i} \eta_{ii}$. Note that $-\eta_{ii}$ is positive. If condition (4.17) is met and all the

¹Note that the pricing equation (2.4) has an exponential decay term, which we ignore for clarity. We can perform an analysis on the full equation by multiplying the U_i^{n+1} by a factor in order to isolate the solution from the decay term, as was done in [81].

η_{ij} are nonnegative, then it can be shown that the scheme possesses local maximum and minimum principles. In general, it is not possible to ensure that all the η_{ij} will be nonnegative for equation (2.4) because the diffusion terms are nonconstant. If the η_{ij} are not all nonnegative and U satisfies a Lipschitz condition, then as will be shown in Chapter 5

$$U_i^{\min} + \mathcal{O}(h) \leq U_i^{n+1} \leq U_i^{\max} + \mathcal{O}(h),$$

where h denotes the mesh spacing, when scheme (4.11) is used.

If all the η_{ij} are nonnegative then the following scheme

$$U_{ij+\frac{1}{2}}^{n+1} = \begin{cases} \frac{1}{2}(U_i^{n+1} + U_j^{n+1}) & \text{if } \frac{\bar{L}_{ij} \cdot \mathbf{V}(x_i, y_i)}{2} + \eta_{ij} \geq 0 \\ U_{upij}^{n+1} + \frac{\phi(q_{ij+\frac{1}{2}}^{n+1})}{2}(U_{downij}^{n+1} - U_{upij}^{n+1}) & \text{if } \frac{\bar{L}_{ij} \cdot \mathbf{V}(x_i, y_i)}{2} + \eta_{ij} < 0 \end{cases} \quad (4.18)$$

can be used instead of (4.11). After some tedious algebra, it can be shown that scheme (4.18) will possess local maximum and minimum principles when the following condition is satisfied

$$\sum_{\substack{j \in \Omega_i \\ \frac{\alpha_{ij}}{2} + \eta_{ij} < 0}} -\bar{\alpha}_{ij} \gamma_{ij+\frac{1}{2}} + \sum_{j \in \Omega_i} \frac{\bar{\beta}_{ij}}{2} + \sum_{\substack{j \in \Omega_i \\ \frac{\alpha_{ij}}{2} + \eta_{ij} \geq 0}} \frac{\bar{\alpha}_{ij}}{2} - \bar{\eta}_{ii} \leq \frac{1}{(1-\theta)}. \quad (4.19)$$

There are two advantages to using scheme (4.18) instead of (4.11) when all the η_{ij} are nonnegative. First, the amount of additional numerical diffusion is usually reduced. Secondly, condition (4.19) is usually less strict than condition (4.17).

Although it can only be shown that local extrema will be approximately bounded

if we have negative η_{ij} , scheme (4.18) can be modified to

$$U_{ij+\frac{1}{2}}^{n+1} = \begin{cases} U_{upij}^{n+1} + \frac{\phi(q_{ij+\frac{1}{2}}^{n+1})}{2}(U_{downij}^{n+1} - U_{upij}^{n+1}) & \text{if } \vec{L}_{ij} \cdot \mathbf{V}(x_i, y_i) < 0 \\ & \text{and } \frac{\vec{L}_{ij} \cdot \mathbf{V}(x_i, y_i)}{2} + \eta_{ij} < 0, \\ \frac{1}{2}(U_i^{n+1} + U_j^{n+1}) & \text{otherwise,} \end{cases} \quad (4.20)$$

when we have negative η_{ij} , in order to use any positive η_{ij} in an attempt to reduce the amount of additional numerical diffusion (relative to scheme (4.11)). Scheme (4.20) maintains the nonzero structure of the Jacobian matrix which results from central weighting, and reverts to scheme (4.18) when all the η_{ij} are nonnegative. Note that if we were to use scheme (4.18) when we have negative η_{ij} , there would be an increase in the number of nonzeros in the Jacobian. Scheme (4.20) will be used for the numerical examples in this work.

4.2 Summary

In this chapter we developed a compact positive coefficient scheme for convection. The scheme utilizes any true diffusion in order to reduce the use of the flux limiter, thereby reducing the amount of augmenting diffusion required to ensure nonnegative coefficients. In other words, the positive coefficient scheme maximizes the use of central weighting. Since the discretization is nonconservative, the positive coefficient scheme is compact (the resulting Jacobian has the same nonzero structure as a Jacobian constructed using central weighting). The compact positive coefficient scheme does not appear to have been previously developed.

Chapter 5

Discretization Analysis: Diffusion

This chapter describes the conditions under which discretizations of the $(\mathbf{D}\nabla) \cdot \nabla U$ term in equation (2.4) using the finite volume method with linear shape functions are guaranteed to produce nonnegative coefficients. Recall that in Chapter 4 it was shown that the convective term can be discretized using arbitrary temporal weighting, such that the coefficients are positive. For clarity, we will only examine discretizing the diffusion term in equation (2.4). That is, we will drop time dependence and the convective term. Note that, for explicit and partially explicit schemes, a condition on the time step size must also be satisfied in order to obtain nonnegative coefficients. This chapter also relates previous work in finance to the current topic. In addition, it is shown that when a scheme produces negative coefficients, that local maximum and minimum principles are approximately satisfied.

5.1 Constant Coefficients

Consider discretizing

$$(\mathbf{D}\nabla) \cdot \nabla U = 0 \quad (5.1)$$

when \mathbf{D} is constant and symmetric positive definite using the finite volume approach used in discretization (3.10), which in this case is equivalent to a standard Galerkin finite element approach with $U = \sum_j U_j N_j$, where $U_j = U(\mathbf{x}_j, \mathbf{y}_j)$. Then for each U_i we obtain

$$\sum_{j \in \Omega_i} \eta_{ij}(U_j - U_i) = 0 \quad (5.2)$$

where, for constant \mathbf{D} ,

$$\eta_{ij} = - \int_{\Omega} \nabla' N_i \mathbf{D} \nabla N_j d\Omega. \quad (5.3)$$

For

$$\mathbf{D} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (5.4)$$

equation (5.3) becomes

$$\eta_{ij} = - \int_{\Omega} \nabla' N_i \nabla N_j d\Omega. \quad (5.5)$$

Consider two nodes, i and j , in the interior of the mesh. Let Δ_1 and Δ_2 be two triangles which share the common edge k (see Figure 3.4). Also, let l_i^1 be the length of edge i for Δ_1 and \bar{n}_i^1 be the outward pointing unit normal to edge i for Δ_1 . After noting that $\nabla N_j = -\frac{l_j \bar{n}_j}{2|\Delta|}$ for C^0 triangular shape functions, equation (5.5) can be rewritten as

$$\eta_{ij} = -\frac{l_i^1 l_j^1 \bar{n}_i^1 \cdot \bar{n}_j^1}{4|\Delta_1|} - \frac{l_i^2 l_j^2 \bar{n}_i^2 \cdot \bar{n}_j^2}{4|\Delta_2|} \quad (5.6)$$

for interior nodes. Equation (5.6) gives

$$\eta_{ij} = -\frac{l_i^1 l_j^1 \cos(\pi - \theta_{k_1})}{4|\Delta_1|} - \frac{l_i^2 l_j^2 \cos(\pi - \theta_{k_2})}{4|\Delta_2|}, \quad (5.7)$$

where $k_1, k_2 \in \Omega_i$, θ_{k_1} is θ_k (see Figure 3.4) for Δ_1 and $|\Delta_1|$ is the area of Δ_1 .

Equation (5.7) can be simplified to

$$\eta_{ij} = \frac{l_i^1 l_j^1 \cos \theta_{k_1}}{4|\Delta_1|} + \frac{l_i^2 l_j^2 \cos \theta_{k_2}}{4|\Delta_2|}.$$

Noting that $|\Delta_1| = (l_i^1 l_j^1 \sin \theta_{k_1})/2$ gives us

$$\eta_{ij} = \frac{\cot \theta_{k_1} + \cot \theta_{k_2}}{2}. \quad (5.8)$$

Hence, η_{ij} will be nonnegative if and only if $\theta_{k_1} + \theta_{k_2} \leq \pi$ [29]. In other words, the sum of the angles opposite each interior edge must be less than or equal to π . The condition $\theta_{k_1} + \theta_{k_2} \leq \pi$ is satisfied if the mesh is a Delaunay triangulation [6]. A regular triangular mesh (we define a regular triangular mesh in this work

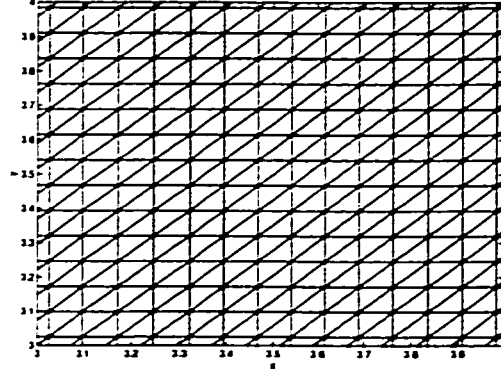


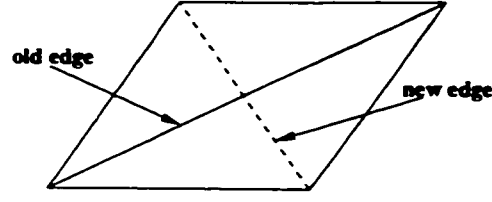
Figure 5.1: A regular triangular mesh.

to be a triangulation where the triangles have an edge parallel to the x axis and another edge parallel to the y axis) is a Delaunay triangulation. Refer to Figure 5.1 for a plot of a regular triangular mesh. Note that, in general, for triangles with a boundary edge we require that the angle opposite the boundary edge be non-obtuse in order to ensure that the η_{ij} for boundary nodes are nonnegative. Such an angle condition may not be met when the triangulation is Delaunay unless a suitable node placement is used near the boundary.

If

$$\mathbf{D} = \begin{pmatrix} k_{xx} & 0 \\ 0 & k_{yy} \end{pmatrix} \quad (5.9)$$

in (5.1), where k_{xx} and k_{yy} are constant and positive, a transformation from x, y to, say, x', y' can be performed so that diffusion tensor (5.9) becomes tensor (5.4) in the x', y' plane and a Delaunay triangulation can be constructed in the transformed plane by performing edge swapping (see [49]). Equation (5.1) can then be solved in

Figure 5.2: *Edge swapping.*

the original coordinate system by mapping the Delaunay triangulation back into the x, y plane. Note that for tensor (5.9) a regular triangulation (which is a Delaunay triangulation) in the transformed plane will also be a regular triangulation in the original coordinate system. For a given set of nodes in the original coordinate system, the transformation can be bypassed by performing edge swapping (refer to Figure 5.2) in the x, y plane to ensure that all the η_{ij} are nonnegative [29]. Swapping edges in such a manner produces a triangulation which corresponds to a Delaunay triangulation in a transformed coordinate system.

For positive definite

$$\mathbf{D} = \begin{pmatrix} k_{xx} & k_{xy} \\ k_{xy} & k_{yy} \end{pmatrix}, \quad (5.10)$$

where k_{xx} , k_{xy} and k_{yy} are constant, a rotation (to remove the cross-partial terms k_{xy}) followed by a transformation can be performed in order to obtain diffusion tensor (5.4). For a given set of nodes, one can again swap edges in the original coordinate system to construct a triangulation which corresponds to a Delaunay triangulation in a rotated/transformed coordinate system.

In summary, for a constant diffusion tensor (5.10), meshes where the η_{ij} for

interior nodes are nonnegative can always be constructed in the original coordinate system for any given set of nodes, but such meshes will generally not be regular triangulations. Regular triangulations will be guaranteed to produce nonnegative η_{ij} only when $k_{xy} = 0$. Similarly, standard finite difference schemes will not produce positive coefficients for equation (5.1) unless $k_{xy} = 0$. Finite element methods using bilinear quadrilateral elements will in general require nonorthogonal meshes (unless $k_{xy} = 0$) to ensure that the η_{ij} are nonnegative. Note that use of bilinear elements will also require, in general, an aspect ratio condition in order to ensure nonnegative η_{ij} [3].

5.2 Relation to Previous Work in Finance

Although the concepts of rotations and nonorthogonal grids for obtaining positive coefficients in numerical schemes have not been explicitly addressed in the finance literature, they are not new to finance. Consider the case where we have two factors, $x = \log(S_1)$ and $y = \log(S_2)$, where S_1 and S_2 follow processes (2.13) and (2.14), respectively. Then, the underlying PDE has the form of equation (2.4) with

$$\nabla = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \end{pmatrix},$$

$$\mathbf{D} = \frac{1}{2} \begin{pmatrix} \sigma_{S_1}^2 & \rho \sigma_{S_1} \sigma_{S_2} \\ \rho \sigma_{S_1} \sigma_{S_2} & \sigma_{S_2}^2 \end{pmatrix}, \quad (5.11)$$

and

$$\mathbf{V} = - \begin{pmatrix} r - \frac{\sigma_{S_1}^2}{2} \\ r - \frac{\sigma_{S_2}^2}{2} \end{pmatrix}. \quad (5.12)$$

If there is correlation ($\rho \neq 0$), it is suggested in [41] that two new uncorrelated variables, ψ_1 and ψ_2 , be defined by performing the following transformation

$$\psi_1 = \sigma_{S_2} x + \sigma_{S_1} y$$

and

$$\psi_2 = \sigma_{S_2} x - \sigma_{S_1} y.$$

The above transformation can be written as

$$\begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \sqrt{2}\sigma_{S_2} & 0 \\ 0 & \sqrt{2}\sigma_{S_1} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \quad (5.13)$$

Observe that

$$\begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

in equation (5.13) rotates an intermediate coordinate system in order to remove the correlation. The underlying PDE now has the form of equation (2.4) with

$$\nabla = \begin{pmatrix} \frac{\partial}{\partial \psi_1} \\ \frac{\partial}{\partial \psi_2} \end{pmatrix},$$

$$\mathbf{D} = \begin{pmatrix} (1 + \rho)\sigma_{S_1}^2 \sigma_{S_2}^2 & 0 \\ 0 & (1 - \rho)\sigma_{S_1}^2 \sigma_{S_2}^2 \end{pmatrix}$$

and

$$\mathbf{V} = - \begin{pmatrix} (r - \frac{\sigma_{S_1}^2}{2})\sigma_{S_2} + (r - \frac{\sigma_{S_2}^2}{2})\sigma_{S_1} \\ (r - \frac{\sigma_{S_1}^2}{2})\sigma_{S_2} - (r - \frac{\sigma_{S_2}^2}{2})\sigma_{S_1} \end{pmatrix}.$$

It can be shown (see Section 5.5.2) that the positive probability two-dimensional nine state lattice scheme in [34] is in fact a scheme which uses a finite element method with bilinear quadrilateral elements for diffusion to discretize the log-transformed two-asset Black-Scholes equation. The lattice scheme differs from conventional finite element methods in that the coefficients in the discrete equations are effectively given positive *a priori* values and then the nodes are placed in accordance to the *a priori* coefficient values. Note that placing nodes in appropriate locations can create higher-order compact schemes [38]. However, such an approach is restricted to the case where the velocity and diffusion tensors in the underlying PDE are constant. The positivity of the lattice scheme in Section 5.5.2 is in part

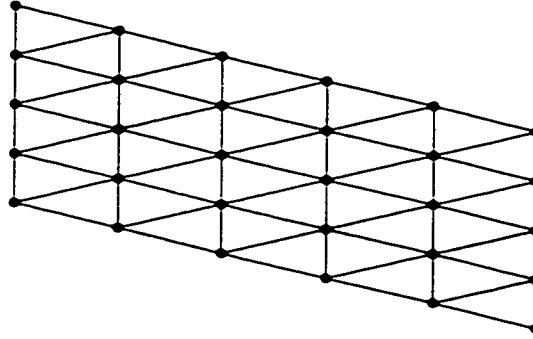


Figure 5.3: *A unique positive coefficient Delaunay triangulation for $\mathbf{D} = \mathbf{I}$.*

due to the fact the quadrilaterals are not rectangles (unless there is no correlation). That is, the underlying mesh/grid is not orthogonal, but skewed.

