Valuation of Convertible Bonds
Modelling and Implementation

A thesis presented

by

Ana Bermúdez

to

Faculty of Finance

in fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Finance

Cass Business School

London, UK

July 2004

Abstract

The objective of this thesis is to improve the understanding of the models arising in convertible bond (CB) valuation, introduce new models incorporating interest rate and credit risk and develop sophisticated numerical methods to implement those models. We carry out our analysis in the CB market because it is rapidly increasing and yet not enough research has been done to accurately and efficiently price those instruments. Moreover, the complexity of the mathematical models arising in CB valuation make this area of finance a particularly challenging and interesting one to research. Despite concentrating on CB pricing we believe that our work has broader implications. This is because we proposed a very general and flexible framework that could be applied to price any American-style contingent claim in a two-factor setting.

In the first part of the thesis we introduce for the first time in finance the method of characteristics/finite elements combined with a Lagrange multiplier method to solve two-factor pricing models for financial derivatives. To demonstrate the applicability of the approach, we solve a convertible bond model with equity and interest rate risk;w e focus on the consistent and rigorous specification of the model, and fully address its practical implementation.

The second part of the thesis explores how to incorporate credit risk in CB valuation. This is a complex task due to the hybrid nature of CBs and there is no consensus in the literature of whether it has to be done in an equity or an asset based framework. For this reason, we introduce new equity based and asset based models, which include stochastic interest rate, and solve them using the numerical technique developed in the first part. Regarding the equity based approach we propose a unified intensity-based framework of which most existing comparable models are special cases; this allows us to implement and analyze previous models as well as introduce new ones. We find that different models lead to significantly different prices and that it is important to consistently specify the process for the stock price, the recovery value and the holder's rights upon default. The flexibility of the approach enables us to generate a great number of default-recovery scenarios. Regarding the asset based approach, we introduce a new model which has both structural and reduced form features and in which recovery is endogenised. Despite the fact that the state variable is unobserved and a simple capital structure is assumed, the possibility that default can be triggered both exogenously and endogenously at a cash-flow time leads to a more realistic formulation than it can be achieved in an equity based approach. Moreover our endogenised recovery potentially allows a greater ability to estimate recovery values from market data.

We conclude that both models have a great potential to explain empirical CB values and given their advantages and disadvantages can be chosen depending on the purpose and the circumstances.

Acknowledgments

This thesis would not have been possible without the help of my supervisor, Professor Nick Webber. I am also grateful to my friends, colleagues and family who supported me during those three years. I would also like to thank María Nogueiras, Carlos Vázquez and my father, Alfredo Bermúdez, with whom I have been able to discuss a number of the technical issues developed in the thesis. Finally, I gratefully acknowledge the financial support of Cass Business School and Fundación Caixa Galicia.

Contents

Contents .... **v**

Introduction **1**

1 Literature Review **5**

1.1 Pricing of Convertible Bonds 5

1.1.1 Credit Risk Modelling 7

1.1.2 Credit Risk in CB Valuation 13

1.2 Numerical Methods in Finance 22

1.2.1 Lattice versus PDEs 24

1.2.2 PDE Models. Variational Inequalities 29

1.2.3 Existence and Uniqueness of Solution. Convergence of Numerical Schemes 32

1.2.4 Truncation of the Domain. Boundary Conditions 34

1.2.5 Numerical Solution of PDEs: FE versus FD 35

1.2.6 Convection Dominance. The Method of Characteristics 41

1.2.7 Numerical Solution of Free Boundary Problems. Lagrange Multiplier Method 43

2 Two-factor Pricing for a Class of Contingent Claims. **46**

2.1 Modelling Framework 48

2.1.1 Partial Differential Inequalities 49

2.1.2 Variational Formulation 50

2.2 Numerical Solution of the Mixed Problem 54

2.2.1 Dealing with the Free Boundaries: Lagrange-Multiplier Method 55

2.2.2 Time and Space Discretization: a Characteristics/Finite Element Method........ 57

2.3 Conclusions 63

3 A Two-factor Equity Based Convertible Bond Model Ignoring Credit Risk 65

3.1 The Model 67

3.1.1 The Governing Equation 67

3.1.2 Convertible Bond Valuation as a Free Boundary Problem 68

3.1.3 Auxiliary Conditions 71

3.1.4 The Interest Rate Model 72

3.2 The Numerical Solution 77

3.2.1 A Note Regarding the Unilateral Conditions 78

3.2.2 Finite Truncation of the Domain and Boundary Conditions 79

3.2.3 The System of Characteristics 82

3.3 Numerical Results 86

3.3.1 Benchmarking 86

3.3.2 Contract Specifications and Theoretical Convertible Bond Prices 93

3.3.3 Empirical Results 94

3.4 Conclusion 101

4 A Unified Framework for Pricing CBs with Interest Rate and Credit Risk **103**

4.1 The Governing Equation 106

4.2 Splitting Procedures 111

4.2.1 Splitting 1. U_T: asset or nothing call, W_T: cash or nothing put 113

4.2.2 Splitting 2. U_T: equity, W_T: equity premium (put) 115

4.2.3 Splitting 3. U_T: risk premium (warrant), W_T: bond floor 116

4.3 Detailed Specification of the Model 117

4.3.1 The Interest Rate Model 117

4.3.2 The Hazard Rate Process 117

4.3.3 The Recovery Value 118

4.3.4 Conversion Rights upon Default 120

4.3.5 Nested Previous Models 121

4.3.6 The Dual Recovery Model 123

4.4 Analytical Solutions for Bond Convertible just at Expiration 124

4.4.1 Analytical Solution with Recovery of Bond Part 125

4.4.2 Analytical Solution with Recovery of Par 128

4.4.3 Analytical Solution with Recovery of Market Value 129

4.4.4 Analytical Solution with Recovery of Bond and Equity Part 129

4.4.5 Comparison of Nested Models Based on Analytical Solutions 130

4.5 Numerical Solution 135

4.5.1 Solution of Coupled System of Equations 140

4.6 Numerical Results 141

4.6.1 Benchmarking 142

4.6.2 Comparison of Nested Models based on Numerical Solutions 147

4.6.3 Pricing in the Dual Recovery Model 152

4.7 Conclusions 161

5 An Asset Based Model of Defaultable CBs with Endogenised Recovery 164

5.1 An Asset Based Model for Convertible Bond Valuation 167

5.1.1 Detailed Specification of the Model 169

5.1.2 A PDI for a Convertible Bond 174

5.2 The Solution Method 177

5.2.1 Boundary Conditions 178

5.3 Numerical Results 183

5.3.1 Benchmarking 184

5.3.2 The Recovery Specification 186

5.3.3 Exercise Conditions 187

5.3.4 Parameter Deltas and Gammas 192

5.3.5 The Default Specification 192

5.3.6 The Effect of a Stochastic Interest Rate 194

5.4 Conclusions 195

Conclusion 197

Bibliography 201

A Finite Element Calculations 212

A.1 Discretization: Galerkin's method 214

A.2 Lagrange triangular finite elements 216

A.3 Coefficients matrix and independent term 220

A.3.1 Change to the reference element 225

B Solution of the System of Characteristics 231

B.1 Vasicek model 231

B.2 Hull and White Model 234

Introduction

Convertible bonds (CBs) are sophisticated financial instruments playing a major role in the financing of companies. Typically, they are corporate debt securities or structured products that offer investors the right to forgo future coupon and/or principal payments in exchange to a specified number of shares of common stock. This hybrid feature of convertible bonds provides investors with the downside protection of ordinary bonds and the upside return of equities and fund managers with asset allocation strategies that take advantage of both fixed-income and equity markets. Despite the obvious economic importance of this product, relatively little research has been done to accurately and efficiently price those instruments. In this thesis we attempt to provide a better understanding of the mathematical models arising in convertible bond valuation, together with the numerical methods required in the case that a closed form solution is not available. Moreover, we propose a general pricing framework, which can be made market consistent, and provides a balanced trade-off between speed and accuracy.

This class of financial products has been chosen due to its complexity and because neither the mathematical model for its valuation nor the numerical techniques for its solution are yet standardized.

This thesis is in two parts. The first part presents a general approach for solving two-factor partial differential inequalities (PDI) arising in option-pricing problems with American-style embedded options. The method of characteristics and finite elements is proposed for time and space discretization respectively, together with a Lagrange multiplier method to deal with the free-boundary problem arising from inequality constraints on the solution such as: early-exercise opportunities, conversion provisions, call and put provisions. As an application of the proposed numerical approach, the pricing of callable, putable convertible bonds with stochastic interest rates is carried out.

The second part of the thesis is concerned with the modelling of credit risk in convertible bond valuation. This is a complex task given the hybrid nature of convertible bonds. The literature is divided on whether default risk should be incorporated in an equity based or an asset based framework, given that both approaches offer some advantages and disadvantages. For those reasons, we introduce two different models: an equity based and an asset based model. Both are implemented using the numerical technique proposed in the first part.

The thesis has contributions from four original papers, two of which have been published. The first paper, published in Mathematical Models and Methods for Applied Sciences, develops a general methodology to solve partial differential inequalities arising in the valuation of financial derivatives. The other three papers develop different two-factor models for valuing of CBs that are solved numerically using the tailored characteristics/finite elements discretization, coupled with a Lagrange multiplier method for free boundaries proposed in the first paper.

In the second paper, published in the Journal of Economic Dynamics and Control, a two-factor PDE pricing model for convertible bonds is solved. The two factors are the stock price and the interest rate. The model fits the observed term structure, calibrates the volatility parameters to market data and allows for correlation between the state variables.

As an empirical exercise we compare prices of actual market issues with prices forecasted by the model, leading to very promising results.

The third paper incorporates credit risk into the two-factor pricing framework for convertible bonds proposed in the second paper. This paper specifies and implements an intensity-based default risk model for convertible bonds in a "two and a half" factor setting. The factors are the stock price and the interest rate, together with default risk. We model the hazard rate as a deterministic function of the stock, the interest rate and time. We account explicitly for the stock price behaviour and the CB holder's rights in the event of default as well as the recovery value on the bond. Most comparable existing models are special cases of this general setting. We find that different models lead to significantly different convertible bond values. We also introduce new models for the recovery value. We describe and implement an algorithm to solve the coupled system of partial differential inequalities arising from the model. In a variational formulation we discretize using characteristics and finite element methods. An iterative algorithm is applied over the discretized problem to deal with free boundaries. We benchmark using special cases for which an analytical solution is available. We study the convergence of the numerical method. We compare the consequences of different model specifications, including assumptions about the hazard rate, the recovery value and the stock price behaviour.

Finally, the fourth paper proposes a two-factor model for CBs in which the issuer may default, and where the state variables are the firm's asset value and an interest rate. The CB defaults either at the unpredictable jump time of a counting process, or when the firm is required to make a cashflow to the CB holder. Recovery upon default is endogenised into the model by assuming that the firm can invoke temporary protection against its creditors. As before we solve for CB values using a tailored characteristics/finite elements discretization, coupled with a Lagrange multiplier method for free boundaries. The formulation enables the specification of financially consistent boundary conditions for the convertible bond.

Summarizing, the thesis has two main contributions:

- Introduce new models for valuing CBs incorporating interest rate and default risk.
- Use sophisticated numerical methods to solve the valuation problems. Those methods have not been used before in finance and offer clear advantages over most of the currently used numerical techniques.

Chapter 1
Literature Review

In this Chapter we look at the previous literature on pricing of convertible bonds and on numerical methods commonly used in finance when the valuation models can not be solved analytically.

1.1 Pricing of Convertible Bonds

Convertible bonds are ordinary bonds with an option for the bondholder to convert the bond into common stock at some contractual price at the bond maturity time or some prespecified dates during the life of the option. Like ordinary bonds, the issuer may pay regular coupons to the holder. The conversion option has value to the bond holders, who benefit from future potential growth of the company. Convertible bonds provide investors with the downside protection of bonds and the upside return of equities. Holders of convertible bonds have both a debt claim against the company's assets and an equity claim. For most convertible bonds, the issuer reserves the right to call the bonds. Upon a call, the bondholder may either convert the bond into shares or redeem it at the call price. Some restrictions on the calling privilege may be imposed. There may also be a put feature incorporated into a convertible bond, allowing the holder to sell back the bond to the issuing company in return for a fixed sum. The call feature will decrease the value of a convertible, while the put feature increases it.

Interest in convertibles in the financial markets has increased considerably in recent years. However, neither the mathematical models nor the numerical techniques to solve them are yet standardized.

There are three main issues on the modelling side: whether the stock value or the firm value is the main underlying factor; whether there are additional stochastic factors, such as an interest rate; how default is modelled and what happens upon default to the state variables, the CB holders' rights and the convertible value.

The early models of convertible bonds (Ingersoll (1977*a*) and Brennan and Schwartz (1977)) follow Merton (1973) in using the value of the firm with geometric Brownian motion as the sole state variable. Brennan and Schwartz (1980) and more recently Nyborg (1996) and Carayannopoulos (1996) include in addition a stochastic interest rate. Brennan and Schwartz and Nyborg assume the short rate follows a mean reverting log-normal process; Carayannopoulos assumes the short rate follows the Cox, Ingersoll and Ross (1985) model. Default risk is usually incorporated structurally by capping payouts to the bond by the value of the firm.

Recent literature, on the other hand, mainly uses the stock price as a state variable and either ignores credit risk (Zhu and Sun (1999), Epstein, Wilmottt and Haber (2000), Barone-Adesi, Bermúdez and Hatgioannides (2003), Bermúdez and Nogueiras (2004)), incorporates it via a credit spread (McConnell and Schwartz (1986), Cheung and Nelken (1994), Ho and Pfeffer (1996)) or models it in a reduced form setting as an exogenously specified default process (see Duffie and Singleton (1999)). However, some authors have pointed out (see Schonbucher (2003)) that given the hybrid nature of convertibles, asset based models are the right class to consider in order to account for credit risk. Arvanitis and Gregory (2001) implement and compare both type of models for CB valuation. Bermúdez and Webber (2004) (see Chapter 5) propose an asset based model that incorporates both endogenous and exogenous default, as well as endogenised recovery.

In the equity based approach most authors use a single factor model, although some allow interest rates to be stochastic. The Vasicek (1977) or else the extended Vasicek (Hull and White (1990)) models are used by Epstein, Haber and Wilmott (2000), Barone-Adesi, Bermúdez and Hatgioannides (2003), Bermúdez and Nogueiras (2004), and Davis and Lischka (2002). Ho and Pfeffer (1996) use the Black, Derman and Toy (1990) model; and Zvan, Forsyth and Vetzal (1998*a*) and Yigitbasioglu (2002) use the Cox, Ingersoll and Ross (1985) model. Cheung and Nelken (1995) adopt the model developed by Kalotay, Williams and Fabozzi (1993).

Very few authors model the hazard rate stochastically (Davis and Lischka (2002), Arvanitis and Gregory (2001)). However from the implementation point of view, stochastic hazard rates offer the same complexity as stochastic interest rates, given that the dynamics for both process are often very similar, and their role in the valuation PDE is analogous (Duffie and Singleton (1999)). Most recent papers model the hazard rate as a deterministic function of the state variables (also called a quasi-factor or half factor).

1.1.1 Credit Risk Modelling

In general, credit risk models fall into two main categories, structural and reduced form.

Structural Models

In structural models the state variable is usually the value of the firm or firm asset value, which moves randomly. All claims on the firm's value are modelled as derivative securities with the firm value as underlying.

Default occurs when the value of the firm hits or crosses a boundary, the barrier level. It is necessary to specify the process for the firm value, the location of the barrier, and the form and amount of recovery upon default. This approach was introduced by Black-Scholes (1973) and Merton (1973), who allow default only at maturity. Black and Cox (1976) relax this assumption allowing default prior to maturity.

The main advantage of these models are: (1) they provide a link between the equity and debt instruments issued by a firm which may be necessary for example in the valuation of CBs and callable bonds, (2) they can be used, at least in theory, to optimize the capital structure and, (3) default risk is endogenised and measured based upon the share price and fundamental data only.

The main disadvantage of structural models is that the firm value is unobservable and often difficult to model. Specially the volatility of the firm value is hard to estimate. Also, models become too complex for reasonable capital structures. Finally, they are not well suited for pricing and hedging of credit instruments.

Most authors assume the firm follows a geometric Brownian motion and interest rates are deterministic (Merton (1974) , Geske (1977) , Hull and White (1995) , Nielsen, Saa-Requejo and Santa-Clara (1993)). A problem with these models is that the implied credit spreads are very low and tend to zero as time to maturity approaches zero. To overcome this Schonbucher (1996) and Zhou (1997) introduce jumps into the firm value process.

Longstaff and Schwartz (1995) give a semi-closed form solution for defaultable bonds, when interest rates follow Vasicek model and may be correlated with the firm value. They assume the firm value follows a GBM, that the barrier level is constant and that the recovery value is a fraction of the bond principal.