5.3 Nonconstant Coefficients

In option pricing models the diffusion coefficients are typically nonconstant. If transformations can be performed that produce constant diffusion coefficients, then, as was pointed out in Section 5.1, a mesh where all the η_{ij} in discretization (3.10) are nonnegative for interior nodes can always be constructed for any node placement. When the diffusion coefficients are nonconstant, the finite volume discretization of equation (5.1) becomes

$$\sum_{j \in \Omega_i} \eta_{ij} (U_j - U_i) = 0,$$

where

$$\eta_{ij} = - \int_{\Omega} \nabla' N_i \mathbf{D}_i \nabla N_j d\Omega.$$

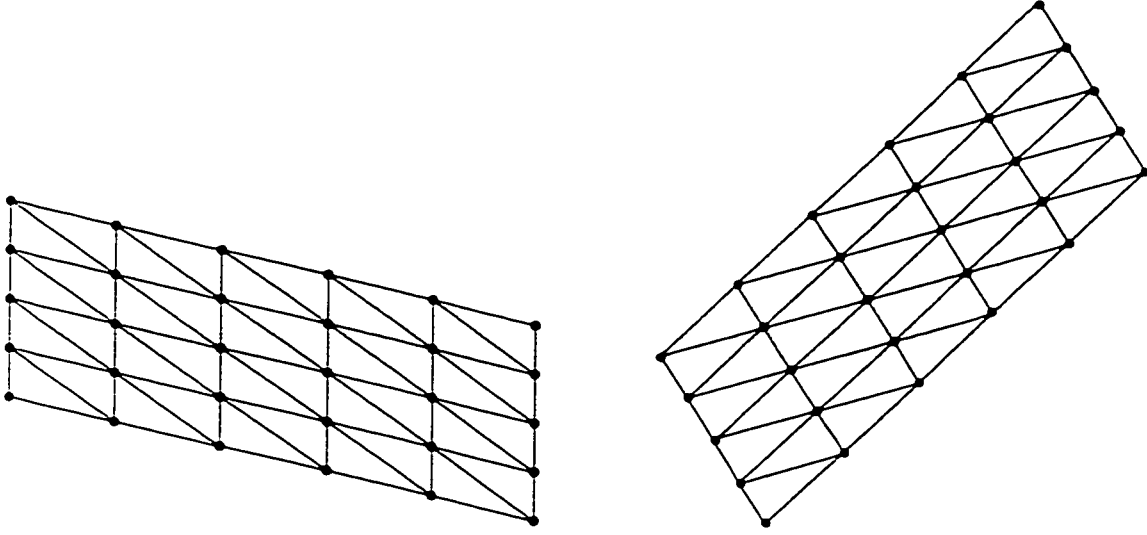


Figure 5.4: A unique positive coefficient triangulation (left) for diffusion tensor (5.14) which corresponds to the unique Delaunay triangulation (right) in a rotated/transformed coordinate system where $\mathbf{D} = \mathbf{I}$.

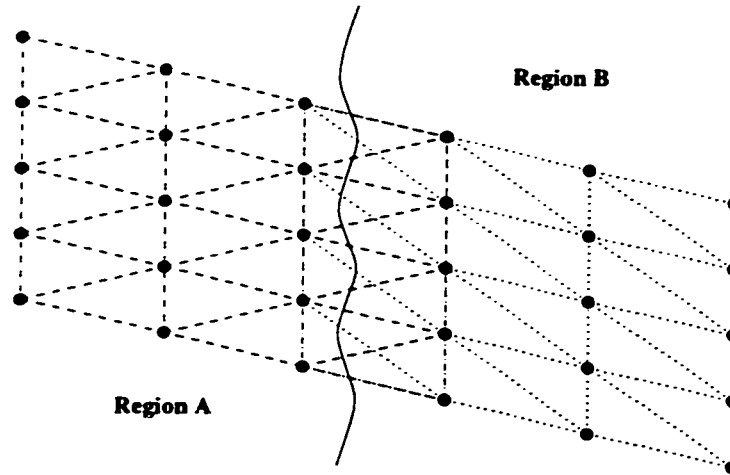


Figure 5.5: Two regions of nodes with different diffusion tensors. In region A $\mathbf{D} = \mathbf{I}$ and in region B the diffusion tensor is tensor (5.14). A triangulation which ensures that the η_{ij} are nonnegative for interior nodes in both regions cannot be constructed because nodes in each region require a different unique triangulation in order to obtain nonnegative η_{ij} .

One cannot ensure, in general, that the η_{ij} will be nonnegative for interior nodes when the diffusion tensor is nonconstant for a given node placement.

We can see that nonnegative η_{ij} cannot be ensured by means of a counter example. Consider the set of uniformly spaced nodes in Figure 5.3. By construction, for the triangulation in Figure 5.3, we have the angles opposite each interior edge summing to less than π and the angles opposite each boundary edge are less than $\frac{\pi}{2}$. Thus, if $\mathbf{D} = \mathbf{I}$, then all the η_{ij} are nonnegative. The triangulation in Figure 5.3 is a unique Delaunay triangulation since no local edge swap can be performed such that the sum of the two angles opposite the new edge will be less than or equal to π .

Now, consider the diffusion tensor

$$\mathbf{D} = \begin{pmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{pmatrix}. \quad (5.14)$$

Let x', y' be the rotated/transformed coordinate system where the transformed diffusion tensor $\mathbf{D}' = \mathbf{I}$. If the original nodes shown in Figure 5.3 are transformed into the x', y' coordinate system and an edge swap is performed, we obtain the unique Delaunay triangulation shown in the right panel of Figure 5.4. Mapping this triangulation back into the x, y plane results in the mesh shown in the left panel of Figure 5.4. This mesh ensures that all the η_{ij} are nonnegative and it is unique.

Referring to Figure (5.5), we now divide the given set of nodes into two regions (A and B) with different diffusion tensors. In region A we let $\mathbf{D} = \mathbf{I}$ and in region B

we let the diffusion tensor be tensor (5.14). Note that because of the quadrature rule used in the finite volume method, the η_{ij} for nodes in one region are independent of the different diffusion tensor in the other region. As indicated in Figure 5.5, there are no triangulations which will ensure nonnegative η_{ij} for interior nodes in both regions, because each region requires a different unique triangulation.

In the special case where

$$\mathbf{D} = \begin{pmatrix} k_{xx}(x, y) & 0 \\ 0 & k_{yy}(x, y) \end{pmatrix}, \quad (5.15)$$

which occurs in option pricing problems when $\rho = 0$, we will show that a sufficient condition for the nonnegativity of the η_{ij} is that the mesh be a regular triangulation. That is, the mesh consists of triangles with an edge parallel to the x axis and another edge parallel to the y axis. For (5.15), equation (5.6) becomes

$$\eta_{ij} = -\frac{l_i^1 l_j^1 \bar{\mathbf{n}}_i^1 \cdot \mathbf{D}_i \bar{\mathbf{n}}_j^1}{4|\Delta_1|} - \frac{l_i^2 l_j^2 \bar{\mathbf{n}}_i^2 \cdot \mathbf{D}_i \bar{\mathbf{n}}_j^2}{4|\Delta_2|}. \quad (5.16)$$

η_{ij} will be nonnegative if $\bar{\mathbf{n}}_i^1 \cdot \mathbf{D}_i \bar{\mathbf{n}}_j^1 \leq 0$ and $\bar{\mathbf{n}}_i^2 \cdot \mathbf{D}_i \bar{\mathbf{n}}_j^2 \leq 0$. In (5.16) \mathbf{D}_i determines the linear mapping $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ defined by $\bar{\mathbf{n}}_j^1 \mapsto \mathbf{D}_i \bar{\mathbf{n}}_j^1$. Since \mathbf{D}_i has nonnegative entries, $\mathbf{D}_i \bar{\mathbf{n}}_j^1$ is in the same quadrant as $\bar{\mathbf{n}}_j^1$. Thus, $\bar{\mathbf{n}}_i^1 \cdot \mathbf{D}_i \bar{\mathbf{n}}_j^1 \leq 0$ if $\bar{\mathbf{n}}_i^1$ and $\bar{\mathbf{n}}_j^1$ are separated by a quadrant or if they lie on opposite axes, which is a property of regular triangles.

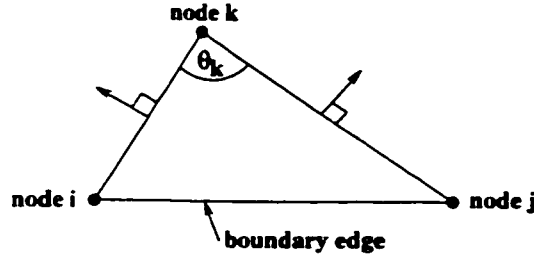


Figure 5.6: *Boundary triangle with boundary edge parallel to the x axis.*

5.3.1 Boundaries

We will now show that for boundary nodes, if there is no diffusion normal to a boundary which is parallel to the x or y axis, then there is no constraint on the size of the angle opposite to the boundary edge that is required to ensure that the η_{ij} for boundary nodes are nonnegative. Such a situation is common for option pricing problems where the original PDE is solved at the boundary. In this case, in order to have a well-posed problem, there must be no diffusion normal to the boundary and any convective flow must be outgoing. If such conditions are not met, asymptotic Dirichlet conditions are usually imposed.

Without loss of generality, we can let the boundary be parallel to the x axis (see Figure 5.6). For a node i on the boundary

$$\eta_{ij} = -\frac{l_i l_j \vec{n}_i \cdot \mathbf{D}_i \vec{n}_j}{4|\Delta|}. \quad (5.17)$$

When there is no diffusion normal to the boundary

$$\mathbf{D}_i = \begin{pmatrix} k_{xx}(x_i, y_i) & 0 \\ 0 & 0 \end{pmatrix}. \quad (5.18)$$

Let \tilde{n}_{i_x} be the x component of the unit outward normal vector to edge i (the edge opposite node i). Diffusion tensor (5.18) implies that equation (5.17) can be written as

$$\eta_{ij} = -\frac{l_i l_j k_{xx}(x_i, y_i) \tilde{n}_{i_x} \tilde{n}_{j_x}}{4|\Delta|}.$$

Since the boundary edge is parallel to the x axis, $\tilde{n}_{i_x} \tilde{n}_{j_x} \leq 0$ for any θ_k (see Figure 5.6). Hence, η_{ij} will be nonnegative for any θ_k .

5.4 Approximate Local Maximum and Minimum Principles

It cannot be shown that discrete local maximum and minimum principles hold for a scheme when any of the η_{ij} are negative. That is, it cannot be proved that $\min_{j \in \Omega_i} U_j \leq U_i \leq \max_{j \in \Omega_i} U_j$. However, we will now show that if U satisfies a Lipschitz condition, then $\min_{j \in \Omega_i} U_j + \mathcal{O}(h) \leq U_i \leq \max_{j \in \Omega_i} U_j + \mathcal{O}(h)$, where h denotes the mesh spacing, for a finite volume discretization of equation (5.1) when

$$\mathbf{D} = \begin{pmatrix} k_{xx}(x, y) & k_{xy}(x, y) \\ k_{xy}(x, y) & k_{yy}(x, y) \end{pmatrix}.$$

Assume that the computational domain is bounded. After discretizing (5.1) using

a finite volume approach, we have

$$\left(\sum_{\Delta_m \in \Delta_{i+}} -\eta_{ii}^m + \sum_{\Delta_k \in \Delta_{i-}} -\eta_{ii}^k \right) U_i(x_i, y_i) = \sum_{\Delta_m \in \Delta_{i+}} \sum_{j \in \Omega_i} \eta_{ij}^m U_j(x_j, y_j) + \sum_{\Delta_k \in \Delta_{i-}} \sum_{j \in \Omega_i} \eta_{ij}^k U_j(x_j, y_j), \quad (5.19)$$

where x_i and y_i are the coordinates at node i , and η_{ii}^m is the value of η_{ii} over triangle Δ_m . In (5.19), $\Delta_{i+} \cup \Delta_{i-}$ is the set of triangles that have node i as a vertex, where Δ_{i+} is the set of triangles where each η_{ij} is nonnegative and Δ_{i-} is the set of triangles where an η_{ij} is negative. Let $j_{+k} \in \Omega_i$ be the set of nodes where $\eta_{ij} \geq 0$, and $j_{-k} \in \Omega_i$ be the set of nodes where $\eta_{ij} < 0$. Also, let $x_{j_{-k}} = x_{j_{+k}} + \alpha_{j_{-k}} h$ and $y_{j_{-k}} = y_{j_{+k}} + \beta_{j_{-k}} h$, where $\alpha_{j_{-k}}$ and $\beta_{j_{-k}}$ are constants. Then equation (5.19) can be rewritten as

$$\begin{aligned} \left(\sum_{\Delta_m \in \Delta_{i+}} -\eta_{ii}^m + \sum_{\Delta_k \in \Delta_{i-}} -\eta_{ii}^k \right) U_i(x_i, y_i) &= \sum_{\Delta_m \in \Delta_{i+}} \sum_{j \in \Omega_i} \eta_{ij}^m U_j(x_j, y_j) \\ &+ \sum_{\Delta_k \in \Delta_{i-}} [\eta_{ij_{+k}}^k U_{j_{+k}}(x_{j_{+k}}, y_{j_{+k}}) \\ &+ \eta_{ij_{-k}}^k U_{j_{-k}}(x_{j_{+k}} + \alpha_{j_{-k}} h, y_{j_{+k}} + \beta_{j_{-k}} h)]. \end{aligned} \quad (5.20)$$

If U is Lipschitz continuous, equation (5.20) becomes

$$\begin{aligned} \left(\sum_{\Delta_m \in \Delta_{i+}} -\eta_{ii}^m + \sum_{\Delta_k \in \Delta_{i-}} -\eta_{ii}^k \right) U_i(x_i, y_i) &= \sum_{\Delta_m \in \Delta_{i+}} \sum_{j \in \Omega_i} \eta_{ij}^m U_j(x_j, y_j) \\ &+ \sum_{\Delta_k \in \Delta_{i-}} [\eta_{ij_{+k}}^k U_{j_{+k}}(x_{j_{+k}}, y_{j_{+k}}) \\ &+ \eta_{ij_{-k}}^k (U_{j_{+k}}(x_{j_{+k}}, y_{j_{+k}}) + \mathcal{O}(h))] \end{aligned}$$

which, after noting that $-\eta_{ii}^k = \eta_{ij+k}^k + \eta_{ij-k}^k$, can be simplified to

$$\begin{aligned} \left(\sum_{\Delta_m \in \Delta_{i+}} -\eta_{ii}^m + \sum_{\Delta_k \in \Delta_{i-}} -\eta_{ii}^k \right) U_i &= \sum_{\Delta_m \in \Delta_{i+}} \sum_{j \in \Omega_i} \eta_{ij}^m U_j \\ &+ \sum_{\Delta_k \in \Delta_{i-}} -\eta_{ii}^k U_{j+k} + \sum_{\Delta_k \in \Delta_{i-}} \eta_{ij-k}^k \mathcal{O}(h). \end{aligned} \quad (5.21)$$

Note that $\eta_{ij-k}^k = \bar{n}_i^k \cdot \mathbf{D}_i \bar{n}_{j-k}^k / (2 \sin \theta_{j+k}^k)$, where \bar{n}_i^k is the outward pointing unit normal to edge i for Δ_k (see Figure 3.4) and θ_{j+k}^k is θ_{j+k} for Δ_k . Since the computational domain is bounded, $\sum_{\Delta_k \in \Delta_{i-}} \eta_{ij-k}^k$ is a finite quantity assuming that the triangles are nondegenerate. Equation (5.21) can be written as

$$\begin{aligned} \left(\sum_{\Delta_m \in \Delta_{i+}} -\eta_{ii}^m + \sum_{\Delta_k \in \Delta_{i-}} -\eta_{ii}^k \right) U_i &= \sum_{\Delta_m \in \Delta_{i+}} \sum_{j \in \Omega_i} \eta_{ij}^m U_j \\ &+ \sum_{\Delta_k \in \Delta_{i-}} -\eta_{ii}^k U_{j+k} + \mathcal{O}(h). \end{aligned} \quad (5.22)$$

Since it can be shown that $-\eta_{ii} > 0$, and recalling that $\eta_{ij}^m > 0$ ($\Delta_m \in \Delta_{i+}$), we can deduce an approximate maximum principle. By defining

$$U_i^{\max} = \max_{j \in \Omega_i} U_j$$

we can write equation (5.22) as

$$\begin{aligned} \left(\sum_{\Delta_m \in \Delta_{i+}} -\eta_{ii}^m + \sum_{\Delta_k \in \Delta_{i-}} -\eta_{ii}^k \right) U_i &\leq \\ \left(\sum_{\Delta_m \in \Delta_{i+}} -\eta_{ii}^m + \sum_{\Delta_k \in \Delta_{i-}} -\eta_{ii}^k \right) U_i^{\max} &+ \mathcal{O}(h), \end{aligned} \quad (5.23)$$

which gives, after recalling $-\eta_{ii} > 0$,

$$U_i \leq U_i^{\max} + \mathcal{O}(h).$$

Similarly, if we define

$$U_i^{\min} = \min_{j \in \Omega_i} U_j,$$

then

$$U_i \geq U_i^{\min} + \mathcal{O}(h).$$

Hence,

$$\min_{j \in \Omega_i} U_j + \mathcal{O}(h) \leq U_i \leq \max_{j \in \Omega_i} U_j + \mathcal{O}(h).$$

Thus, discrete local maximum and minimum principles are approximately satisfied as $h \rightarrow 0$. Similar bounds were derived in [11] for finite difference operators.

The above result can be extended to incorporate time dependence with arbitrary temporal weighting. For example, if we incorporate time dependence and use a fully implicit scheme, equation (5.23) becomes

$$\begin{aligned} \left(1 + \frac{\Delta t}{A_i} \left(\sum_{\Delta_m \in \Delta_{i+}} -\eta_{ii}^m + \sum_{\Delta_k \in \Delta_{i-}} -\eta_{ii}^k \right) \right) U_i^{n+1} \leq \\ \left(1 + \frac{\Delta t}{A_i} \left(\sum_{\Delta_m \in \Delta_{i+}} -\eta_{ii}^m + \sum_{\Delta_k \in \Delta_{i-}} -\eta_{ii}^k \right) \right) U_i^{\max} + \mathcal{O}(h) \frac{\Delta t}{A_i}, \end{aligned} \quad (5.24)$$

where A_i is the area of the control volume and $U_i^{\max} = \max(U_i^n, U_{j \in \Omega_i}^{n+1})$. After noting that $A_i = \mathcal{O}(h^2)$, equation (5.24) can be written as

$$U_i^{n+1} \leq U_i^{\max} + \frac{1}{(h^2/\Delta t) + \mathcal{C}} \mathcal{O}(h), \quad (5.25)$$

where \mathcal{C} is some positive constant. Since $\frac{1}{(h^2/\Delta t) + \mathcal{C}} \leq \frac{1}{\mathcal{C}}$, equation (5.25) becomes

$$U_i^{n+1} \leq U_i^{\max} + \mathcal{O}(h).$$

5.5 Lattices as Finite Differences/Elements

In this section we will show that the positive probability two-dimensional lattice scheme described in [34] is equivalent to an explicit discretization which uses a finite element method with bilinear quadrilaterals for diffusion. The scheme has positive coefficients, in part, because the quadrilaterals are not rectangles (unless there is no correlation).

Before examining the two-dimensional scheme, we will show that the single-factor binomial method is equivalent to a known explicit finite difference scheme.

5.5.1 The Binomial Method

In this section we will show that the binomial method is an explicit finite difference discretization of the log-transformed Black-Scholes equation

$$\frac{\partial U}{\partial t} = \left(r - \frac{\sigma_S^2}{2}\right) \frac{\partial U}{\partial x} + \frac{\sigma_S^2}{2} \frac{\partial^2 U}{\partial x^2} - rU \quad (5.26)$$

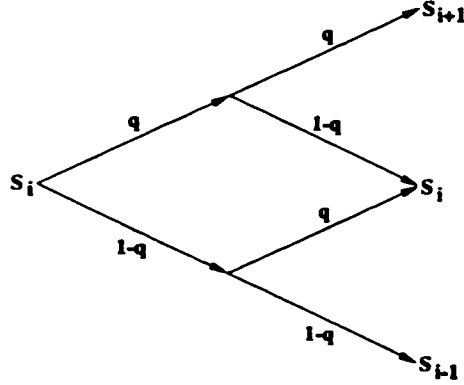


Figure 5.7: A two period binomial process.

where $x = \log(S)$, plus an $\mathcal{O}(\Delta t)$ term, where the space-like derivatives are approximated using central differences.

In the binomial method, the asset price (S) is assumed to follow a discrete multiplicative binomial process. Figure 5.7 shows the possible price movements after two time steps, where S_{i+1} is the price after two upward movements and S_{i-1} is the price after two downward movements. The value of an option U written on S can be computed using the binomial method [19]:

$$U(S_i, t_{n+1}) = e^{-r2\Delta t} [q^2 U(S_{i+1}, t_n) + 2q(1-q)U(S_i, t_n) + (1-q)^2 U(S_{i-1}, t_n)], \quad (5.27)$$

where $t_{n+1} = t_n + 2\Delta t$ and $q = (e^{r\Delta t} - e^{-\sigma_S \sqrt{\Delta t}}) / (e^{\sigma_S \sqrt{\Delta t}} - e^{-\sigma_S \sqrt{\Delta t}})$ is a risk-neutral probability. Note that $t = T - t^*$. Denoting the $U(S_i, t_n)$ as U_i^n in equation (5.27) gives us

$$U_i^{n+1} = e^{-r2\Delta t} [q^2 U_{i+1}^n + 2q(1-q)U_i^n + (1-q)^2 U_{i-1}^n]. \quad (5.28)$$

After noting that $2q(1 - q) = 1 - q^2 - (1 - q)^2$, equation (5.28) can be written as

$$U_i^{n+1} = e^{-r2\Delta t} [U_i^n + q^2(U_{i+1}^n - U_i^n) + (1 - q)^2(U_{i-1}^n - U_i^n)]$$

which results in

$$\begin{aligned} U_i^{n+1} = & e^{-r2\Delta t} U_i^n + e^{-r2\Delta t} \tilde{q}^2 \frac{(U_{i+1}^n - 2U_i^n + U_{i-1}^n)}{h^2} + \\ & e^{-r2\Delta t} (1 - 2\tilde{q}) \frac{(U_{i-1}^n - U_i^n)}{h}, \end{aligned} \quad (5.29)$$

where $\tilde{q} = e^{r\Delta t} - e^{-\sigma_S \sqrt{\Delta t}}$ and $h = e^{\sigma_S \sqrt{\Delta t}} - e^{-\sigma_S \sqrt{\Delta t}}$. By performing Taylor series expansions, the coefficients in equation (5.29) can be expressed as

$$e^{-r2\Delta t} = 1 - r2\Delta t + \mathcal{O}(\Delta t^2), \quad (5.30)$$

$$e^{-r2\Delta t} \tilde{q}^2 = \sigma_S^2 \Delta t + 2\sigma_S (r - \frac{\sigma_S^2}{2}) \Delta t^{\frac{3}{2}} + \mathcal{O}(\Delta t^2), \quad (5.31)$$

$$e^{-r2\Delta t} (1 - 2\tilde{q}) = (-2r + \sigma_S^2) \Delta t + \mathcal{O}(\Delta t^2). \quad (5.32)$$

Substituting (5.30) to (5.32) into equation (5.29) gives us

$$\begin{aligned} U_i^{n+1} = & (1 - r2\Delta t) U_i^n + (\sigma_S^2 \Delta t + 2\sigma_S (r - \frac{\sigma_S^2}{2}) \Delta t^{\frac{3}{2}}) \frac{(U_{i+1}^n - 2U_i^n + U_{i-1}^n)}{h^2} + \\ & (-2r + \sigma_S^2) \Delta t \frac{(U_{i-1}^n - U_i^n)}{h} + \mathcal{O}(\Delta t^2), \end{aligned}$$

which gives us, after noting that $h = 2\sigma_S \sqrt{\Delta t} + \mathcal{O}(\Delta t^{\frac{3}{2}})$,

$$U_i^{n+1} - U_i^n = -r2\Delta t U_i^n + \sigma_S^2 \Delta t \frac{(U_{i+1}^n - 2U_i^n + U_{i-1}^n)}{h^2} +$$

$$\begin{aligned}
& (r - \frac{\sigma_S^2}{2})\Delta t \frac{(U_{i+1}^n - 2U_i^n + U_{i-1}^n)}{h} + \\
& (-2r + \sigma_S^2)\Delta t \frac{(U_{i-1}^n - U_i^n)}{h} + \mathcal{O}(\Delta t^2). \tag{5.33}
\end{aligned}$$

Equation (5.33) assumes that U^n is smooth (at least in C^2), which will be the case for $t_n > 0$ (see [71]).¹ Simplifying equation (5.33) and dividing by $2\Delta t$ gives

$$\begin{aligned}
\frac{U_i^{n+1} - U_i^n}{2\Delta t} = & -rU_i^n + \frac{\sigma_S^2}{2} \frac{(U_{i+1}^n - 2U_i^n + U_{i-1}^n)}{h^2} + \\
& (r - \frac{\sigma_S^2}{2}) \frac{(U_{i+1}^n - U_{i-1}^n)}{2h} + \mathcal{O}(\Delta t), \tag{5.34}
\end{aligned}$$

which is an explicit finite difference scheme (where the space-like derivatives are approximated using central differences and the time step size is $2\Delta t$) for the log-transformed Black-Scholes equation plus a term of order Δt .

5.5.2 2-D Positive Probability Lattice Schemes

In this section we will show that the two-dimensional lattice scheme described in [34] is equivalent to a scheme which uses a finite element method with bilinear quadrilaterals for diffusion, characteristics [64] for convection and mass lumping. For clarity, we will first look at the one-dimensional case before examining the two-dimensional case.

We will now show that one time step of the one-dimensional lattice scheme of [34] is an explicit discretization of the log-transformed Black-Scholes equation which

¹If a discrete barrier is not present. For certain node placements, smoothing of the payoff function may be required (see [75]).

uses central differencing for the diffusion term and characteristics for convection.

Consider the case where $x = \log(S)$ and $dx = \mu_x dt + \sigma_S dW_x$, where under the assumption of risk neutrality $\mu_x = r - \frac{\sigma_S^2}{2}$. Using the scheme presented in [34], the value of an option (U) written on S after one time step (Δt) is approximated by²

$$U(x_{n+1}, t_{n+1}) = e^{-r\Delta t} \left(\frac{1}{6}U(x_{n+1} + \mu_x \Delta t + \sigma_S \sqrt{3\Delta t}, t_n) + \frac{2}{3}U(x_{n+1} + \mu_x \Delta t, t_n) + \frac{1}{6}U(x_{n+1} + \mu_x \Delta t - \sigma_S \sqrt{3\Delta t}, t_n) \right) \quad (5.35)$$

where $t = T - t^*$ and $t_{n+1} = t_n + \Delta t$. After noting that $e^{-r\Delta t} = 1 - r\Delta t + \mathcal{O}(\Delta t^2)$, equation (5.36) can be written as

$$\begin{aligned} U(x_{n+1}, t_{n+1}) &= \left(\frac{1}{6}U(x_{n+1} + \mu_x \Delta t + \sigma_S \sqrt{3\Delta t}, t_n) \right. \\ &+ \frac{2}{3}U(x_{n+1} + \mu_x \Delta t, t_n) + \frac{1}{6}U(x_{n+1} + \mu_x \Delta t - \sigma_S \sqrt{3\Delta t}, t_n) \Big) \\ &- r\Delta t \left(\frac{1}{6}U(x_{n+1} + \mu_x \Delta t + \sigma_S \sqrt{3\Delta t}, t_n) \right. \\ &+ \frac{2}{3}U(x_{n+1} + \mu_x \Delta t, t_n) \\ &+ \left. \frac{1}{6}U(x_{n+1} + \mu_x \Delta t - \sigma_S \sqrt{3\Delta t}, t_n) \right) + \mathcal{O}(\Delta t^2). \end{aligned} \quad (5.36)$$

After denoting $U(x_{n+1}, t_{n+1})$ as U_i^{n+1} and $U(x_{n+1} + \mu_x \Delta t, t_n)$ as U_j^n , equation (5.36) equals

$$\frac{U_i^{n+1} - U_j^n}{\Delta t} = \frac{\sigma_S^2}{2} \left(\frac{U_{j+1}^n - 2U_j^n + U_{j-1}^n}{h^2} \right) - \frac{r}{6}(U_{j+1}^n + 4U_j^n + U_{j-1}^n) + \mathcal{O}(\Delta t), \quad (5.37)$$

²At first glance it appears odd that the coefficients (or risk-neutral probabilities) are constant. That is, they do not appear to depend on the time step size and node spacing. However, the node spacing is coupled to the time step size. The nodes are always placed so that the coefficients equal the specified values for any time step size.

where $h = \sigma_S \sqrt{3\Delta t}$. Equation (5.37) is an explicit discretization, using central differencing for diffusion and characteristics for convection, of the log-transformed Black-Scholes equation plus a term of order Δt . Note that the discounting term $(-rU)$ in equation (5.26) is approximated using a weighted average of U_{j+1}^n , U_j^n and U_{j-1}^n .

We will now show that one time step (Δt) of the two-dimensional scheme in [34] is equivalent (to order Δt) to a discretization of the log-transformed two-dimensional Black-Scholes equation (PDE (2.4) with $\nabla = (\frac{\partial}{\partial x}, \frac{\partial}{\partial y})'$, diffusion tensor (5.11) and velocity tensor (5.12)) that uses characteristics for velocity terms, a finite element method for the diffusion terms and mass lumping.

Let $x = \log(S_1)$ and $y = \log(S_2)$ follow processes

$$dx = \mu_x dt + \sigma_{S_1} dW_x$$

and

$$dy = \mu_y dt + \sigma_{S_2} dW_y,$$

where, assuming risk-neutrality, $\mu_x = r - \frac{\sigma_{S_1}^2}{2}$ and $\mu_y = r - \frac{\sigma_{S_2}^2}{2}$. Using the scheme outlined in [34], the value of an option, U , written on S_1 and S_2 can be approximated by

$$U(x_{n+1}, y_{n+1}, t_{n+1}) = e^{-r\Delta t} \sum_{i=-1}^1 \sum_{j=-1}^1 p(i, j) U(x_{n+1} + \mu_x \Delta t + i\sigma_{S_1} \sqrt{3\Delta t}, \\ y_{n+1} + \mu_y \Delta t + (i\rho + j\sqrt{1-\rho^2})\sigma_{S_2} \sqrt{3\Delta t}, t_n), \quad (5.38)$$

where $t_{n+1} = t_n + \Delta t$, $t = T - t^*$ and with probabilities

$$(p(i, j)) = \begin{pmatrix} \frac{1}{36} & \frac{1}{9} & \frac{1}{36} \\ \frac{1}{9} & \frac{4}{9} & \frac{1}{9} \\ \frac{1}{36} & \frac{1}{9} & \frac{1}{36} \end{pmatrix} \quad i, j = -1, 0, 1.$$

After denoting $U(x_{n+1}, y_{n+1}, t_{n+1})$ as $U_{k,l}^{n+1}$ and $U(x_{n+1} + \mu_x \Delta t + i\sigma_{S_1} \sqrt{3\Delta t}, y_{n+1} + \mu_y \Delta t + (i\rho + j\sqrt{1-\rho^2})\sigma_{S_2} \sqrt{3\Delta t}, t_n)$ as $U_{i,j}^n$, equation (5.38) can be rewritten as

$$U_{k,l}^{n+1} = e^{-r\Delta t} \sum_{i=-1}^1 \sum_{j=-1}^1 p(i, j) U_{i,j}^n,$$

which, after noting that $e^{-r\Delta t} = 1 - r\Delta t + \mathcal{O}(\Delta t^2)$, equals

$$\begin{aligned} \frac{U_{k,l}^{n+1} - U_{0,0}^n}{\Delta t} + r \sum_{i=-1}^1 \sum_{j=-1}^1 p(i, j) U_{i,j}^n + \mathcal{O}(\Delta t) &= \frac{1}{\Delta t} \sum_{\substack{i=-1 \\ i \neq 0}}^1 \sum_{\substack{j=-1 \\ j \neq 0}}^1 p(i, j) U_{i,j}^n \\ &\quad - \frac{5}{9\Delta t} U_{0,0}^n. \end{aligned} \quad (5.39)$$

A discretization of the log-transformed two-dimensional Black-Scholes equation using characteristics for convection, a weighted average for the discounting term, mass lumping and a Galerkin approach with bilinear quadrilateral elements for diffusion is given by

$$\frac{U_{k,l}^{n+1} - U_{0,0}^n}{\Delta t} + r \sum_{i=-1}^1 \sum_{j=-1}^1 p(i, j) U_{i,j}^n = \frac{1}{\Delta t} \sum_{\substack{i=-1 \\ i \neq 0}}^1 \sum_{\substack{j=-1 \\ j \neq 0}}^1 \frac{1}{18} U_{i,j}^n - \frac{4}{9\Delta t} U_{0,0}^n. \quad (5.40)$$

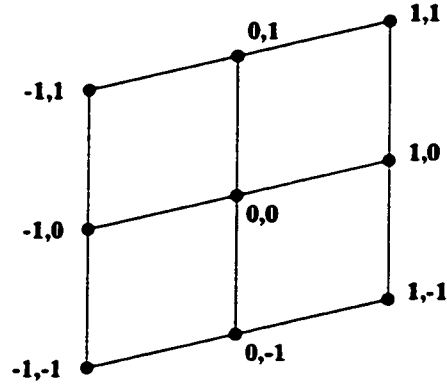


Figure 5.8: *Quadrilateral finite element mesh when correlation is positive, where node $(-1,1)$ refers to the node at which $i = -1$ and $j = 1$.*

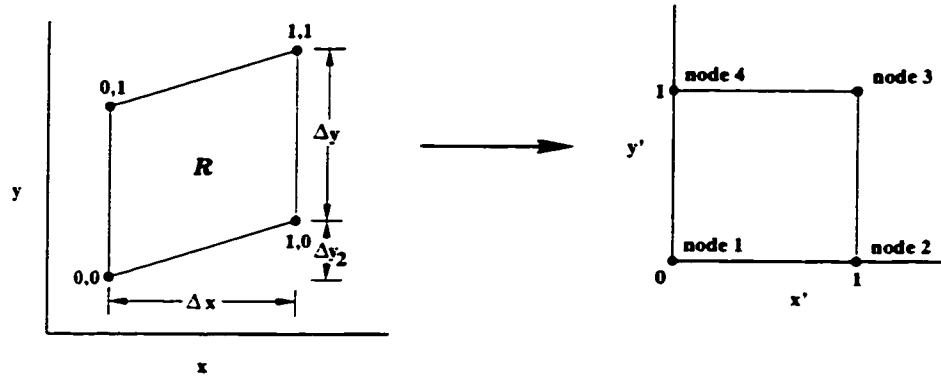


Figure 5.9: *Transforming a quadrilateral in x, y to a standard square in x', y' .*

The underlying mesh for discretization (5.40) is represented in Figure 5.8. Note that the quadrilateral elements are not rectangular (unless there is no correlation).