Briys and de Varenne (1997) also get a closed form solution for risky bonds assuming the same process as Longstaff and Schwartz for the firm value and the stochastic interest rate. The difference is that they model the barrier level as a fraction of the discounted face value and the recovery as a fraction of the barrier level.

Reduced Form Models

In reduced form models default is exogenous, occurring at the first jump time τ of a counting process, N_t, with jump intensity λ_t. The main issues in reduced form models are the specification of processes for the riskless short rate r_t, the hazard rate λ_t, and the recovery value.

This approach has been followed, among others, by: Ramaswamy and Sundaresan (1986), Jarrow and Turnbull (1995), Hull and White (1995), Duffie and Singleton (1994), (1997), (1999), Lando (1998) etc.

The random time of default τ is a stopping time. The hazard rate is the local arrival probability of the stopping time per time.

The hazard rate is closely related to the credit spread implied by the market price of bonds issued by a defaultable obligor and the price of similar default-free instruments For this reason the terms modelling of the credit spread and modelling of the hazard rate are interchangeable. However, only if default and interest rates are uncorrelated, hazard rates, or equivalently default probabilities, can be obtained directly from market credit spreads. The fundamental relation between defaultable zero-coupon bond prices and implied default probabilities appears for example in Das (1998) and Duffie (1999).

The simplest specification for the counting process N is a Poisson process. A Poisson process with intensity $\lambda > 0$ is a non-decreasing, integer-valued process with initial value $N(0) = 0$ whose increments are independent and satisfy, for all $0 \leq t < T$

$$P[N(T) - N(t) = n] = \frac{1}{n!}(T-t)^n \lambda^n \exp(-(T-t)\lambda).$$

When default is triggered by a Poisson process with constant intensity, the term structure of spreads is flat and it does not change over time. This can be improved by modelling the hazard rate as a function of time, leading to a so called inhomogeneous Poisson process. In that case it is possible to fit a given term structure of defaultable bond prices.

A more realistic specification will allow the credit spread to move randomly; a Poisson process with stochastic intensity is a so-called Cox process. Different specifications have been proposed for the risk-free interest rate and the hazard rate. Ideally, both should be stochastic, correlated and maybe follow a multifactor model. On the other hand, it should be possible to work out closed-form solution for simple instruments, such that the model can be easily calibrated.

The implied term structure of credit spreads can be used to price many simple credit derivatives, but not convertible bonds. In practice, the calibration of the credit spread curve is not easy due to the lack of data. Usually there will be more than one curve consistent with the market price of the calibration instruments. To guarantee uniqueness of the spread curve either a parametric form with few parameters is given, or extra smoothness conditions are imposed. The methods for calibration of the spread curve are just adaptation of the methods to calibrate risk-free interest rate curves. Authors who calibrate the credit spread imposing a parametric form are for example Nelson and Siegel (1987) and Svensson (1994), (1995). Vasiceck and Fong (1982), Adams and van Deventer (1994), and Waggoner (1997) use methods based on splines, which impose smoothness.

Recovery Modelling

Regarding the recovery of defaultable claims, many models (as reviewed by Schonbucher (2003) and Bielecki and Rutowski (2002)) have been proposed in the literature: recovery of treasury (RT), recovery of par (RP) multiple defaults (MD), recovery of market value (RMV), zero recovery (ZR) and stochastic recovery. As Schonbucher points out they try to model the value of the settlement, not the actual outcome of the bankruptcy process, they just measure this outcome in different units.

The RT is very convenient from the computational point of view. The reason is that the price of a defaultable issue under RT is a weighted average of the default-free instrument and the price under zero recovery, which is usually easy to compute. However the RT can lead to unrealistic shapes of spread curves, and lead to recoveries above 100%. The RP and

RMV model are similar for issues close to part. The RMV is more consistent for the pricing of credit risk derivatives, but it does worse in pricing downgraded and distressed debt. The RMV is very elegant, in the sense that pricing of financial instruments can be done by discounting with the adjusted defaultable rate $r + \lambda(1 - R)$, where λ is the hazard rate and R is the recovery rate. In the RP pricing is more complicated. Both models are suited for the calibration of the implied credit spreads, although in the RMV it is not possible to separate the calibration of the hazard rate, λ, and the loss rate, $(1 - R)$. The RMV cannot be used with firm value models, whereas the RP can be used in intensity based and in firm value models. Finally, the intuition behind both models is different: the RMV is motivated by the idea of reorganization and renegotiation of debts; the RP is motivated by the idea of bankruptcy proceedings under an authority ensuring strict relative priority.

In general recovery rates are difficult to imply from market prices. Bakshi, Madan and Zhang (2001) do so for the RP model and a modification of the RT under restrictive assumptions for the default intensities. A more general calibration algorithm has been proposed by Unal, Madan and Guntay (2001).

In absence of implied recovery rates, historical recovery rates could be used as benchmark. In that case they need to be adjusted for risk premia. Hamilton, Gupton and Berthault (2000) study the recovery rates and losses given by default of defaulted public debt from 1981 to 2000. They show that recovery rates have extremely high variability across different default events. They try to explain recovery rates using variables that are endogenous to the defaulted obligor or the defaulted bond issue. They find some empirical indicators that explain this variability, but much of the uncertainty cannot be explained. Altman, Resti and

Sirone (2001) consider a different important issue, the systematic dependence of recovery rate across different defaults. There are numerous statistical models to predict recovery rates based on firm-specific variables, but their explanatory power is generally not too high. For pricing, studies using market prices should be more relevant.

1.1.2 Credit Risk in CB Valuation

Since the important work by Ingersoll (1977*a*), (1977*b*) and Brennan and Schwartz (1977), the contingent claims approach (Black and Scholes (1973), Merton (1973)) to pricing convertible bonds is the definitive choice. Traditional methods such as "break-even period" analysis, "discount cashflow" analysis and "synthetics" have serious shortcomings as discussed by Cheung and Nelken (1994). As such, the theoretical equilibrium price of a convertible bond is defined as the value that offers no arbitrage opportunity to either the holder or the issuer. Usual provisions such as the possibility of early conversion, callability by the issuer and putability by the holder, make the issuer to follow a call policy (referred to as optimal call) that minimizes the value of a convertible bond, and the investor to follow conversion (referred to as optimal conversion) and redemption (referred to as optimal redemption) strategies that maximize the value of the convertible bond at each point in time. Following the analysis of Black Scholes and Merton, Ingersoll (1977*a*) and Brennan and Schwartz (1977) derive a partial differential equation for the value of a callable convertible bond as a function of the firm value only. They both determine the optimal conversion strategy for investors and the optimal call policy for the issuer via the criterion of dominance, and provide the inequality constraints that must be satisfied by the convertible bond value when these embedded options are considered.

Ingersoll (1977*a*) and Brennan and Schwartz (1977) assume that the firm's value is the underlying stochastic variable that the convertible bond depends upon. Although theoretically very attractive, the existence of senior debt, preferred equity and multiple classes of common equity in a typical firm's capital structure makes the valuation of convertibles in such a context difficult in practice. Furthermore, the availability of credible data on non-publicly traded issues poses serious additional problems.

Brennan and Schwartz (1980) extend their own previous work by allowing for the uncertainty inherent in interest rates. They compare the two-factor with the single factor model and they conclude that CB values under deterministic interest rates are higher than values under stochastic interest rate, although overall the differences are very small. This conclusion has encouraged some authors to use one-factor models. Nevertheless, since convertible bonds have long lifespans, the assumption of a flat term structure is not in general valid.

Following Merton (1973), Ingersoll (1977*a*) and Brennan and Schwartz (1977) assume that the firms' outstanding securities consist solely of common stock and convertible securities and they allow for the possibility that the firm will default either prior or at maturity. When in 1980 they extend their own work by adding the possibility of senior debt in the firm's capital structure, both the conversion and the maturity condition in the model depend upon the value of the senior debt given that the convertibles are no longer outstanding. Therefore when solving for the convertible bond, the value of the senior debt must be found first as the solution of the same PDE as the convertible with the appropriate definition of total rate of cash distribution to the firms security holders, and final and bankruptcy conditions. These is a clear disadvantage of choosing the firm value as the state variable. These structural valuation approaches, as reviewed by Nyborg (1996), account for credit risk but, since they use the total value of the firm as the stochastic variable, involve many unobservable parameters (notably, the volatility of the firm's value instead of the underlying equity) that make them difficult to use.