Determining the coefficients on the R.H.S. of equation (5.40) involves a fair amount of tedious algebra. The following provides a sketch of how the coefficients were determined. Let $\Delta x = \sigma_{S_1} \sqrt{3\Delta t}$, $\Delta y = \sigma_{S_2} \sqrt{1 - \rho^2} \sqrt{3\Delta t}$ and $\Delta y_2 =$

$\rho\sigma_{S_2}\sqrt{3\Delta t}$. The transformation

$$\begin{aligned} x' &= (x - x_{0,0})/\Delta x \\ y' &= \left(y - y_{0,0} - \frac{\rho\sigma_{S_2}}{\sigma_{S_1}}x \right) / \Delta y, \end{aligned}$$

where $x_{0,0}$ and $y_{0,0}$ are the coordinates at node $i = 0, j = 0$, converts the quadrilaterals in Figure 5.8 into standard squares (refer to Figure 5.9). Assume, without loss of generality, that we are working in the upper right quadrant of x', y' . Let \bar{N} denote the standard basis functions in x', y' and $\bar{\nabla} = (\frac{\partial}{\partial x'}, \frac{\partial}{\partial y'})'$. Then

$$\bar{\nabla} \bar{N}_1 = \begin{pmatrix} y' - 1 \\ x' - 1 \end{pmatrix}, \bar{\nabla} \bar{N}_2 = \begin{pmatrix} 1 - y' \\ -x' \end{pmatrix} \quad \text{and} \quad \bar{\nabla} \bar{N}_3 = \begin{pmatrix} y' \\ x' \end{pmatrix}.$$

Let \mathcal{R} denote the region covered by the quadrilateral in Figure 5.9,

$$J^{-1} = \begin{pmatrix} \frac{\partial x'}{\partial x} & \frac{\partial x'}{\partial y} \\ \frac{\partial y'}{\partial x} & \frac{\partial y'}{\partial y} \end{pmatrix} = \begin{pmatrix} \frac{1}{\Delta x} & 0 \\ -\frac{\rho\sigma_{S_2}}{\sigma_{S_1}} & \frac{1}{\Delta y} \end{pmatrix}$$

and recall that

$$\mathbf{D} = \frac{1}{2} \begin{pmatrix} \sigma_{S_1}^2 & \rho\sigma_{S_1}\sigma_{S_2} \\ \rho\sigma_{S_1}\sigma_{S_2} & \sigma_{S_2}^2 \end{pmatrix}.$$

Then

$$\iint_{\mathcal{R}} \nabla N_{0,0} \mathbf{D} \nabla N_{1,0} dx dy = \int_0^1 \int_0^1 \bar{\nabla} \bar{N}_1' J^{-1} \mathbf{D} (\bar{\nabla} \bar{N}_2' J^{-1})' |J| dx' dy' = \frac{\Delta x \Delta y}{36 \Delta t},$$

where $|J| = \Delta x \Delta y$ is the determinant of J . By symmetry,

$$\iint_{\mathcal{R}} \nabla N_{0,0} \mathbf{D} \nabla N_{0,1} dx dy = \iint_{\mathcal{R}} \nabla N_{0,0} \mathbf{D} \nabla N_{1,0} dx dy = \frac{\Delta x \Delta y}{36 \Delta t}.$$

Finally,

$$\iint_{\mathcal{R}} \nabla N_{0,0} \mathbf{D} \nabla N_{1,1} dx dy = \int_0^1 \int_0^1 \bar{\nabla} \bar{N}'_1 J^{-1} \mathbf{D} (\bar{\nabla} \bar{N}'_3 J^{-1})' |J| dx' dy' = \frac{\Delta x \Delta y}{18 \Delta t}.$$

By again using symmetry arguments, we can use the above results to determine the contributions from the remaining three quadrilaterals. Summing the contributions and dividing by $\Delta x \Delta y$ (which arises from mass lumping) gives us the coefficients on the R.H.S. of equation (5.40).

Now, subtracting the R.H.S. of (5.40) from the R.H.S. of (5.39) gives (after yet more tedious algebra)

$$\frac{1}{\Delta t} \sum_{\substack{i=-1 \\ i \neq 0}}^1 \sum_{\substack{j=-1 \\ j \neq 0}}^1 \left(p(i, j) - \frac{1}{18} \right) U_{i,j}^n - \frac{1}{9} U_{0,0}^n = \frac{1}{\Delta t} \left(\frac{\partial^4 U_{0,0}^n}{\partial x^2 \partial y^2} \mathcal{O}(\Delta t^2) \right) = \mathcal{O}(\Delta t),$$

which assumes U^n is smooth (at least in C^4), which is the case for $t_n > 0$. Thus, equation (5.39) equals

$$\begin{aligned} \frac{U_{k,l}^{n+1} - U_{0,0}^n}{\Delta t} &= -r \sum_{i=-1}^1 \sum_{j=-1}^1 p(i, j) U_{i,j}^n \\ &+ \frac{1}{\Delta t} \sum_{\substack{i=-1 \\ i \neq 0}}^1 \sum_{\substack{j=-1 \\ j \neq 0}}^1 \frac{1}{18} U_{i,j}^n - \frac{4}{9 \Delta t} U_{0,0}^n + \mathcal{O}(\Delta t). \end{aligned}$$

Therefore, one time step of the two-dimensional scheme in [34] is equivalent to a discretization of the two-dimensional log-transformed Black-Scholes equation which uses a finite element method with bilinear quadrilateral elements for diffusion, characteristics for velocity, mass lumping, a weighted average for the discounting term plus a term of order Δt .

5.6 Summary

For any given set of nodes, if the diffusion tensor is constant, then a mesh can always be constructed such that the coefficients are nonnegative. However, it was shown in this chapter that if the diffusion tensor is nonconstant, then a mesh which ensures positive coefficients cannot be constructed in general.

Fortunately, it was also shown that the finite volume method will approximately satisfy discrete local maximum and minimum principles when the discretization of the diffusion term produces negative coefficients.

Finally, it was shown that the binomial method and the two-dimensional positive probability scheme in [34] are in fact equivalent to known finite difference/element schemes (to order Δt). Since the binomial method is a consistent finite difference scheme, the convergence of the binomial method to the solution of the Black-Scholes equation as $\Delta t \rightarrow 0$ can be proved without resorting to the use of probabilistic arguments. With respect to the two-dimensional scheme in [34], the probabilities (coefficients) are positive, in part, because the scheme uses skewed meshes when correlation is not zero.

The above results highlight a common misconception in finance. It is sometimes

stated that lattice methods and finite difference schemes are qualitatively different. That is, lattice schemes are in some way implementations of discrete models while finite differences solve continuous models. The results in this chapter show that lattice methods and finite differences/elements are equivalent numerical techniques.

Chapter 6

Results

This chapter contains results obtained discretizing equation (2.4) using (3.10). The results are divided into two parts (I and II). In Part I, we address

- convergence
- efficiency gains
- convection dominance/degeneracy
- control volume construction

by examining several different option pricing problems. Part II addresses the issue of constructing meshes such that the discretization of the diffusion term in equation (2.4) does not produce nonnegative coefficients.

In both Parts I and II, the runs were performed using ILU-CGSTAB with level one fill for the solution of the Jacobian. The discrete equations (3.10) are nonlinear because of the flux limiter (4.8) or the penalty term. Consequently, full Newton

iteration was used to solve the algebraic equations. The absolute Newton iteration tolerance was 10^{-3} and the inner iteration tolerance was

$$\frac{\|res^m\|_2}{\|res^0\|_2} < 10^{-4},$$

where $\|res^0\|_2$ and $\|res^m\|_2$ are the 2-norms of the initial residual and the residual after m iterations, respectively.¹ All the computational domains were chosen such that increasing them had no effect on the solution at the region of interest, to at least five figures. Of course, the size of the computational domains may have to be increased for longer term options and options with higher volatilities.

In Part I all the triangulations were Delaunay and the control volumes were constructed using perpendicular bisectors (see Figure 6.1) unless stated otherwise. The Crank-Nicolson method ($\theta = \frac{1}{2}$) with a constant time step was used for the computations. It should be noted that computational savings can be obtained for long-term pricing problems by using an adaptive time-stepping routine [83].

In Part II, in order to ensure that any negative coefficients were only caused by the discretization of the diffusion term, flux limiting scheme (4.20) was always used for convection and the discretization was fully implicit ($\theta = 1$). The control volumes were constructed using triangle centroids to allow comparisons to be made between regular triangulations (which are Delaunay) and positive coefficient non-Delaunay triangulations. The sensitivities (deltas) were calculated using the method outlined in [79].

¹For pricing problems where the solution values are large, a relative tolerance should be used for the Newton iterations.

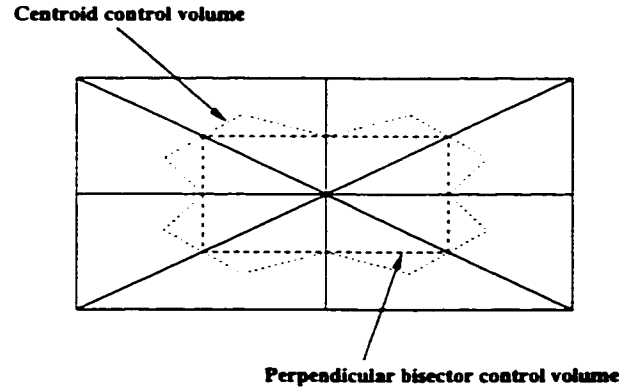


Figure 6.1: *Perpendicular bisector and centroid control volumes.*

6.1 Part I

6.1.1 Worst of Two Asset Options

To demonstrate the convergence of discretization (3.10) on an irregular mesh (see Figure 6.2), European put options on the worst of two assets were priced since an analytic solution [68] is known for such problems. Although the underlying PDE (equation (2.4) with $\nabla = (\frac{\partial}{\partial S_1}, \frac{\partial}{\partial S_2})'$, diffusion tensor (2.15) and velocity tensor (2.16)) is not convection dominated in this case, flux limiting scheme (4.20) was used in order to test the full method. The results for half year European puts with various exercise prices (K) computed on successively finer (refer to Figure 6.3) irregular meshes are contained in Table 6.1. The results demonstrate that numerical solutions of high accuracy (no more than 0.006% of the exercise price away from the analytic solution for the cases considered) can be obtained using a mesh with a relatively small number of nodes.

Table 6.2 contains the values of half year American put options on the worst

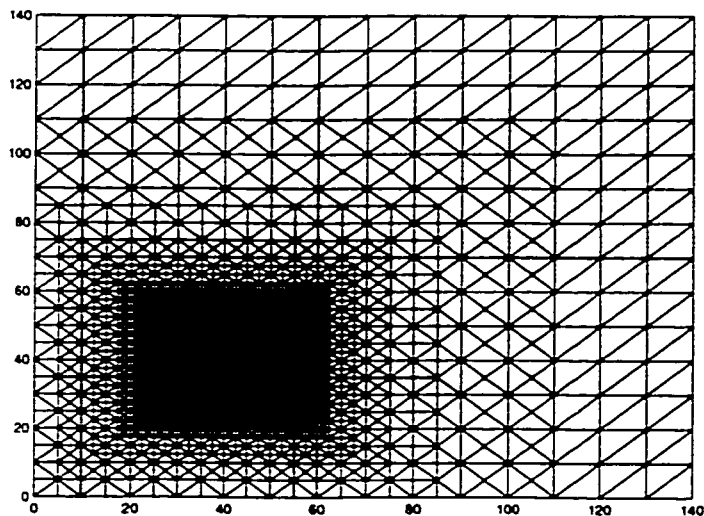


Figure 6.2: *An irregular triangular mesh with 3558 nodes.*

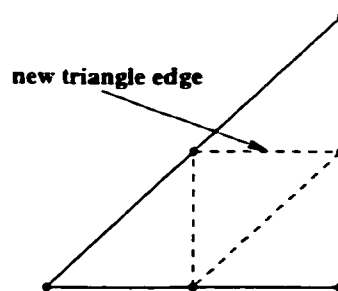


Figure 6.3: *Example of mesh refinement. The original triangle is denoted by the solid lines.*

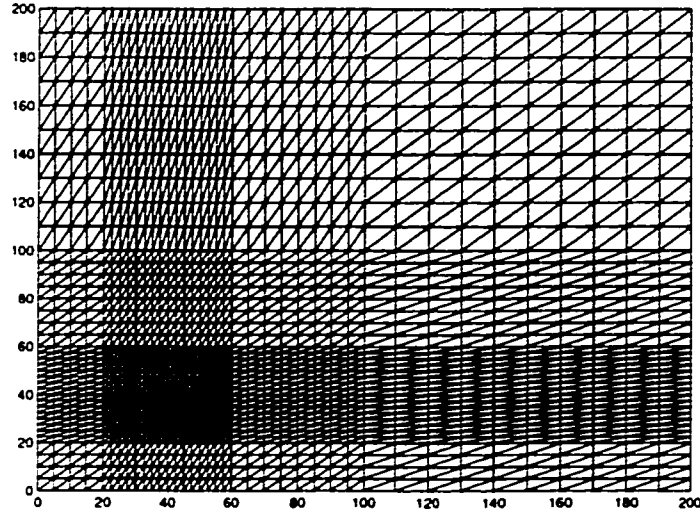


Figure 6.4: A regular triangular mesh.

Nodes		3588	13140	50220	Analytic
Δt		0.02	0.01	0.005	
	35	1.671	1.674	1.675	1.675
K	40	4.262	4.267	4.268	4.268
	45	7.986	7.990	7.991	7.991
Normalized exec. time		1	8.11	71.60	

Table 6.1: *European put options on the worst of two assets when $r = 0.05$, $\sigma_{S_1} = \sigma_{S_2} = 0.30$, $\rho = 0.5$, $T - t^* = 0.5$ and $S_1 = S_2 = 40$. The solutions were computed on successively finer irregular meshes using the modified van Leer flux limiter. The normalized execution times were obtained by using the coarse grid (3588 nodes and $\Delta t = 0.02$) execution time as the base time.*

of two assets computed using flux limiting scheme (4.20). The American case differs from the European case because the early-exercise constraint (2.18) is being imposed (using the penalty method described in Chapter 3). Successively finer regular meshes (see Figure 6.4) were used to calculate the American put prices in Table 6.2. The results in Table 6.2 appear to indicate quadratic convergence. Regular triangular meshes are similar to orthogonal grids (which are typically used when solving PDE models in finance) in the sense that if nodes are added to a region of interest, then the number of nodes in other parts of the domain will also increase. Unlike regular meshes, irregular meshes allow one to add nodes to the region of interest without increasing the number of nodes in other parts of the domain. Consequently, substantial computational savings can be achieved by using irregular meshes. Table 6.3 contains American put values computed using flux limiting scheme (4.20) and irregular meshes. Tables 6.2 and 6.3 indicate that an irregular mesh can be used to price options to within \$0.01 with an order of magnitude less computation time, relative to when a regular mesh is used. With respect to the computed values in Table 6.3, the initial irregular mesh (similar to the mesh in Figure 6.2, but with 2664 nodes) was refined by inserting nodes only near the exercise price.

Table 6.4 contains values computed using central weighting (4.3) on irregular meshes. The values in Table 6.4 are identical to the values obtained using the flux limiting scheme on the same irregular meshes (refer to Table 6.3). Scheme (4.20) uses diffusion in the underlying equation in an attempt to reduce the need for additional numerical diffusion. The results suggest that scheme (4.20) does indeed

Nodes		5929	23409	93025
Δt		0.02	0.01	0.005
K	35	1.690	1.700	1.702
	40	4.330	4.343	4.346
	45	8.127	8.138	8.142
Normalized exec. time		1.50	12.08	82.54

Table 6.2: *American put options on the worst of two assets when $r = 0.05$, $\sigma_{S_1} = \sigma_{S_2} = 0.30$, $\rho = 0.5$, $T - t^* = 0.5$ and $S_1 = S_2 = 40$. The solutions were computed on successively finer regular meshes using the modified van Leer flux limiter. The normalized execution times were obtained by using the irregular coarse grid (2664 nodes and $\Delta t = 0.02$, see Table 6.3) execution time as the base time.*

introduce little augmenting diffusion. Figure 6.5 is a plot of American put option values computed using the flux limiting scheme on an irregular mesh.

6.1.2 Asian Options

The values of quarter year European fixed strike Asian call options are contained in Table 6.5. The results were obtained by using scheme (4.20) on irregular meshes (similar to the irregular mesh in Figure 6.2). Valuing Asian options is numerically one of the more difficult pricing problems because the underlying PDE (2.4) with $\nabla = (\frac{\partial}{\partial S}, \frac{\partial}{\partial A})'$, has a convection term (see velocity tensor (2.38)) but no diffusion term (see diffusion tensor (2.37)) in one of the spatial dimensions. Hence, the PDE is degenerate. Computing the solution of such a model can be viewed as a standard problem for verifying the robustness of a numerical scheme. The cases considered in Table 6.5 are particularly difficult because of the low volatility ($\sigma_S = 0.10$). The numerical results are comparable with the results obtained using various methods

Nodes		2664	4744
Δt		0.02	0.01
K	35	1.696	1.701
	40	4.336	4.344
	45	8.131	8.139
Normalized exec. time		1	3.63

Table 6.3: American put options on the worst of two assets when $r = 0.05$, $\sigma_{S_1} = \sigma_{S_2} = 0.30$, $\rho = 0.5$, $T - t^* = 0.5$ and $S_1 = S_2 = 40$. The solutions were computed on irregular meshes using the modified van Leer flux limiter. The normalized execution times were obtained by using the coarse grid (2664 nodes and $\Delta t = 0.02$) execution time as the base time.

Nodes		2664	4744
Δt		0.02	0.01
K	35	1.696	1.701
	40	4.336	4.344
	45	8.131	8.139
Normalized exec. time		0.97	3.47

Table 6.4: American put options on the worst of two assets when $r = 0.05$, $\sigma_{S_1} = \sigma_{S_2} = 0.30$, $\rho = 0.5$, $T - t^* = 0.5$ and $S_1 = S_2 = 40$. The solutions were computed on irregular meshes using central weighting. The normalized execution times were obtained by using the coarse grid (2664 nodes and $\Delta t = 0.02$) execution time, when the modified van Leer flux limiter was used (see Table 6.3), as the base time.

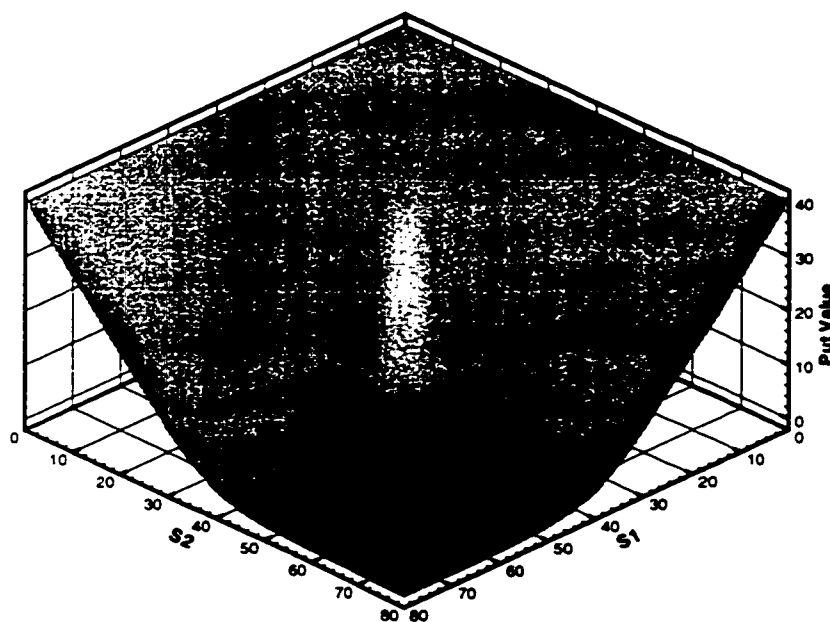


Figure 6.5: *Values for an American put option on the worst of two assets when $\tau = 0.05$, $\sigma_{S_1} = \sigma_{S_2} = 0.30$, $\rho = 0.5$, $T - t^* = 0.5$ and $K = 40$. The solution was computed on an irregular mesh with 4744 nodes using the modified van Leer flux limiter.*

Nodes		2853	10226	38634
Δt		0.01	0.005	0.0025
K	95	6.113	6.118	6.119
	100	1.829	1.848	1.852
	105	0.163	0.152	0.150
Normalized exec. time		1	7.20	69.72

Table 6.5: *European fixed strike Asian call options computed using the van Leer flux limiter on successively finer meshes when $r = 0.10$, $\sigma_S = 0.10$, $T - t^* = 0.25$, and $S = 100$. The normalized execution times were obtained by using the coarse grid (2853 nodes and $\Delta t = 0.01$) execution time as the base time.*

cited in [24]. Table 6.5 appears to again indicate quadratic convergence.

Figure 6.6 demonstrates the oscillations that can result when central weighting is used to discretize the convective term. Figure 6.6 contrasts sharply with Figure 6.7, where Figure 6.7 is a plot of a solution obtained using scheme (4.20) that is free of oscillations. Note that exactly the same parameters and irregular mesh were used to compute both solutions.

6.1.3 Convertible Bonds

As mentioned in Chapter 1, there is evidence to suggest that centroid control volumes can deteriorate accuracy [6]. Both centroid and perpendicular bisector control volume constructions were examined when a ten year convertible bond was priced. The convertible bond was continuously callable. Parameter values for the pricing problem are outlined in Table 6.6. The bond pays a 5% coupon semi-annually.

Due to the fact that the underlying variables (interest rate and stock price, see Section 2.4) differ by orders of magnitude, a rescaling of the variables was necessary

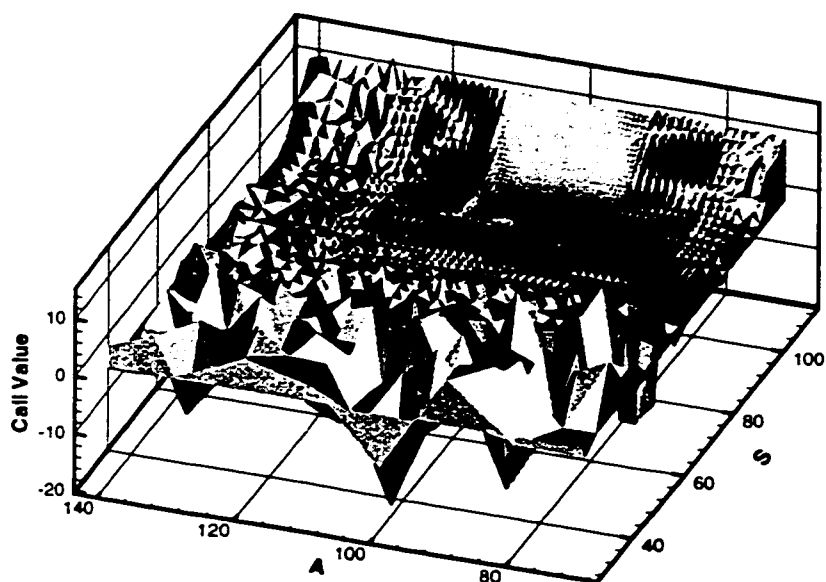


Figure 6.6: *European fixed strike Asian call option calculated using central weighting on an irregular mesh with 10226 nodes when $r = 0.10$, $\sigma_S = 0.10$, $T - t^* = 0.25$ and $K = 100$.*

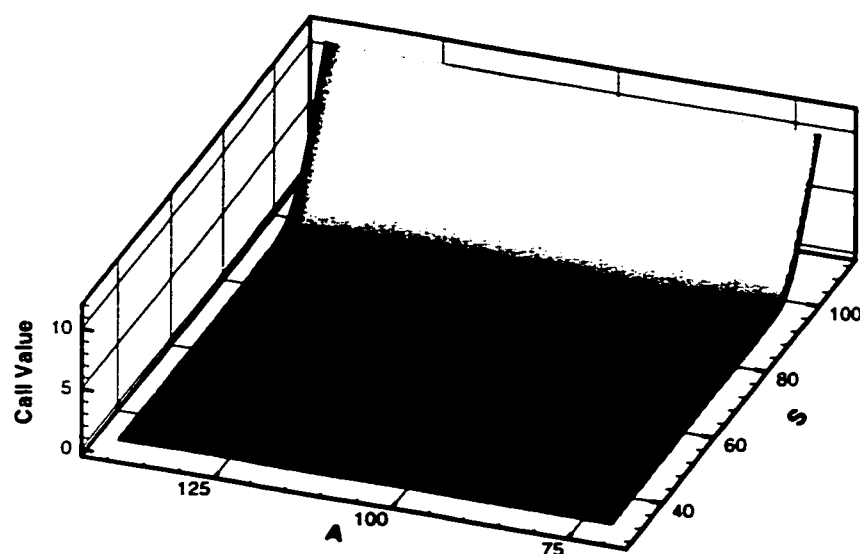


Figure 6.7: *European fixed strike Asian call option calculated using the modified van Leer flux limiter on an irregular mesh with 10226 nodes when $r = 0.10$, $\sigma_S = 0.10$, $T - t^* = 0.25$ and $K = 100$.*

Parameter	Value	Parameter	Value
σ_S	0.25	c	0.5
σ_r	0.03	λ	0.0
ρ	-0.5	ω	2.0
a	0.58	Cp	105.0
b	0.0345		

Table 6.6: Convertible bond parameter values.

Nodes		2698	10327	40214		
Δt		0.125	0.0625	0.03125		
S	50	104.189	104.231	104.231	0.04	
	40	100.462	100.489	100.488	0.08	r
	30	95.095	95.128	95.123	0.12	
Normalized exec. time		1	8.01	122.49		

Table 6.7: Values of a ten year convertible bond (at $T - t^* = 10.0$) which is continuously callable and pays a 5% coupon semi-annually. The solutions were calculated on successively finer irregular meshes using centroid control volumes and central weighting. The normalized execution times were obtained by using the coarse grid (1541 nodes and $\Delta t = 0.125$) execution time as the base time.

in order to construct Delaunay triangulations (which are required for perpendicular bisector control volumes). A new variable $y = \frac{S}{100}$ was defined, and the equations were transformed using this new variable.

The values in Table 6.7 were computed using central weighting and irregular meshes (similar to the mesh in Figure 6.2) with centroid control volumes. Table 6.8 contains values obtained using flux limiting scheme (4.20) with centroid control volumes on the same irregular meshes used when the solutions were computed using central weighting. Unlike the solutions for American put options, there is a noticeable difference between central weighting and the flux limiting scheme (4.20),

Nodes		2698	10327	40214		
Δt		0.125	0.0625	0.03125		
S	50	104.189	104.231	104.231	0.04	
	40	100.428	100.471	100.479	0.08	r
	30	95.036	95.096	95.110	0.12	
Normalized exec. time		1.32	11.38	139.49		

Table 6.8: Values of a ten year convertible bond (at $T - t^* = 10.0$) which is continuously callable and pays a 5% coupon semi-annually. The solutions were calculated on successively finer irregular meshes using centroid control volumes and the modified van Leer flux limiter. The normalized execution times were obtained by using the coarse grid (1541 nodes and $\Delta t = 0.125$) execution time, when central weighting was used (see Table 6.7), as the base time.

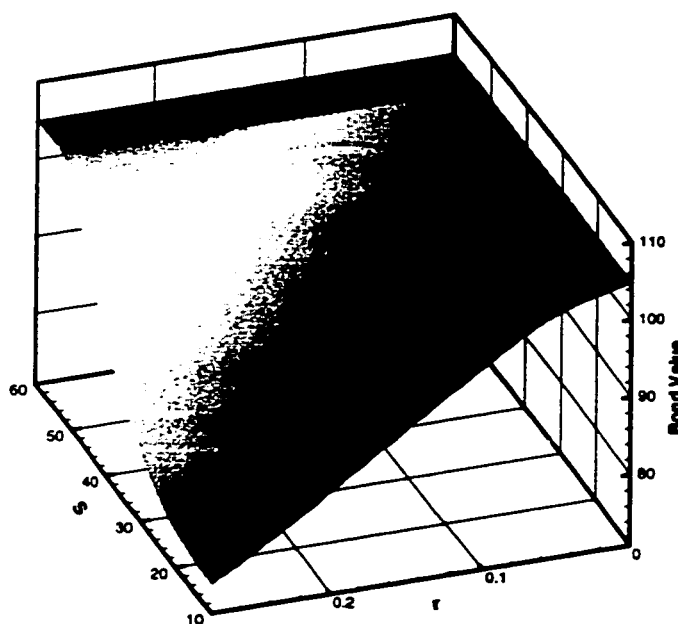


Figure 6.8: A ten year convertible bond (at $T - t^* = 10.0$) which is continuously callable and pays a 5% coupon semi-annually. The solution was computed on an irregular mesh with 5647 nodes using centroid control volumes and the modified van Leer limiter.

Nodes		2698	10327	40214		
Δt		0.125	0.0625	0.03125		
S	50	104.189	104.231	104.231	0.04	
	40	100.467	100.491	100.490	0.08	r
	30	95.097	95.132	95.129	0.12	
Normalized exec. time		1	7.09	102.81		

Table 6.9: Values of a ten year convertible bond (at $T - t^* = 10.0$) which is continuously callable and pays a 5% coupon semi-annually. The solutions were calculated on successively finer irregular meshes using perpendicular bisector control volumes and central weighting. The normalized execution times were obtained by using the coarse grid (1541 nodes and $\Delta t = 0.125$) execution time as the base time.

with the flux limiter appearing to be more slowly convergent.

In contrast, Tables 6.9 and 6.10 contain values obtained with perpendicular bisector control volumes using central weighting and flux limiting scheme (4.20), respectively. The values were computed using the same irregular meshes used for the results obtained with centroid control volumes. In this case (using perpendicular bisector control volumes), the values obtained using central weighting and the flux limiting scheme are very similar, and these results are also comparable to the results computed using central weighting with centroid control volumes (Table 6.7). The results suggest that the differences between the values calculated with centroid control volumes were due to the fact that such control volumes have a diffusive effect on the solution, particularly when the flux limiting scheme is used. Similar results have been observed in the computation of the Euler equations [30].

It is interesting to note that the increases in execution time after mesh and time step refinements were larger for the convertible bond pricing problem than for the

Nodes		2698	10327	40214		
Δt		0.125	0.0625	0.03125		
S	50	104.189	104.231	104.231	0.04	r
	40	100.465	100.490	100.489	0.08	
	30	95.084	95.125	95.128	0.12	
Normalized exec. time		1.13	9.57	120.28		

Table 6.10: Values of a ten year convertible bond (at $T - t^* = 10.0$) which is continuously callable and pays a 5% coupon semi-annually. The solutions were calculated on successively finer irregular meshes using perpendicular bisector control volumes and the modified van Leer flux limiter. The normalized execution times were obtained by using the coarse grid (1541 nodes and $\Delta t = 0.125$) execution time, when central weighting was used (see Table 6.9), as the base time.

two-asset and Asian option pricing problems.

Figure 6.8 contains a plot of the convertible bond values computed using scheme (4.20) on an irregular mesh. Referring to Figure 6.8, the effect of the call provision (constraint (2.49)) can clearly be seen for large S .