If we assume that the value of the convertible depends upon the value of the issuer's common stock, credit risk is easily incorporated in a convertibles' model by adding a constant option-adjusted spread or effective credit spread to the riskless interest rate as, for example, in Ho and Pfeffer (1996),[1] McConnell and Schwartz (1986) and Cheung and Nelken (1994). The main drawback of these early papers is that they model credit risk in an ad hoc manner. It is necessary to specify how default is triggered and what happens upon default with the state variable, the holders' rights and the convertible value.

The first authors to have modelled default exogenously in the spirit of reduced form models, are Davis and Lischka (2002) (DL hereafter) and later Takahashi, Kobayahashi and Nakagawa (2001) (TKN hereafter) and Arvanitis and Gregory (2001). They assume that default occurs at the first jump of a Poisson process and they model the intensity of the jump as a deterministic function of the stock price. They assume that upon default the stock price jumps to zero. DL and Arvanitis and Gregory (2001) model the recovery as a constant fraction R of the par value of the bond, whereas in TKN model the recovery

[1] The credit risk of the bond is captured by the option adjusted spread (OAS), which added to the one period Treasury rate, is used to discount the bond's cash flows.

is a fraction of the market value of the bond prior to default. Recently, Andersen and Buffum (2003) implement a model very similar to the one of DL focusing their attention on calibration issues. Hung and Wang (2002) implement in a lattice a model very similar to TKN. Carayannopoulos and Kalimipalli (2003) do an empirical investigation of the TKN model, extended to allow for stochastic interest rates.

However, it can be argued that the approach followed in all the above references, penalizes unnecessarily the credit risk-free equity upside of the convertible bond. At low equity prices, when the conversion option is worth little, the convertible is essentially a pure bond and it is clearly correct to price (i.e. discount cash flows) with the full credit spread of the issuer. However, it is generally held that a company's ability to issue stock is not strongly influenced by its credit rating. Accordingly the value contributed to the bond by its conversion rights should not be subject to the same risky discounting as the fixed payments. The value of a convertible bond has components of different default risk. The equity upside has zero default risk since the issuer can always deliver its own stock. On the other hand, coupon, principal payments and any put provisions depend on the issuer's timely access to the required cash amounts, which crucially are not known in advance, and thus introduce credit risk. A number of modelling choices could be made to achieve this result[2] with different advantage and disadvantages. These are the so called blended discount rate models. In its most simple form, a blended discount rate model uses the full risky rate to discount all cash flows and only the risk free rate for the equity component.[3]

---

[2] A Quantitave Research publication of Deutche Bank (see www.dbquant.com) proposes five alternative blended discount rate models in a one-factor PDE model. They use a constant deteministic credit spread to discount the fixed payments but not the equity value.

[3] In a lattice framework, Connolly (1998) and Phillips (1997) suggest using a different discount factor along

More sophisticated discount rate models, like the probability conversion model of Goldman Sachs (1994) or Tsiveriotis and Fernandes (1998) approach, use a weighted blended discount factor that accounts for the moneyness of the embedded conversion option.

In the Goldman Sachs model the discount rate is equal to $r + (1 - P_c)h$, where P_c is the risk neutral conversion probability, which is a derivative of the underlying stock S. At maturity, P_c is set to 1 if the CB is converted, and is set to 0 otherwise. For any other time, P_c is re-set to 1 if it is optimal to convert the CB. The conversion probability is a derivative on the underlying state variables, and therefore can be found as the solution of a Black-Scholes type PDE or in a lattice (by setting the discounting to zero).

On the other hand, Tsiveriotis and Fernandes (1998) (TF) split the convertible into an equity component and a bond component which are discounted differently to reflect their different credit risk. Yigitbasioglu (2002) extends the TF model by allowing a stochastic interest rate and FX risk.

Tsiveriotis and Fernandes (1998) carry on to define a hypothetical derivative security, the cash-only part of the convertible bond (COCB) that follows the same dynamics as the convertible's value. The resulting valuation equation for the COCB explicitly involves the issuer's credit spread. On the other hand, the part of the value of the convertible bond related to payments in equity is discounted using the risk-free rate.

The holder of the COCB is entitled to all cash flows from the bond part but not any equity cash flows. This means that at time t the value of the COCB, V_{co}, is set to zero if it is optimal to convert. By definition, V_{co} is a function of the underlying stock price and

the tree for nodes where conversion, call or put has been exercised and nodes where the bond is still alive.

the instantaneous interest rate. Therefore V_{co} should follow the same Black-Scholes type differential equation with the discount rate set to be the full risky rate

$$\frac{\partial V_{co}}{\partial t}+\frac{1}{2}\sigma^2 S^2\frac{\partial^2 V_{co}}{\partial S^2}+\rho S\sigma w\frac{\partial^2 V_{co}}{\partial S\partial r}+\frac{1}{2}w^2\frac{\partial^2 V_{co}}{\partial r^2}+\left(rS-D\left(S,t\right)\right)\frac{\partial V_{co}}{\partial S}$$
$$+\left(u-\lambda_r w\right)\frac{\partial V_{co}}{\partial r}-(r+h)V_{co}=0. \tag{1.1}$$

The difference between the CB value V and V_{co} is the share related part of the bond (SOCB) and its price V_{so} naturally satisfies the same differential equation with the discount rate equal to the risk-free rate r

$$\frac{\partial V_{so}}{\partial t}+\frac{1}{2}\sigma^2 S^2\frac{\partial^2 V_{so}}{\partial S^2}+\rho S\sigma w\frac{\partial^2 V_{so}}{\partial S\partial r}+\frac{1}{2}w^2\frac{\partial^2 V_{so}}{\partial r^2}+\left(rS-D\left(S,t\right)\right)\frac{\partial V_{so}}{\partial S}$$
$$+\left(u-\lambda_r w\right)\frac{\partial V_{so}}{\partial r}-rV_{so}=0. \tag{1.2}$$

From the above two equations, the convertible bond price $V=V_{co}+V_{so}$ satisfies

$$\frac{\partial V}{\partial t}+\frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2}+\rho S\sigma w\frac{\partial^2 V}{\partial S\partial r}+\frac{1}{2}w^2\frac{\partial^2 V}{\partial r^2}+\left(rS-D\left(S,t\right)\right)\frac{\partial V}{\partial S}$$
$$+\left(u-\lambda_r w\right)\frac{\partial V}{\partial r}-(r+h)V_{co}-rV_{so}=0. \tag{1.3}$$

Equivalently

$$\frac{\partial V}{\partial t}+\frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2}+\rho S\sigma w\frac{\partial^2 V}{\partial S\partial r}+\frac{1}{2}w^2\frac{\partial^2 V}{\partial r^2}+\left(rS-D\left(S,t\right)\right)\frac{\partial V}{\partial S}$$
$$+\left(u-\lambda_r w\right)\frac{\partial V}{\partial r}-(r+h^*)V=0, \tag{1.4}$$

where $h^* = h\frac{V_{co}}{V}$ is the conversion adjusted credit spread. When the equity price is very high, the convertible is certain to convert, V_{co} is very small and the risk-free rate is used to discount the corresponding cash flows. On the other hand, if the equity price is very low, the convertible bond behaves like an ordinary debt bond and h^* is close to h, which means that the full risky rate is used to discount the cash flows. Therefore TF approach is similar to the probability conversion model with a different technique to adjust the discount rate.

TF use a constant deterministic credit spread. Moreover, they do not specify what happens upon default to the stock price, the convertible bond value, or holder's rights. Ayache, Forsyth and Vetzal (2002), (2003) (AFV hereafter) extend previous literature by proposing a general specification of default in which the stock price jumps by a given percentage η upon default and the issuer has the right either to convert or recover a given fraction R of the bond part of the convertible. The way they define the bond part is different from the original definition of Tsiveriotis and Fernandes. In our understanding, the definition of the bond part in AFV lacks financial intuition and makes any implementation extremely complex. Indeed, even though AFV decompose the convertible bond value (and therefore write the model as a coupled system of equations) they provide most of their numerical results just for two extreme situations in which the model reduces to one equation only.

Summarizing, another important issue on the modelling side is, given the hybrid debt-equity nature of convertibles, whether it is necessary to split its value in order to apply a different credit regime to the debt and equity components. This was the approach taken in the first valuation models, which were pricing convertibles as a portfolio of bonds and warrants. Unfortunately this decomposition is valid only if the bond is convertible only at expiry and there are no other options embedded in the bond. In general, the value of the debt and equity components will be linked, and the valuation problem is to solve a coupled system of equations. How to split the convertible value though, is an open an controversial matter.

Whether or not we split the convertible into bond and equity parts, a key issue is how to model the credit spreads. Some papers use a constant hazard rate despite a body of empirical evidence (Duffie and Singleton (1997) and Schwartz (1998)).[4] Other papers model the hazard rate stochastically but make the interest rate deterministic (Davis and Lischka (2002), Arvanitis and Gregory (2001)).

Most authors model the hazard rate as a deterministic function of the state variables. Many parameterizations could be applied. Table 1.1 shows some specifications that have been used in the literature. Some authors model the hazard rate as a function of the stock price only, and impose negative correlation via a power function or an exponential one. In both specifications the spread is a monotonic decreasing function of the stock price; but only the power function guarantees an infinite hazard rate for zero stock price, which is a desirable property (see Olsen (2002)). Recently Das and Sundaram (2004) have combined an exponential dependency on the interest rate with a power dependency on the stock price, and have added the time to maturity. They calibrate the hazard rate using market prices of CDS and historical data and they test this approach empirically by pricing some real convertible bonds. Their results seem very satisfactory. Andersen and Buffum (2003) are concerned with the simultaneous calibration of the hazard rates and the volatility smiles; they point out the need to make the hazard rate time dependent to avoid mispricing.

---

[4] A tutorial from the Convertible Bonds Research group at Barclays Capital (2002) points out that traditional models perform poorly for volatile, high-yielding, low parity convertibles because the credit spread does not vary with the stock price. They tend to over-estimate gamma and vega and to under-estimate delta.

Davis and Lischka (2002)	$p_t = c + k/S^a$
Olsen (2002)	
Takahasi, Kobayahashi, Nakagawa (2001)	
Ayache Forsyth and Vetzal (2003)	
Arvanitis and Gregory (2001)	$p_t = k \exp\left[-aS_t + d\right]$
Das and Sundaram (2004)	$p_t = k \exp\left[br_t + c(T-t) + d\right]/S^a$

Table 1.1. Models for the hazard rate

To model the credit spread as a function of the state variables is very intuitive and appears to provide realistic valuations, sensitivities and implied parameters, but it does constrain the credit spread to have an explicit relationship with the stock price. In reality, credit spreads contain some randomness. This suggests developing a model in which both stock prices and credit spreads follow separate but correlated random processes, as proposed by Davis and Lischka (2002). As these authors (2002) point out, although there are three sources of uncertainty- stock price, interest rate and credit spread- practitioners tend to avoid more than two factors. Therefore, and in order to reduce the problem to two factors, we will either take the hazard rate to be deterministic and model the interest rate stochastically or we will assume the interest rate to be deterministic and the hazard to be stochastic. From the modelling perspective, both approaches appear to be symmetrical since the stochastic processes assumed for interest rate and hazard rate are in general very similar.

In general, the pricing equation for convertibles cannot be solved analytically, hence numerical methods need to be used. It is often the case that a number of the complex features of convertible bonds are ignored because of the limitations of the adopted numerical scheme or because of the difficulty of implementing and using the model.

Previous work has focused on finite difference schemes (Brennan and Schwartz (1980), McConnell and Schwartz (1986), Tsiveriotis and Fernandes (1998), Nyborg (1996), Epstein and Wilmott (2000), Zhu and Sun (1999), Yigitbasiouglu (2002), Andersen and Buffum (2003)) or lattice methodologies (see for example Cheung and Nelken (1994), Carayannopoulos (1996), Ho and Pfeffer (1996), Philips (1997), Connolly (1998), Davis and Lischka (2002), Arvanitis and Gregory (2001), Takahashi, Kobayahashi and Nakagawa (2001), Hung and Wang (2002), Das, Sundaram (2004)).

The numerical schemes used in previous work present some difficulties. In general, no special treatment is done to account for the convection dominance problem. Moreover, the treatment of the early exercise in the three shapes it appears in CB valuation (call, put and conversion) is almost always explicit, with all consequent inaccuracy problems. Finally, when lattice and explicit FD are used, strong restrictions in the definition of the time and space step need to be imposed, leading to very slow inefficient algorithms for long date instruments in a two-factor setting. In the following Section we examine those problems in depth in the context of numerical methods used in finance for contingent claim valuation.

1.2 Numerical Methods in Finance

In the absence of closed form solutions, there are three numerical techniques that are commonly employed in finance to price derivative products: (i) numerical solution of partial differential equations (PDEs), (ii) lattice methods (binomial and multinomial trees) and (iii) Monte Carlo simulation. Whereas lattice approaches and Monte Carlo start with a particular description of the asset price dynamics, PDE methods focus on the dynamics of the no-arbitrage portfolio, only implicitly considering the asset price dynamics.

Each method has characteristics that make it more appropriate for certain type of instruments and valuation models and less appropriate for others.

The Monte Carlo technique is clearly very powerful and general. It tends to be numerically more efficient than other procedures when there are three or more stochastic variables and provides a standard error for the estimates that are made. Monte Carlo simulation can accommodate complex payoffs and complex stochastic processes and it easily handles path dependent derivatives. Nevertheless, and despite recent developments (see for instance Carr and Yang (1997), (1998), Broadie and Glasserman (1997), Longstaff and Schwartz (1998), Andersen (1999)), it is not easy to implement this method in the case of American-style derivatives. Thus we will concentrate on PDE methods and lattice models, which could handle the early exercise premium.

In this Section we do a review of the numerical techniques that are commonly employed in finance, focusing our attention on the numerical solution of partial differential equations. First we present a comparative analysis of PDE and lattice methods. Then we introduce the concept of variational formulation of a PDE model, and we discuss the issues of existence and uniqueness of a solution, as well as the convergence of numerical schemes. Then we proceed to talk about boundary conditions and the different discretization schemes, namely finite differences (FD) and finite elements (FE). In relation to the discretization scheme we discuss the convection dominance problem and finally we address the numerical solution of free boundary problems.

1.2.1 Lattice versus PDEs

Lattice methods are perhaps the most widely used numerical methods in finance. The popularity of lattice methods can be in part attributed to their intuitiveness (easy to understand) and simplicity (easy to implement), at least for relatively basic derivative pricing problems. These methods assume discrete approximations of the underlying stochastic process. In that case, the pricing models can be written as discrete sets of difference equations, and therefore can be implemented directly on a computer. On the contrary PDE methods assume that the underlying stochastic processes are continuous and the fair price of the contingent claim solves a partial differential equation. Most of these PDEs do not admit simple closed form solution, hence numerical techniques are required. In order to solve a PDE numerically we must replace the problem by one with a finite number of degrees of freedom i.e., reduce the continuous PDE to a discrete set of difference equations that can be solved in a computer. The most common types of discretization are finite differences and finite elements.

The binomial option pricing model was first introduce by Cox, Ross and Rubinstein (1979) (CRR hereafter). Binomial trees provide a convenient way to model the asset price process using a discrete binomial distribution to approximate the normal distribution of log returns assumed in the Black-Scholes analysis. They can be configured in various ways according to different parameter choices for probabilities of up and down movements and incremental price changes. These parameters are chosen to match the volatility and expected return of the underlying asset(s). However, since there are three parameters and we are only trying to match two values there is a free choice for one of them.

Whereas in CRR-type trees centering occurs on the current asset price, the Jarrow Rudd (JR) type tree centers the lattice on the forward price. In the JR tree both the binomial and the lognormal process have the same first two moments for any number of time steps, whereas in the CRR tree the choice of parameters ensures equality of the variance only in the limit. Also JR parameters insure that the probabilities remain always positive whereas CRR approach may lead to negative probabilities for large number of time steps.[5] In any case, convergence of the binomial option pricing formula to the Black-Scholes formula can be guaranteed using the central limit theorem.

Discrete and continuous models are not directly comparable. Among other things, lattice methods provide the price at the initial time just for the current value of the state variable, whereas PDE methods provide the price for any level of the state variables in the computational domain. Nevertheless, it is interesting to notice that the discretized versions of the continuous models can themselves be interpreted as discrete probabilistic models. In fact, already Brennan and Schwartz (1977) show that the explicit finite difference method is equivalent to a binomial lattice approach and the implicit finite differences method corresponds to a multinomial lattice where, in the limit, the underlying variable can move from its value to infinite possible values at next timestep. A proof of this equivalence in a more general setting can be found in Lapeyre, Sulem and Talay (2004).