6.2 Part II

To investigate the effect of negative η_{ij} , we will examine pricing several types of options written on two assets (S_1 and S_2). We chose to examine pricing European call options on the maximum of two assets, a European digital call option and a European discrete double barrier call option on the maximum of two assets. The pricing problems were chosen based on the fact that the European call options on the maximum of two assets are standard pricing problems, the digital option has a discontinuous payoff function, and the discrete barrier option introduces discontinu-

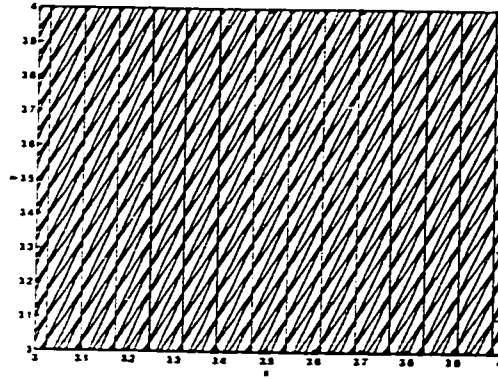


Figure 6.9: A positive coefficient mesh constructed using edge swapping when $\sigma_{S_1} = 0.10$, $\sigma_{S_2} = 0.30$ and $\rho = 0.70$.

ities at monitoring dates. The price processes in all the problems follow geometric Brownian motion.

In order to construct positive meshes for comparison purposes, we perform the change of variables $x = \log(S_1)$ and $y = \log(S_2)$ so that diffusion tensor (2.15) will be transformed into a constant tensor. After performing the transformation, the underlying PDE is equation (2.4) with $\nabla = (\frac{\partial}{\partial x}, \frac{\partial}{\partial y})'$, diffusion tensor (5.11) and velocity tensor (5.12). Note that such a transformation to produce a constant \mathbf{D} can be performed because the underlying price processes follow geometric Brownian motion. Transformations that produce constant diffusion tensors are not possible in general.

Table 6.11 contains results for quarter year European call options on the maximum of two assets. The values in Table 6.11 were computed using regular meshes and meshes constructed using edge swapping (refer to Figure 6.9). The edge-swapped meshes ensured that all the coefficients in the discrete equations for unknowns at interior nodes were positive. Note that the nodes are at the same posi-

tions for both meshes. Table 6.11 appears to indicate that the two different meshes produced comparable results (with the edge-swapped mesh being more slowly convergent). However, Figures 6.10 and 6.11 seem to suggest that the positive coefficient edge-swapped mesh produced solutions of poor quality. That is, the level curves of values and of the deltas ($\frac{\partial U}{\partial S_1}$ and $\frac{\partial U}{\partial S_2}$) are *jagged* when the edge-swapped mesh was used, but the level curves are relatively smooth when the regular mesh was used. The jaggedness in the contour levels persists even when the edge-swapped mesh is refined (see Figure 6.12).

It should be noted that the jaggedness is not caused by convection - the effect occurs even if the convection term is removed entirely from equation (2.3). Nor is the jaggedness in the level curves of the deltas an artifact of the method that was used to compute the deltas, since it is clear that the jaggedness is already present in the level curves of values.

A possible cause of the jaggedness is that the *kinks* in the payoff function occur within elements for the edge-swapped meshes and not, as is the case for the regular meshes used in the numerical examples, only at element edges. Accuracy may deteriorate when the second derivatives of the initial (terminal) condition do not exist within elements [75]. To test this hypothesis the initial condition was smoothened by projecting it onto the space spanned by the basis functions using

$$\left(\int_x \int_y \mathbf{N} \mathbf{N}' dy dx \right) \mathbf{U}^0 = \int_x \int_y g(x, y, T) \mathbf{N} dy dx, \quad (6.1)$$

where \mathbf{N} are the basis functions, \mathbf{U}^0 are the smoothened initial data and $g(\cdot)$ is the payoff function. As pointed out in [75], this should restore optimal convergence

ρ	Mesh	Numerical Solution			Analytic
0.50	Regular	2.811	2.964	3.001	3.001
	Swapped	2.797	2.948	2.994	
0.70	Regular	2.693	2.847	2.884	2.891
	Swapped	2.707	2.820	2.873	
Δt		0.005	0.0025	0.00125	
Nodes		6724	26569	106276	

Table 6.11: *Values of European call options on the maximum of two assets when $r = 0.05$, $\sigma_{S_1} = 0.10$, $\sigma_{S_2} = 0.30$, $T - t^* = 0.25$, $S_1 = S_2 = 40$ and $K = 40$. The solutions were computed on successively finer regular meshes and positive coefficient meshes constructed using edge swapping.*

rates. Note that except for the convective term in some cases, the finite volume discretization (3.10) is equivalent to a standard finite element approach when \mathbf{V} and \mathbf{D} are constant. It was found that smoothing the initial condition had little effect on the quality of the contour plots.

The jaggedness would seem to contradict the fact that the discretization with the edge-swapped mesh is a positive coefficient scheme. Although discrete local maximum and minimum principles hold for the discretization when positive coefficient meshes are used, it is interesting to note that nodes are not connected to their nearest spatial neighbours in an edge-swapped mesh. Consequently, the value at a node may not be bounded by the values at its spatially nearest neighbouring nodes.

Figure 6.13 contains a plot of level curves of values (where the values at the nodes are indicated) for a European call option on the maximum of two assets when the solution was computed using an edge-swapped mesh. Figure 6.13 demonstrates that

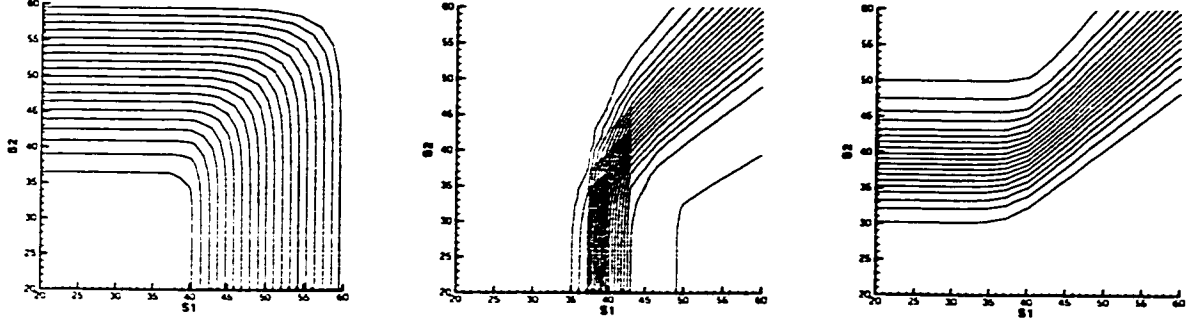


Figure 6.10: Level curves of values, deltas with respect to S_1 ($\frac{\partial U}{\partial S_1}$) and deltas with respect to S_2 ($\frac{\partial U}{\partial S_2}$) of a European call option on the maximum of two assets when $r = 0.05$, $\sigma_{S_1} = 0.10$, $\sigma_{S_2} = 0.30$, $\rho = 0.70$, $T - t^* = 0.25$ and $K = 40$. The solutions were computed with $\Delta t = 0.005$ using a regular mesh with 6724 nodes.

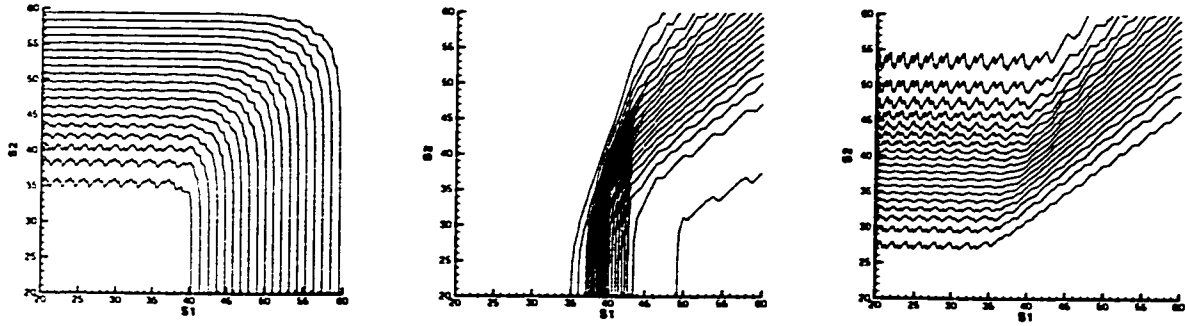


Figure 6.11: Values, deltas with respect to S_1 and deltas with respect to S_2 of a European call option on the maximum of two assets when $r = 0.05$, $\sigma_{S_1} = 0.10$, $\sigma_{S_2} = 0.30$, $\rho = 0.70$, $T - t^* = 0.25$ and $K = 40$. The solutions were computed with $\Delta t = 0.005$ using a positive coefficient edge-swapped mesh with 6724 nodes.

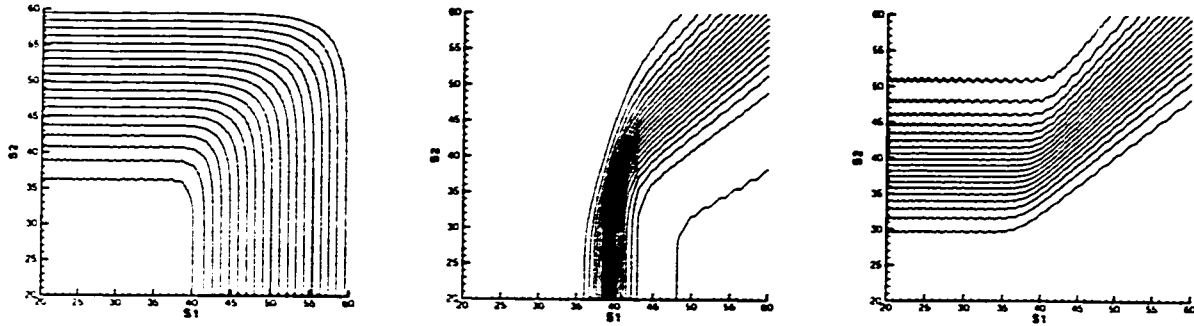


Figure 6.12: Values, deltas with respect to S_1 and deltas with respect to S_2 of a European call option on the maximum of two assets when $r = 0.05$, $\sigma_{S_1} = 0.10$, $\sigma_{S_2} = 0.30$, $\rho = 0.70$, $T - t^* = 0.25$ and $K = 40$. The solutions were computed with $\Delta t = 0.0025$ using a positive coefficient edge-swapped mesh with 26569 nodes.

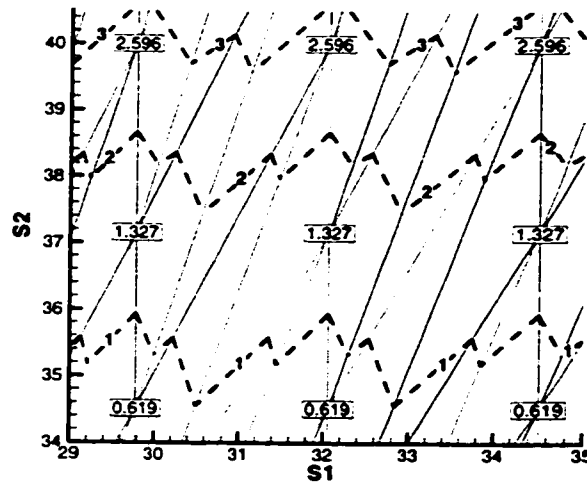


Figure 6.13: Level curves of values and values at nodes for a European call option on the maximum of two assets when $r = 0.05$, $\sigma_{S_1} = 0.10$, $\sigma_{S_2} = 0.30$, $\rho = 0.70$, $T - t^* = 0.25$ and $K = 40$. The solutions were computed with $\Delta t = 0.005$ using a positive coefficient edge-swapped mesh with 6724 nodes.

although the level curves appear to indicate an oscillatory solution, the computed values at the nodes are actually well behaved. Hence, the jaggedness is an artifact of calculating the level curves. More specifically, linear interpolation on the edge-swapped meshes produced poor results.

The poor quality of level curves computed on edge-swapped meshes is most likely due to the fact that the elements are long and thin (or *stretched*). Although stretched elements are optimal for interpolation with respect to minimizing the error bound when there is a direction with dominant curvature, the placement of the elements must be data dependent (see [22, 60]). That is, the triangles should be short in directions where the curvature of U is high, and long in directions where the curvature of U is low. Edge swapping to ensure positive coefficients ignores the curvature of U . Thus, elements may be placed with their long side in the direction of high curvature.

Poor interpolation results may be of concern when solving a model if jump conditions are present. For example, in the presence of discrete dividends, interpolation will generally be performed at ex-dividend dates.

An alternative to edge swapping is to place nodes in such a manner as to ensure positive coefficients. Recall that a regular mesh will ensure positive coefficients when the coordinate system has been rotated to eliminate the cross-partial terms. Rotating such a mesh back into the original coordinate system will ensure that the coefficients are positive without generating stretched elements (see Figure 6.14). To remove the cross-partial terms from diffusion tensor (5.11), the coordinate system must be rotated by an angle of $\frac{1}{2}\tan^{-1}\left(\frac{2\rho\sigma_1\sigma_2}{\sigma_1^2-\sigma_2^2}\right)$.

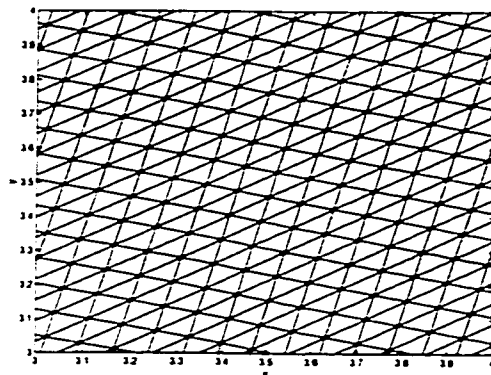


Figure 6.14: A positive coefficient mesh which corresponds to a regular mesh when the coordinate system is rotated by an angle of $\frac{1}{2}\tan^{-1}(\frac{2\rho\sigma_{S_1}\sigma_{S_2}}{\sigma_{S_1}^2 - \sigma_{S_2}^2})$ when $\sigma_{S_1} = 0.10$, $\sigma_{S_2} = 0.30$ and $\rho = 0.70$.

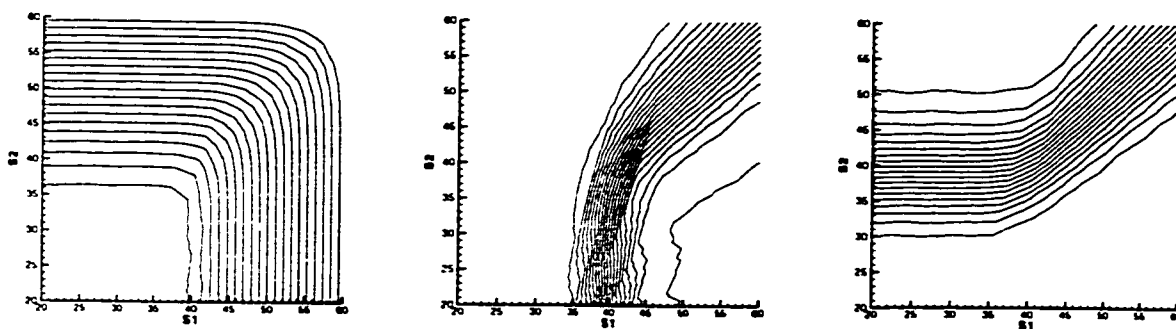


Figure 6.15: Values, deltas with respect to S_1 and deltas with respect to S_2 of a European call option on the maximum of two assets when $r = 0.05$, $\sigma_{S_1} = 0.10$, $\sigma_{S_2} = 0.30$, $\rho = 0.70$, $T - t^* = 0.25$ and $K = 40$. The solutions were computed with $\Delta t = 0.005$ using a positive coefficient rotated mesh with 6614 nodes.

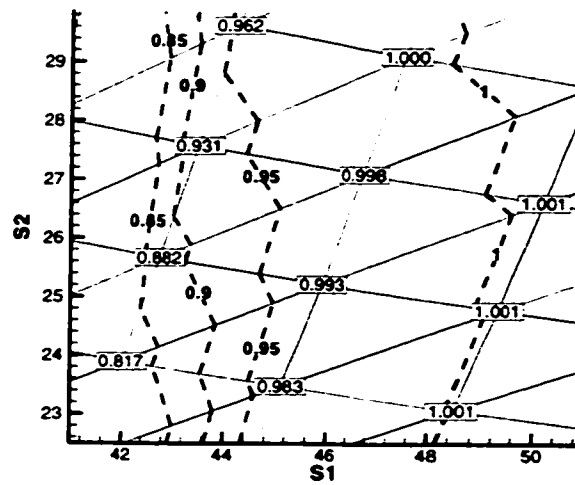


Figure 6.16: *Level curves of deltas with respect to S_1 and deltas at nodes for a European call option on the maximum of two assets when $r = 0.05$, $\sigma_{S_1} = 0.10$, $\sigma_{S_2} = 0.30$, $\rho = 0.70$, $T - t^* = 0.25$ and $K = 40$. The solutions were computed with $\Delta t = 0.005$ using a positive coefficient rotated mesh with 6614 nodes.*

Figure 6.15 contains contour plots of values and deltas of a quarter year European call on the maximum of two assets computed using a positive coefficient rotated mesh. The node positions in the rotated mesh differ from the node positions in the regular mesh, but the spacings are identical. The contours calculated using the rotated mesh are better than the contours computed using the edge-swapped mesh. However, the rotated mesh contours are not as smooth as the contours calculated using regular meshes.