PDE models are usually more flexible. Although lattices can be adapted to accommodate underlying assets paying discrete cashflows (like dividends or coupons) and time-

[5] A third possible model is due to Tian (1993) who proposes binomial and trinomial trees in which the parameters are derived as unique solution of equation systems where the first two moments of the continuous and discrete asset distributions are matched.

varying parameters (see for example Hull (2002) or Clewlow and Strickland (1998)), these features are more easily handled in a PDE framework (as described for example in Wilmott (1998)). Also hedge sensitivities may be approximated in a binomial tree using finite difference ratios, but in a PDE approach are given as a by-product.

Geske and Shastri (1985) compare binomial and finite difference methods applied to vanilla option pricing models with one stochastic variable. The comparison is made with respect to differences in both the approximation theory and the efficiency of the computation algorithms, although the latter has not been done very rigorously and is not supported analytically. Clearly the discrepancies are sensible to the particular scheme and implementation. But they make one point clear, namely that lattices loose efficiency when dealing with discrete dividends or with American-Style options. Besides, despite of what the authors say, the generalization to two or more factors will change things substantially.

In general PDE models are more easily analyzed than lattice models. The continuous models lead to more manageable discrete models, and there is a well-understood theory of convergence and error analysis for discretization of continuous partial differential equations.

Theoretical proofs of convergence of discrete time models to their continuous time analogues have been frequently addressed in the literature.[6] However little attention has been given to the rate of convergence, which appears to be one of the central properties of a discrete time model. The rate or order of convergence of a discrete model measures the

[6] Amin and Khana (1994) for example, prove convergence for American options prices in the case where the limiting process of the discrete-time model is a diffusion process satisfying convergence with the first two moments and one higher moment.

asymptotic trade-off between speed and accuracy of a numerical method. A key question is how do we choose the number of time steps to achieve a required accuracy or in other words, what is the error incurred as a function of the number of time steps. Clearly an approximation algorithm should have a rate of convergence as great as possible. Knowledge of the order of convergence allows ranking the available discrete time models according to well measurable and quantifiable criteria. Convergence behaviour is also an issue.

Leisen and Reimer (LR) (1996) examine convergence behaviour and convergence speed for the CRR, JR and Tian models. They show that even for the European call, binomial option prices asymmetrically oscillate with changing amplitude around the Black-Scholes solution for increasing tree refinements. Furthermore the error can actually increase with an increase in the number of time steps. They prove that CRR and JR models converge with order one for the European call and need to be seen as equivalent. They also construct new binomial models where the calculated option prices converge smoothly to the Black-Scholes solution and with order two, although they are not able to give a strict proof of the greater order of convergence in line with their theorem.[7] Finally they compare all three models with respect to speed and accuracy following an approach introduced by Broadie and Detemple (1996).[8] Their model achieves the same degree of accuracy 1400 times faster than the standard models. Nevertheless for American options, in which we concentrate most here, their method exhibits only order one, although on average the same

[7] To get rid of the odd-even convergence they build a tree centred on the strike price and to improve accuracy they exploit findings of mathematical literature on normal approximations to the binomial function.

[8] Option prices are calculated for a large sample of random parameters and then the relative standard deviation to the true solution is calculated and compared to computation time with increasing refinement.

accuracy is achieved ten-times faster than previous binomial models. Despite these promising results, LR trees do not seem to be very popular.

Heston and Zhou (2000) extend and contradict somewhat the work of Leisen and Reimer (1996) by characterizing the rate of convergence of discrete-time multinomial models. They show that the rate of convergence depends on the smoothness of the option payoff function, and is much lower than commonly believed because option payoffs are not continuously differentiable. They show that on the standard binomial tree the rate of convergence cannot uniformly be first order but it is at least $1/\sqrt{n}$. They state that although at the current node the solution may still have the $1/n$ rate of convergence that LR claim for CRR, the nonsmoothness of the payoff function can have an impact sufficient to cause the well known oscillatory pattern of the binomial prices at the current node. In summary, they reinforce the idea of Leisen and Reimer that for standard trees not even the rate of convergence is very low for European options but is in general unknown for American and path dependent options.[9]

Binomial methods can be extended to deal with path dependent options and options whose payoff depends on more than one asset. A lot of research has been done in this respect. Boyle, Evnine and Gibss (1989) develop an n-dimensional extension of the binomial method for valuing multivariate contingent claims. Madan, Milne and Shefrin (1989) and He (1990) generalize the CRR model to a multivariate model and in that context show convergence for prices and hedge sensitivities. Regarding path dependency, Hull and White

[9] They propose both a smoothing approach and an adjustment approach such that the resulting equation achieves its maximum convergence rate across all the nodes and the standard Richardson extrapolation can be used. But the numerical results they provide are not very satisfactory, and cannot be supported mathematically.

(1993) modify the original CRR model for the pricing of path dependent exotic options by linear or quadratic interpolation, like Asian options for example. Cheuk and Vorst (1994) present a model where the payoff of a lookback option is itself modelled in a lattice. Alford and Webber (2001) describe how very high order multinomial lattice methods can be constructed and implemented when the SDE followed by the underlying state variable can be solved. Incorporating both a terminal correction and appropriate truncation methods, they conclude that the heptanomial lattice is the fastest and most accurate of the lattices of higher order.

However, when extended to multi-factor or path dependent options, the trees are no longer easy to understand or implement. Besides, convergence properties are not clear. As Zvan, Forsyth and Vetzal (2000) point out, although most authors have limited themselves to illustrating convergence through numerical examples, this does not prove convergence to the correct solution. Zvan, Forsyth and Vetzal (2000) compare convergence of lattice and PDE methods for pricing Asian options. They show that while it is straightforward to prove that PDE methods are convergent, methods like the lattice based Forward Shooting Method of Barraquand and Pudet (1996) and the method proposed by Hull and White (1993) for path dependent exotic options do not converge for some specific problems. Things become even worse when one moves from prices to hedge parameters.

1.2.2 PDE Models. Variational Inequalities

The value of many financial derivative products is conveniently modelled in terms of two factors, or stochastic space variables, and time. Based on the contingent claims analysis developed by Black and Scholes (1973) and Merton (1973), a PDE for the fair price of these derivatives can be obtained. Valuation PDEs for financial derivatives are usually parabolic and of second order. Furthermore, most PDEs in finance are linear, although non-linear cases can appear as well.

The pricing problem is completed by specifying the final condition (the payoff of the contingent claim) and the boundary conditions (at zero and infinity). This makes PDE methods very general, since just minor changes must be done on the implementation in order to price a wide array of different two-coloured options. Besides, the approach is very flexible to incorporate almost any contract specification; PDE models can easily handle discrete cashflows (jump conditions arising from dividends or coupons), barriers and path dependency in general. Also, algebraic constraints in the solution due to early exercise features can be treated in a uniform manner.

The pricing equation for a two-factor contingent claim is simply the two dimensional convection-diffusion equation together with an exponentially decay term due to a discounting effect. Accurate modelling of the interaction between convective and diffusive processes is one of the most challenging tasks in the numerical approximation of partial differential equations. The choice of the numerical method depends on wether the problem is diffusion dominated or convection dominated. Very often the diffusion is quite small relative to the convection, leading to a so-called covenction-dominated problem. In all such circumstances standard finite elements and finite differences approximations will present difficulties. Thus a very large literature has been built up over the last few decades on a variety of techniques for analyzing and overcoming those difficulties. Books like Morton

(1996) are entirely devoted to the subject. A summary of numerical methods for time-dependent convection-dominated PDEs can be found in Ewing and Wang (2001). They provide a historical review of classical numerical methods and a survey of the recent developments on the Eulerian and characteristics Lagrangian methods. Eulerian methods use the standard temporal discretization, while the main distinguishing feature of characteristic methods is the use of characteristics to carry out the discretization in time.

In the more general case partial differential inequalities must be considered. The inequality arises when the contingent claim's price must satisfy some inequality constraints in order to avoid arbitrage opportunities. Possible constraints include early exercise, conversion, and call and put provisions. These are the so-called free boundary problems because there is (a priori) an unknown boundary separating the regions where inequalities are strict from those where they are saturated.

It is almost always impossible to find an explicit solution to a free boundary problem. Therefore we need numerical techniques. The extra complication in those problems comes from the fact that we do not know where the free boundary is, it is an extra unknown that we need to find as part of the solution procedure. Rigorous methods to deal with free boundaries do a transformation of the original problem into a new one with fixed domain from which the free boundary can be found a posteriori. Two possible transformations are linear complementary problems, usually combined with finite difference methods, and variational inequalities, usually related to finite element methods. The latter has some advantages. First, variational inequalities are the mathematical tool of functional analysis that best suits the rigorous formulation of early exercise problems. Second, they provide an excellent framework to deal with issues such as existence and uniqueness of the solution. Finally, they are appropriate to analyze the error incurred in the numerical methods (numerical analysis).

To rewrite the problem in a variational form we first multiply the equation with a test function from a conveniently chosen functional space and we integrate it over the domain. Then we use a Green's formula to translate second order derivatives into first order derivatives. The resulting integrals over the boundary of the domain can be computed if test functions are chosen appropriately. In particular, if we take the test functions to be zero on the boundary where Dirichlet conditions are specified, the corresponding integrals will vanish. Each solution of the classical formulation is a solution of the variational problem. Conversely, if a solution of the variational problem is twice differentiable, then it is a solution of the original problem in the classical sense. Existence and uniqueness of a classical solution require the final and boundary conditions to be smooth enough (payoff function are not even differentiable). These constraints can be weakened when we use a variational formulation of the problem; the difficulties do not disappear but solutions are sought in more general functional spaces (weighted Sobolev spaces).

1.2.3 Existence and Uniqueness of Solution. Convergence of Numerical Schemes

Kangro and Nicolaides (2000) proved the existence and uniqueness of a classical solution for the initial-boundary value problem arising in a n-dimensional Black-Scholes equation. The idea is that, even if the final condition is not differentiable the diffusion has a regularizing effect so that a regular solution exists for times prior to maturity. Unfortunately the required assumption of geometric Brownian motion process for the state variables excludes the short interest rate as an underlying process.

Theory for the existence of a solution for evolutionary variational inequalities can be seen in several reference books, as for instance Duvaut and Lions (1972), Glowinski, Lions and Tremolières (1973), Bensoussan and Lions (1978) etc. Most existence theorems have been proved under the assumption of coerciveness of the bilinear form associated to the elliptic operator. However, it turns out that partial differential equations arising in finance are usually degenerated because some of their coefficients vanish as any of the independent variables goes to zero. Therefore, the above theorems cannot be applied.

Jaillet, Lamberton and Lapeyre (1990) have weakened some of the regularity assumptions in the general theory, to apply them to one factor pricing models for American put and call options. For more general equations it turns out that the notion of viscosity solution, introduced by Crandall and Lions, is the right class of weak solutions to be considered. The theory of viscosity solutions has been developed to solve linear and non-linear degenerated problems. This concept was first introduced to solve first order equations and subsequently extended to second order elliptic and parabolic equations (see, for instances, Crandall, Ishii, Lions (1992), Lions (1983)). This notion of solution is weak enough to ensure existence and strong enough to guarantee uniqueness. Moreover reasonable numerical schemes provide approximate solutions that converge to viscosity solutions, as described by Barles and Souganides (1991). Particular applications to financial PDEs can be found in Barles, Daher and Romano (1995).

Unfortunately, although the theory of viscosity solutions can be applied to American vanilla options, a rigorous theory of existence for the general two-factor model with inequality constraints has not been found.

1.2.4 Truncation of the Domain. Boundary Conditions

The space-type variables of the pricing equation usually lie in an unbounded subset of $\Re^n$. Clearly, to obtain numerical solutions by finite differences or finite elements this domain must be truncated at large values of the state variables and suitable boundary conditions must be applied. To that purpose it is necessary to understand the behaviour of the solutions at infinity to propose relevant artificial boundary conditions. As shown by Barles, Daher and Romano (1995), artificial boundary conditions lead to small error inside the domain of computation. Kangro and Nicolaides (2000) derive pointwise bounds for the error caused by various boundary conditions (imposed on the artificial boundary for Black-Scholes type equations in multidimensional space) such that it is possible to determine a priori a suitable location for the artificial boundary in terms of a given error tolerance.

Although in some cases it is fairly straightforward to determine the asymptotic form of the PDE using financial reasoning, this becomes a more complicated task for complex contract specifications. Windcliff, Forsyth and Vetzal (2001) use a general boundary treatment, first introduced by Marcozzi (2001), which can be used in many practical situations and does not require the exact specification of Dirichlet behaviour on the boundary of the computational domain. No boundary condition is required numerically as the state variables tend to infinity if a carefully constructed lattice is used. The same can be shown to be true for explicit finite differences.[10] But as the authors point out, specifying boundary conditions is the price one has to pay to avoid timestep size limitations due to stability considerations or to achieve higher rate of convergence (obtained for example with Crank-Nicolson time weighting).

1.2.5 Numerical Solution of PDEs: FE versus FD

The classical discretization techniques are finite differences and finite elements.

It is important to be aware that not all numerical methods perform well on all problems. In order to have convergence of numerical approximation, consistency and stability properties are necessary. In this case we can sometimes obtain that the approximation has a probabilistic interpretation, which means that it can be interpreted as a Markov chain problem (or equivalently a lattice).

Finite Difference Method (FD)

FD have been widely used in finance since they were first suggested by Brennan and Schwartz (1977).

The idea of the finite difference methods is to replace the partial derivatives occurring in partial differential equations by approximations based on Taylor series expansions. When implementing FD we define a grid on the computational domain by splitting the temporal and spatial axis into a finite number of subintervals, and we restrict ourselves to the nodes of the mesh.

---

[10] This is one of the advantages that has been put forward regarding explicit methods (Hull and White (1990), Geske and Shastri (1985)).

The space discretization requires approximation of the first derivatives (convection), second derivatives (diffusion) and the function itself (reaction); we will refer to this combination as the elliptic operator. Centered approximations for the second derivatives are of order two. Central differences for the first derivative are also second order, whereas forward and backward differences are of order one only.

In general, in order to have a stable approximation it is necessary that the matrix which approximates the second order elliptic operator is diagonally dominant. This implies in particular that the matrix is invertible and satisfies the discrete maximum principle. This can be translated into some constraints on the size of the space steps involving the drifts and the diffusion coefficients. Under these conditions the stability for the second order approximation can be guaranteed for uniform elliptic operators. When those conditions are violated or when the elliptic operator is degenerate,[11] which is quite common in financial applications, stable approximations can still be obtained by using a one-sided differencing approximation for the first derivative. If the drift is non negative we use forward differences and if the drift is negative then we use backward differences. However, one-sided approximations have a convergence rate of lower order than the symmetric finite difference scheme (order 1 versus order 2).

Once we discretize the pricing equation in space, we are left with an ordinary differential equation. Several methods are available to solve this final value problem. The most common approximation methods are the θ-schemes, which include the well-known explicit ($\theta = 0$), implicit ($\theta = 1$) and Crank-Nicolson ($\theta = 1/2$) methods. In any case the final

[11] The operator is said to be degenerate when the covariance matrix of the diffusion part of the model is not positive definite.

condition of the system is the discretized final conditions of the parabolic problem (payoff function).

Explicit finite differences ($\theta = 0$) are extremely easy to implement even within a spreadsheet. The method is called explicit because the solution can be found recursively going backwards in a simple iteration from the previous time step. However, convergence of the explicit approximation is conditioned upon a strong constraint on the time step relative to the spatial step. Therefore, the explicit finite difference tends to be rather slow. On the other hand, for the remaining θ-schemes ($\theta \neq 0$) a linear system of equations must be solved at each time step, but the constraint on the time step gets weaker as the parameter increases. In the limit, the fully implicit scheme ($\theta = 1$) always converges when the space and time steps tend to zero; no conditions linking the time and spatial steps are necessary; the scheme is unconditionally stable irrespective of the step size. The Crank-Nicolson ($\theta = 1/2$) approximation is often used in practice since it may provide a second order approximation in time whereas for all other values of θ, the order of the approximation is one only.

Books like Wilmott, Dewynne, Howison (1998) and Tavella, Randall (2000) provide a comprehensive description of this numerical method and its applications in finance. They describe both the pricing models (in terms of its PDE and final and auxiliary conditions) and the numerical solution for an extensive range of products, providing clear evidence of the uniformity of the PDE approach.

Finite Element Method (FE)

The Basic idea of FE is to divide the domain of the differential equation into small non-overlapping parts, the so-called finite elements, and to approximate the unknown solution and the test function of the variational formulation, with functions from a finite dimensional space. This space is usually made up of globally continuous functions that are polynomials in each element of a polygonal mesh of the domain. In Galerkin methods the solution and the test functions are looked for in the same finite dimension space. The solution of the PDE is built as a sum of all these local approximating functions. Usually only the spatial variables are treated in this way, while time is discretized with FD or other methods. With two spatial variables, the domain is partitioned into triangles and/or quadrangles. Three dimensional spatial domains allow partitions into tetrahedrons, hexahedrons, prisms, etc.