The contour plots computed using the rotated mesh suggest an oscillatory solution. Using projection (6.1) to smooth the initial condition again had little effect on the quality of the contour plots. As was the case with the edge-swapped meshes, the computed solutions at the nodes are well behaved (see Figure 6.16). Hence, the jagged contours are once again due to poor interpolation.

The rotated mesh may have produced poor interpolation results, because unlike the regular mesh, the element edges in the rotated mesh are not aligned with the curvature of U . Like edge swapping, ensuring positive coefficients as a criterion for node placement ignores other aspects of the problem.

We next examined the more difficult case of pricing a quarter year digital call option. The problem is considered to be of greater difficulty because there are discontinuities in the payoff function (2.29). Figures 6.17, 6.18 and 6.19 contain contour plots of values and deltas computed using regular, edge-swapped and rotated meshes, respectively. We see the same pattern of results as in the previous pricing problem. That is, the regular mesh produced the best contours, followed by the rotated mesh and the edge-swapped mesh produced the worst contours. The

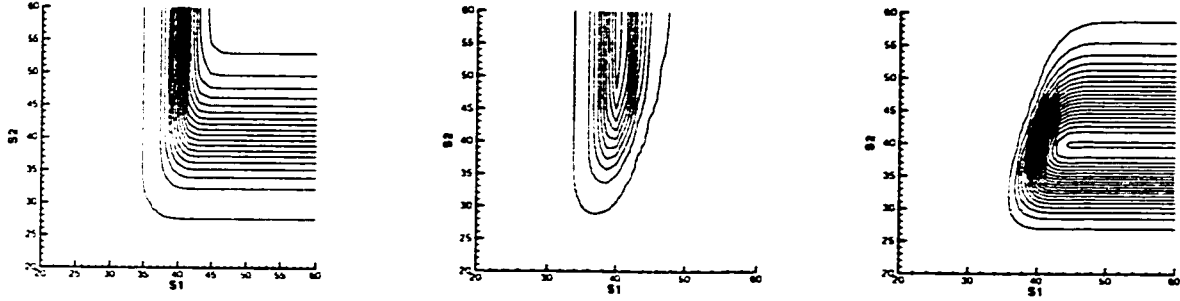


Figure 6.17: *Level curves of values and deltas of a digital call option when $r = 0.05$, $\sigma_{S_1} = 0.10$, $\sigma_{S_2} = 0.30$, $\rho = 0.70$, $T - t^* = 0.25$ and $K = 40$. The solutions were computed with $\Delta t = 0.0025$ using a regular mesh with 26569 nodes.*

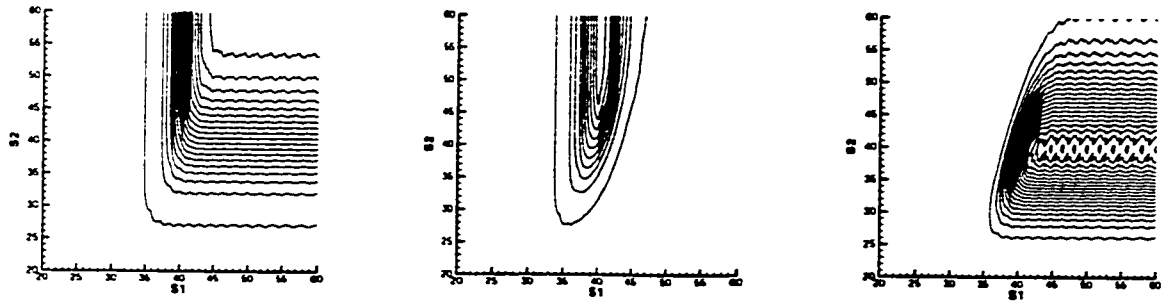


Figure 6.18: *Values, deltas with respect to S_1 and deltas with respect to S_2 of a digital call option when $r = 0.05$, $\sigma_{S_1} = 0.10$, $\sigma_{S_2} = 0.30$, $\rho = 0.70$, $T - t^* = 0.25$ and $K = 40$. The solutions were computed with $\Delta t = 0.0025$ using a positive coefficient edge-swapped mesh with 26569 nodes.*

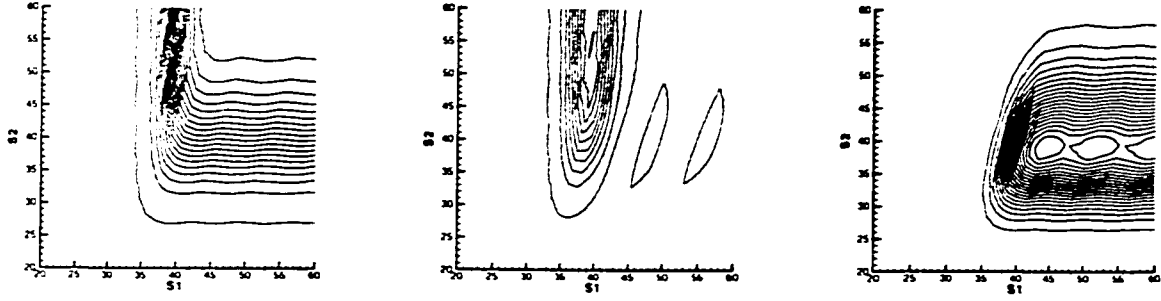


Figure 6.19: Values, deltas with respect to S_1 and deltas with respect to S_2 of a digital call option when $r = 0.05$, $\sigma_{S_1} = 0.10$, $\sigma_{S_2} = 0.30$, $\rho = 0.70$, $T - t^* = 0.25$ and $K = 40$. The solutions were computed with $\Delta t = 0.0025$ using a positive coefficient rotated mesh with 26455 nodes.

jaggedness in the contours computed using the edge-swapped and rotated meshes persisted even if the the initial condition was smoothened using projection (6.1).

The final problem we examined was pricing a quarter year European discrete double barrier call option on the maximum of two assets, where the barrier is applied in time increments of 0.025. For the sample pricing problem $K = 40$, and $H_{lower} = 30$ and $H_{upper} = 50$. Like the digital option, there are discontinuities at maturity. Moreover, each application of the barrier introduces discontinuities into the solution.

Contour plots of solutions computed using regular, edge-swapped and rotated meshes are contained in Figures 6.20, 6.21 and 6.22, respectively. Again we see that the regular mesh produced the smoothest contours, while the edge-swapped mesh produced the worst contours.

Note that when using a rotated mesh, one cannot generally place nodes such that they line up with the barrier. This again highlights the fact that if ensuring

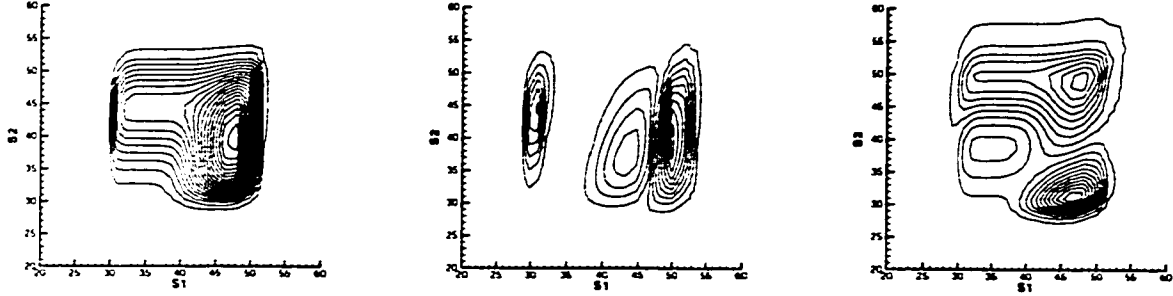


Figure 6.20: *Level curves of values, deltas with respect to S_1 and deltas with respect to S_2 of a European discrete double barrier call option on the maximum of two assets when $r = 0.05$, $\sigma_{S_1} = 0.10$, $\sigma_{S_2} = 0.30$, $\rho = 0.70$, $T - t^* = 0.25$ and $K = 40$. The barrier is applied in time increments of 0.025, and $H_{\text{lower}} = 30$ and $H_{\text{upper}} = 50$. The solutions were computed with $\Delta t = 0.0025$ using a regular mesh with 26569 nodes.*

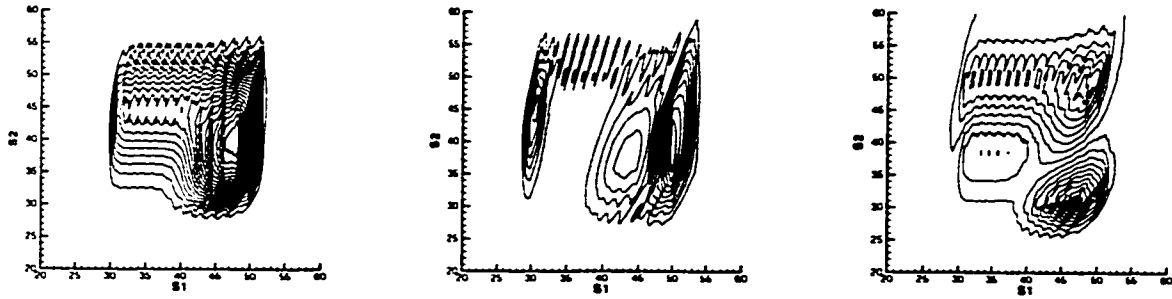


Figure 6.21: *Values, deltas with respect to S_1 and deltas with respect to S_2 of a European discrete double barrier call option on the maximum of two assets when $r = 0.05$, $\sigma_{S_1} = 0.10$, $\sigma_{S_2} = 0.30$, $\rho = 0.70$, $T - t^* = 0.25$ and $K = 40$. The barrier is applied in time increments of 0.025, and $H_{\text{lower}} = 30$ and $H_{\text{upper}} = 50$. The solutions were computed with $\Delta t = 0.0025$ using a positive coefficient edge-swapped mesh with 26569 nodes.*

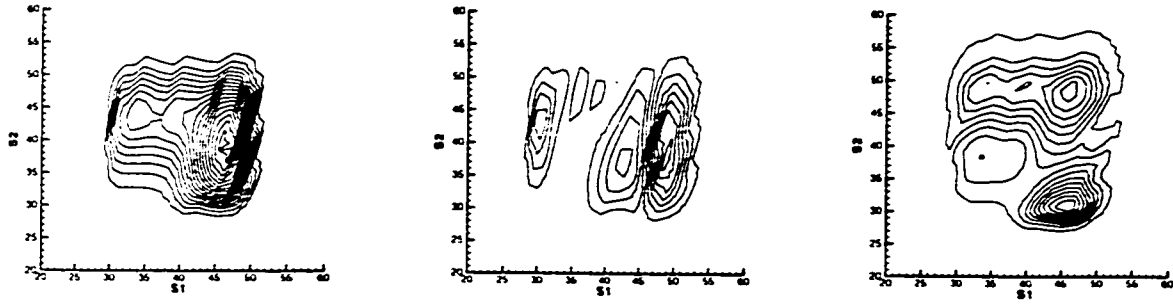


Figure 6.22: *Values, deltas with respect to S_1 and deltas with respect to S_2 of a European discrete double barrier call option on the maximum of two assets when $r = 0.05$, $\sigma_{S_1} = 0.10$, $\sigma_{S_2} = 0.30$, $\rho = 0.70$, $T - t^* = 0.25$ and $K = 40$. The barrier is applied in time increments of 0.025, and $H_{\text{lower}} = 30$ and $H_{\text{upper}} = 50$. The solutions were computed with $\Delta t = 0.0025$ using a positive coefficient rotated mesh with 26455 nodes.*

positive coefficients is used as a criterion for constructing the mesh, other aspects of the problem, such as, the curvature of U or constraints on the solution may have to be ignored.

6.3 Summary

Although the computational domains for option pricing problems are typically rectangular, irregular triangular meshes are nonetheless advantageous because they can be used to improve computational efficiency. As the results demonstrated, sufficiently accurate solutions can be obtained using irregular meshes in an order of magnitude less time than is required when meshes which are analogous to the finite difference grids typically employed in finance are used.

The results in this chapter appear to indicate that the finite volume discretiza-

tion (3.10) has a second-order rate of convergence. Furthermore, the results suggest that the flux limiting scheme (4.20) can handle convection dominated problems without introducing excessive amounts of augmenting diffusion. For problems that were not convection dominated, the flux limiting scheme and central weighting produced comparable results if the control volumes were constructed using perpendicular bisectors. The use of centroid control volumes appears to increase the amount of numerical diffusion introduced by the flux limiting scheme. The results suggest that perpendicular bisector control volumes should be used whenever possible.

As shown in Chapter 5, it is not possible, in general, to construct meshes that will ensure that discretizations of the diffusion term will produce nonnegative coefficients. In those cases where positive coefficient meshes can be constructed, the numerical results suggest that solutions computed using such meshes are no better than solutions computed with meshes that do not ensure positive coefficients. Moreover, poor interpolation of the price and deltas resulted when positive coefficient meshes were used. The poor results appear to be due to the fact that the positive coefficient meshes do not place nodes in such a way as to capture rapid changes in the solution. Since, lattice methods are forms of finite difference/element schemes (as shown in Chapter 5), we conjecture that positive probability lattice schemes will have similar problems.

Chapter 7

Conclusions

We have developed a general finite volume framework for two-factor PDE option pricing models. In order to ensure generality, a nonconservative finite volume method was formulated. This method is specifically designed to handle degenerate equations (especially at boundaries) and convection dominated situations.

It is often the case that option pricing problems are posed without explicit boundary conditions. At large (but finite) values of the independent variables in the computational domain, asymptotic forms can usually be used to specify artificial boundary conditions. Since the nodes where these conditions are imposed are usually far from regions of interest, any errors in the asymptotic form will typically have a negligible effect. However, inappropriate boundary conditions near zero may prevent convergence to the true solution in the area of interest because of its proximity to these conditions. It is often the case that on the boundary near the origin, the domain of dependence of the underlying PDE is on the interior domain and boundary. Hence, an appropriate discretization will not need to impose

conditions on these portions of the boundary. The nonconservative finite volume approach presented in this work handles such boundary points correctly.

To ensure the robustness of the framework, the prevention of spurious oscillations caused by convection dominance was addressed in this work. In the case of the pure convection problem, the FVM, through the use of a positive coefficient scheme satisfies discrete local maximum and minimum principles on an unstructured mesh. The positive coefficient scheme employs a combination of central weighting and a flux limiter, which is possible since the equations are nonconservative. In the case of the complete problem with convection and diffusion, the scheme minimizes the use of the flux limiter. Thus, the amount of additional numerical diffusion is reduced. Furthermore, this method is a compact scheme, where the nonzero structure of the Jacobian matrix is no different from that obtained when central weighting or first-order upstream weighting is used. To the best of our knowledge, a compact positive coefficient scheme has not been previously developed.

The numerical examples presented in this work indicate that the use of the flux limiting scheme is superior to central weighting for degenerate or convection dominated cases. On the other hand, when there is enough diffusion in the problem so that central weighting can be used, both methods give comparable results, particularly if the control volumes are constructed using perpendicular bisectors. However, if centroid control volumes are used, then there is some degradation in accuracy when the flux limiting scheme is employed. To summarize, if a Delaunay triangulation can be constructed, then the flux limiting scheme is a robust method which produces good results in degenerate cases, and is comparable to central weighting

when central weighting can be used.

Although the computational domains for option pricing problems are typically rectangular, accurate solutions are only required in a small subregion near the exercise price. As demonstrated in the numerical examples, an unstructured grid is very useful in such cases, since more nodes can be inserted near the region of interest without introducing additional nodes elsewhere in the computational domain.

We used the finite volume discretization to investigate the effect of negative contributions to coefficients from the discretization of the diffusion term. For any given set of nodes, a mesh can be constructed through edge swapping that will ensure that all coefficients in a finite volume discretization are nonnegative in the presence of correlation when the diffusion tensor is constant. However, an edge-swapped mesh will often contain stretched or skewed elements. In order to avoid such elements or when using standard finite differences, one needs to rotate the coordinate system.

It was shown that when meshes that allow negative coefficients are used, discretizations will approximately satisfy discrete maximum and minimum principles as the mesh size parameter approaches zero.

For the pricing problems considered in this work, it was demonstrated that meshes which ensured nonnegative coefficients produced poor interpolation results compared to regular meshes which did not ensure nonnegative coefficients. This may be of concern when certain jump conditions are present, for example, when there are discrete dividends. The poor interpolation results appear to be due to the fact that ensuring positive coefficients ignores the curvature of the solution, as

well as other aspects of the pricing problem.

The results suggest that meshes where the edges are aligned with features of the payoff function (which may not result in positive coefficients) will produce solutions of high quality. It is in fact fortunate that it does not appear necessary to enforce the positive coefficient condition for option pricing problems, because it is generally not possible to produce a positive coefficient discretization for a given set of nodes when the diffusion/volatility tensor is nonconstant.

This work has also shown that the binomial method and the two-dimensional lattice scheme in [34] are in fact equivalent (to order Δt) to a known finite difference scheme and a finite element scheme, respectively. Hence, it can be proved that the binomial method converges to the solution of the one-dimensional Black-Scholes equation without appealing to probabilistic arguments. As well, the lattice scheme in [34] has positive probabilities (coefficients), in part, because the scheme uses skewed meshes when correlation is not zero.

The framework developed in this work differs from the approach that is often employed in the finance literature, where a separate numerical technique is developed for each class of option pricing model. The discretization method used in this work typically does not require special knowledge of the details of any particular option pricing model. The general approach allows for the development of software that isolates model specifications. For example, in an object-oriented implementation, model classes can be developed which specify the velocity and diffusion tensors, boundary conditions and constraints. Thus, different options can be priced by simply constructing new classes, no other modifications to the software

are required.

7.1 Suggestions for Future Research

There are two natural future research avenues arising from this work:

- The convergence of the nonconservative finite volume discretization of diffusion has only been demonstrated pointwise through numerical experimentation. An analytic proof of convergence would be desirable.
- It would be natural to extend the two-factor discretization to three factors. It is likely that the properties of the two-dimensional discretization will hold in three dimensions. Furthermore, the gains in efficiency achieved by using irregular meshes in conjunction with the compact positive coefficient scheme will be even more dramatic in three dimensions.

Bibliography

- [1] K. I. Amin. On the Computation of Continuous Time Option Prices Using Discrete Approximations. *Journal of Financial and Quantitative Analysis*, 26(4):477–495, 1991.
- [2] P. Arminjon and A. Dervieux. Construction of TVD-like Artificial Viscosities on Two-Dimensional Arbitrary FEM Grids. *Journal of Computational Physics*, 106:176–198, 1993.
- [3] O. Axelsson and V. A. Barker. *Finite Element Solution of Boundary Value Problems*. Academic Press, Inc., Toronto, 1984.
- [4] G. Barles. Convergence of Numerical Schemes for Degenerate Parabolic Equations Arising in Finance Theory. In L. C. G. Rogers and D. Talay, editors, *Numerical Methods in Finance*. Cambridge University Press, 1997.
- [5] J. Barraquand and T. Pudet. Pricing of American Path-Dependent Contingent Claims. *Mathematical Finance*, 6(1):17–51, 1996.

- [6] T. J. Barth. Aspects of Unstructured Grids and Finite-Volume Solvers for the Euler and Navier-Stokes Equations. In *Lecture Series 1994-05: Computational Fluid Dynamics*. von Karman Institute for Fluid Dynamics, 1994.
- [7] T. J. Barth and J. A. Sethian. Numerical Schemes for the Hamilton-Jacobi and Level Set Equations on Triangulated Domains. *Journal of Computational Physics*, 145:1–40, 1998.
- [8] F. Black and M. Scholes. The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 81:637–654, 1973.
- [9] M. Blunt and B. Rubin. Implicit Flux Limiting Schemes for Petroleum Reservoir Simulation. *Journal of Computational Physics*, 102:194–210, 1992.
- [10] P. P. Boyle. A Lattice Framework for Option Pricing with Two State Variables. *Journal of Financial and Quantitative Analysis*, 13(1):1–12, 1988.
- [11] A. Brandt. Generalized Local Maximum Principles for Finite-Difference Operators. *Mathematics of Computation*, 27(124):685–718, 1973.
- [12] M. J. Brennan and E. S. Schwartz. Convertible Bonds: Valuation and Optimal Strategies for Call and Conversion. *Journal of Finance*, 32(5):1699–1715, 1977.
- [13] M. J. Brennan and E. S. Schwartz. Finite Difference Methods and Jump Processes Arising in the Pricing of Contingent Claims: A Synthesis. *Journal of Financial and Quantitative Analysis*, 13:461–474, 1978.
- [14] M. J. Brennan and E. S. Schwartz. Analyzing Convertible Bonds. *Journal of Financial and Quantitative Analysis*, 15(4):907–929, 1980.

- [15] P. Chalasani, S. Jha, F. Egriboyun, and A. Varikooty. A Refined Binomial Lattice for Pricing American Asian options. 1998. 8th Annual Derivatives Securities Conference, Boston.
- [16] W. Cheung and I. Nelken. Costing the Converts. *RISK*, 7(7):47–49, 1994.
- [17] P. G. Ciarlet. *The Finite Element Method for Elliptic Problems*. North-Holland Publishing Company, New York, 1978.
- [18] J. C. Cox, J. E. Ingersoll, and S. A. Ross. A Theory of the Term Structure of Interest Rates. *Econometrica*, 53(2):385–407, 1985.
- [19] J. C. Cox, S. A. Ross, and M. Rubinstein. Option Pricing: A Simplified Approach. *Journal of Financial Economics*, 7:229–263, 1979.
- [20] M. G. Crandall, H. Ishii, and P. L. Lions. User’s Guide to Viscosity Solutions of Second Order Partial Differential Equations. *Bulletin (New Series) of the American Mathematical Society*, 27(1):1–67, 1992.
- [21] M. G. Crandall and P. L. Lions. Two Approximations of Solutions of the Hamilton-Jacobi Equations. *Mathematics of Computation*, 43:1–19, 1984.
- [22] E. F. D’Azevedo and R. B. Simpson. On Optimal Interpolation Triangle Incidences. *SIAM Journal on Scientific and Statistical Computing*, 10(6):1063–1075, 1989.
- [23] M. A. H. Dempster and J. P. Hutton. Pricing American Stock Options by Linear Programming. *Mathematical Finance*, 9(3):229–254, 1999.

- [24] M. A. H. Dempster, J. P. Hutton, and Richards D. G. LP Valuation of Exotic American Options Exploiting Structure. *The Journal of Computational Finance*, 2(1), 1998.
- [25] J. N. Dewynne and P. Wilmott. A Note on Average Rate Options with Discrete Sampling. *SIAM Journal on Applied Mathematics*, 55(1):267–276, 1995.
- [26] D. Duffie. *Dynamic Asset Pricing Theory*. Princeton University Press, Princeton, New Jersey, 1992.
- [27] W. H. Fleming and H. M. Soner. *Controlled Markov Processes and Viscosity Solutions*. Springer-Verlag, New York, 1993.
- [28] Bank for International Settlements. The Global OTC Derivatives Market at End-December 1998. Press release, June 2, 1999.
- [29] P. A. Forsyth. A Control Volume Finite Element Approach to NAPL Groundwater Contamination. *SIAM Journal on Scientific and Statistical Computing*, 12(5):1029–1057, 1991.
- [30] P. A. Forsyth and H. Jiang. Nonlinear Iteration Methods for High Speed Laminar Compressible Navier-Stokes Equations. *Computers & Fluids*, 26:249–268, 1997.
- [31] P. A. Forsyth, K. R. Vetzal, and R. Zvan. Convergence of Lattice and PDE Methods for Asian Options. Working paper, University of Waterloo.

- [32] P. A. Forsyth, K. R. Vetzal, and R. Zvan. A Finite Element Approach to the Pricing of Discrete Lookbacks with Stochastic Volatility. *Applied Mathematical Finance*, 6:87–106, 1999.
- [33] G. E. Forsythe and W. R. Wasow. *Finite-Difference Methods for Partial Differential Equations*. John Wiley & Sons, Inc., New York, 1967.
- [34] B. Gao. Convergence Rate of Option Prices from Discrete- to Continuous-Time. Working paper, University of North Carolina at Chapel Hill, 1997.
- [35] I. I. Gihman and A. V. Skorohod. *Stochastic Differential Equations*. Springer, Berlin, 1972.
- [36] J. B. Goodman and R. J. LeVeque. On the Accuracy of Stable Schemes for 2D Scalar Conservation Laws. *Mathematics of Computation*, 45(171):15–21, 1985.
- [37] A. Harten, J. M. Hyman, and P. D. Lax. On Finite-Difference Approximations and Entropy Conditions for Shocks. *Communications on Pure and Applied Mathematics*, 29:297–322, 1976.
- [38] S. Heston and G. Zhou. On Rate of Convergence of Discrete-Time Contingent Claims. 1997. Working paper, Olin School of Business, Washington University.
- [39] T. S. Y. Ho and D. M. Pfeffer. Convertible Bonds: Model, Value Attribution, and Analytics. *Financial Analysts Journal*, September/October:35–44, 1996.
- [40] J. Hull. *Options, Futures, and Other Derivative Securities*. Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1993. Second Edition.

- [41] J. Hull and A. White. Valuing Derivative Securities Using the Explicit Finite Difference Method. *Journal of Financial and Quantitative Analysis*, 25:87–100, 1990.
- [42] J. Hull and A. White. Efficient Procedures for Valuing European and American Path-Dependent Options. *The Journal of Derivatives*, 1 (Fall):21–31, 1993.
- [43] J. Hull and A. White. Numerical Procedures for Implementing Term Structure Models II: Two-Factor Models. *The Journal of Derivatives*, 2(Winter):37–48, 1994.
- [44] J. Imai. An Alternative Lattice Approach for Multidimensional Geometric Brownian Motions. Working paper, Tokyo Institute of Technology, 1997.
- [45] J. E. Ingersoll. A Contingent-Claims Valuation of Convertible Securities. *Journal of Financial Economics*, 4:289–322, 1977.
- [46] A. Jameson. Positive Schemes and Shock Modelling for Compressible Flows. *International Journal for Numerical Methods in Fluids*, 20:743–776, 1995.
- [47] A. G. Z. Kemna and A. C. F. Vorst. A Pricing Method for Options Based on Average Values. *Journal of Banking and Finance*, 113-129:14, 1990.
- [48] D. Kröner. *Numerical Schemes for Conservation Laws*. John Wiley & Sons Ltd. and B. G. Teubner, Toronto, 1997.
- [49] C. L. Lawson. Software for C^1 Surface Interpolation. In J. R. Rice, editor, *Mathematical Software III*, pages 161–193, New York, 1977. Academic Press.

- [50] R. J. LeVeque. *Numerical Methods for Conservation Laws*. Birkhauser Verlag, Basel, 1990.
- [51] J. J. McConnell and E. S. Schwartz. LYON Taming. *Journal of Finance*, 41(3):561–577, 1986.
- [52] R. C. Merton. Theory of Rational Option Pricing. *Bell Journal of Economics and Management Science*, 4 (Spring):141–183, 1973.
- [53] R. A. Nicolaides. Direct Discretization of Planar Div-Curl Problems. *SIAM Journal on Numerical Analysis*, 29(1):32–56, 1992.
- [54] K. G. Nyborg. The Use and Pricing of Convertible Bonds. *Applied Mathematical Finance*, 3:167–190, 1996.
- [55] B. Øksendal. *Stochastic Differential Equations: An Introduction with Applications*. Springer-Verlag, New York, 1995.
- [56] O. A. Oleĭnik. Discontinuous Solutions of Non-Linear Differential Equations. *Uspekhi Mat. Nauk*, 12:3–73, 1957. English translation in *American Mathematical Society Translations (Series 2)* 26:95–172, 1963.
- [57] D. Pooley, P. Forsyth, K. Vetzal, and B. Simpson. Unstructured Meshing for Two Asset Barrier Options. Working paper, University of Waterloo, 1999.
- [58] M. Putti and C. Cordes. Finite Element Approximation of the Diffusion Operator on Tetrahedra. *SIAM Journal on Scientific Computing*, 19(4):1154–1168, 1998.

- [59] C. S. Rafferty, M. R. Pinto, and R. W. Dutton. Iterative Methods in Semiconductor Device Simulation. *IEEE Transactions on Computer Aided Design*, 4(4):462–471, 1985.
- [60] S. Rippa. Long and Thin Triangles can be Good for Linear Interpolation. *SIAM Journal on Numerical Analysis*, 29(1):257–270, 1992.
- [61] P. J. Roache. *Computational Fluid Dynamics*. Hermosa, Albuquerque, New Mexico, 1972.
- [62] L. C. G. Rogers and Z. Shi. The Value of an Asian Option. *Journal of Applied Probability*, 32(4):1077–1088, 1995.
- [63] L. C. G. Rogers and E. J. Stapleton. Fast Accurate Binomial Pricing. *Finance and Stochastics*, 2(1):3–17, 1998.
- [64] T. F. Russell. Time Stepping Along Characteristics with Incomplete Iteration for a Galerkin Approximation of Miscible Displacement in Porous Media. *SIAM Journal on Numerical Analysis*, 22(5):970–1013, 1985.
- [65] V. Selmin and L. Formaggia. Unified Construction of Finite Element and Finite Volume Discretizations for Compressible Flows. *International Journal for Numerical Methods in Engineering*, 39:1–32, 1996.
- [66] J. Smoller. *Shock Waves and Reaction-Diffusion Equations*. Springer-Verlag, New York, 1994. Second edition.

- [67] S. Spekreijse. Multigrid Solution of Monotone Second-Order Discretizations of Hyperbolic Conservation Laws. *Mathematics of Computation*, 49(179):135–155, 1987.
- [68] R. R. Stulz. Options on the Minimum or the Maximum of Two Risky Assets. *Journal of Financial Economics*, 10:161–185, 1982.
- [69] A. Sulem. Portfolio Selection with Transaction Costs. In L. C. G. Rogers and D. Talay, editors, *Numerical Methods in Finance*. Cambridge University Press, 1997.
- [70] P. K. Sweby. High Resolution Schemes Using Flux Limiters for Hyperbolic Conservation Laws. *SIAM Journal of Numerical Analysis*, 21(5):995–1011, 1984.
- [71] V. Thomée. Some Convergence Results for Galerkin Methods for Parabolic Boundary Value Problems. In C. de Boor, editor, *Mathematical Aspects of Finite Elements in Partial Differential Equations*. Academic Press, 1974.
- [72] J. Thuburn. TVD Schemes, Positive Schemes, and the Universal Limiter. *Monthly Weather Review*, 125:1990–1993, 1997.
- [73] B. van Leer. Towards the Ultimate Conservative Difference Scheme. II. Monotonicity and Conservation Combined in a Second-Order Scheme. *Journal of Computational Physics*, 14:361–370, 1974.
- [74] O. Vasicek. An Equilibrium Characterization of the Term Structure. *Journal of Financial Economics*, 5:177–188, 1977.

- [75] L. B. Wahlbin. A Remark on Parabolic Smoothing and the Finite Element Method. *SIAM Journal on Numerical Analysis*, 17(1):33–38, 1980.
- [76] L. Wang. Numerical Issues in Option Pricing. Master's essay, University of Waterloo, 1998.
- [77] P. Wilmott, J. Dewynne, and J. Howison. *Option Pricing: Mathematical Models and Computation*. Oxford Financial Press, Oxford, 1993.
- [78] O. C. Zienkiewicz. *The Finite Element Method*. McGraw-Hill Limited, Toronto, 1977. Third Edition.
- [79] O. C. Zienkiewicz and J. Wu. Automatic Directional Refinement in Adaptive Analysis of Compressible Flows. *International Journal of Numerical Methods in Engineering*, 37:2189–2210, 1994.
- [80] R. Zvan, P. A. Forsyth, and K. R. Vetzal. Penalty Methods for American Options with Stochastic Volatility. *Journal of Computational and Applied Mathematics*, 91:199–218, 1998.
- [81] R. Zvan, P. A. Forsyth, and K. R. Vetzal. Robust Numerical Methods for PDE Models of Asian Options. *The Journal of Computational Finance*, 1(2):39–78, 1998.
- [82] R. Zvan, P. A. Forsyth, and K. R. Vetzal. Discrete Asian Barrier Options. *The Journal of Computational Finance*, 3(1):41–67, 1999.
- [83] R. Zvan, K. R. Vetzal, and P. A. Forsyth. PDE Methods for Pricing Barrier Options. *Journal of Economic Dynamics and Control*, 2000. To appear.