The distinction between FE and FD is relevant at the theoretical level, i.e. when dealing with the numerical analysis. Once the discrete scheme is written and you are left with algebraic transformations of values at the grid points, the distinction vanishes. On structured meshes finite differences and finite elements plus numerical integration (using for example vertex) can be shown to be equivalent. The contrast should be seen more as variational methods versus finite differences rather than finite elements versus finite differences.

However, FE are more flexible than FD in incorporating boundary conditions and in that they allow unstructured meshes. As shown by Zvan, Forsyth and Vetzal (1998*a*), unstructured meshing can be applied to a wide variety of financial models. The idea is that an accurate solution of the pricing PDE requires in many occasions a fine mesh spacing in certain regions of the domain, usually where the gradient is steep, whereas in regions where the gradient is flat, a coarser mesh can be used. Some studies have indicated for example, the need for small mesh spacing near barriers (Figlewski and Gao, (1997), Zvan, Forsyth and Vetzal (2000)). Pooley, Forsyth, Vetzal and Simpson (2000) show that non-rectangular barriers pose difficulties for finite differences methods using structured meshes. They prove that the finite element method with standard unstructured meshing techniques can lead to significant efficiency gains over structured meshes with a comparable number of vertices. Pironneau and Hetch (2000) present and test the modified metric Voronoi of mesh adaptation for a problem with a free boundary that arises in finance for the pricing of American options, leading to satisfactory results.[12]

FE has some other computational practicalities compared to FD (see Winkler, Apel and Wystup (2001)):

- FE is very suitable for modular programming.
- A solution for the entire domain is computed instead of isolated nodes as with the FD method.
- FE provides accurate "derivatives" (risk management parameters) as a by-product.
- FE can easily deal with irregular domains, whereas this is difficultly done in FD.

[12] They use a characteristics/FE method for the space discretisation and the Brennan and Schwartz algorithm to deal with the American early exercised.

Finite elements, which are a widely spread technique in areas such as computational mechanics, have become quite popular in financial engineering in recent years.

Topper (1998), for instance, uses a commercial package to price a wide array of two-factor exotic options (Barriers, Power, and Basket options) in this framework. The software uses a hybrid of FD for time discretization (more specifically Crank-Nicolson) and FE method for space discretization. He delineates some of the advantages of FE with respect to the more widely used FD schemes, specially when computing hedge parameters. The same author has used a mixture of FD/FE and a penalty method to price a great number of passport options.

Winkler, Apel and Wystup (2001) apply the finite element method to value European vanilla options in Heston's stochastic volatility model.[13] They actually present a quite general two-factor problem and show how it can be written in a variational formulation. They also provide a theorem to characterize the existence and uniqueness of solutions and they describe the FE discretization in detail.

But the most sophisticated PDE techniques for derivatives valuation, and the most comprehensive numerical analysis of the problems has been done by Forsyth et al. ((2002), (2001), (2001) (2000), (2001), (2002), (1998*a*), (1998*b*), (1999)) . They present a general approach for two-factor PDE pricing problems using a non-conservative Galerkin FEM for the diffusion and finite volume methods (FVM)[14] for the convection, combined with a penalty method to deal with American-style features. To avoid spurious oscillation caused

[13] They use FD for time discretisation (general θ- weighted scheme) and FE for space discretisation (Lagrange triangular FE).

[14] Finite Volume Methods can be considered like Galerkin finite element methods with a special quadrature rule (Selmin and Formaggia (1996)).

by convection dominated they formulate a local extremum-diminishing scheme using a flux-limiter and central weighting. They apply this uniform approach to the pricing of convertible bonds, Asian options and two-asset options. They go on to use finite volume methods for both the discretization of the convection and the diffusion terms. The same authors have adjusted this PDE uniform approach for pricing Asian Options, discrete look-back with stochastic volatility and callable bonds for example.

1.2.6 Convection Dominance. The Method of Characteristics

When one solves diffusion/convection PDEs, what happens is that the shape of the solution, starting out with the final condition (the terminal payoff for a contingent claim), gets diffused by the second spatial derivative component of the PDE (the volatility component), and convected (i.e. displaced) by the first spatial derivative (the drift component).

In mean reverting processes, there exist regions of the domain where the drift is so large that the convection effect dominates the diffusion. The consequence for the backward time stepping numerical scheme is that the grid point where you are computing the solution can no longer collect the information from the grid points of the previous timestep that are its direct neighbours, because the shape of the solution has very quickly "convected"or shifted in the meantime to much further (downward or upward, depending on the sign of the convection) regions of the plane. It is widely known that in such situations second order centred time-discretization schemes may lead to spurious oscillations. In the lattice framework this is equivalent to saying that the local drift is so large that branching into the usual binomial or trinomial tree will lead to negative probabilities. Hull and White (1993) have solved this with their alternative branching technique. In a PDE approach one has to resort to first order upwind time-differencing or to the most recent Eulerian (including flux limiters) and characteristics techniques, such as the ones described in Ewing and Wang (2001).

The method of characteristic combined with finite elements is a possible non-centred scheme of the convective term. The combination of both discretization processes is called the characteristics/finite element method or the Lagrange-Galerkin method. In the context of Continuum Mechanics it has been introduced in the eighties by Benqué, Esposito and Labadie (1983), Pironneau (1982), Douglas-Russel (1982). An application in finance has been developed by Vázquez (1998) to solve the one-factor model arising in the valuation of American options and Pironneau and Hetch (2000) to solve the two-factor model arising in the valuation of an American put on the maximum of two assets.

As Ewing and Wang (2001) point out, this method symmetrizes and stabilizes the transport PDE, greatly reducing temporal errors. Therefore, it allows for large timesteps without loss of accuracy. In a characteristic (or Lagrangian) method, the transport is referred to a Lagrangian coordinate system that moves with the velocity in the convective term. The characteristics are paths described by the state variables over time. The time derivative along the characteristics of the advection diffusion equation is expressed as the standard time derivative (in Eulerian system, which is fixed in space) plus the convective term. Consequently, the advection-diffusion-reaction PDE is rewritten as a parabolic diffusion-reaction PDE in a Lagrangian system. In other words, in a Lagrangian coordinate system, one would only see the effect of diffusion, reaction and the right-hand-side terms but not the effect of the convection. Hence the solutions of the advection-diffusion PDEs are much smoother along the characteristics than they are in the time direction. This explains why characteristic methods usually allow large time steps to be used in a numerical simulation while still maintaining its stability and accuracy.

Houston and Suli (2000) propose an adaptive Characteristic/Finite Element method for the unsteady convection-diffusion equation. For this kind of problems the presence of local singularities in the solution may lead to a global deterioration of the numerical approximation. Therefore, it is convenient to implement an adaptive algorithm that is capable of automatically refining the discretization within regions of the computational domain where these transitions layers are located. Moreover, they derive a posteriori error bound and are able to design and implement the algorithm to ensure global control of the error with respect to a user defined tolerance.

The main drawback of the characteristics method is that it is just of order one as opposed to the second order central scheme for the first derivative. However, high-order characteristic/finite element methods have been proposed by Boukir, Maday, Metivet and Razanfindrakotoand (1997) and Rui and Tabata (2001). Bermúdez, Nogueiras and Vázquez (2004*a*), (2004*b*) extend the method in Rui and Tabata to degenerated problems; their approach could directly be applied in our case.

1.2.7 Numerical Solution of Free Boundary Problems. Lagrange Multiplier Method

In practice, the most common method of handling the early exercise condition is simply to advance the discrete solution over a timestep ignoring the restriction and then to make a projection on the set of constraints (see for example Clewlow and Strickland (1998)). This is very easy to implement but has the disadvantage that the solution is in an inconsistent state at the beginning of each timestep, or in other words, a discrete form of the linear complementary problem or the variational inequality is not satisfied (see Wilmott, Dewynne Howison (1993)).

The numerical solution of free boundary problems is difficult, especially when it involves two factors. In the case of a single factor American put the algebraic linear complementary problems are commonly solved using a projected iteration method (PSOR) that captures the unknown exercise boundary at each time step (See Wilmot (1998), Vázquez (1998)).

Clarke and Parrot (1999) suggest a multigrid method to accelerate convergence of the basic relaxation method. They show that the algorithm, when applied to the valuation of American options with stochastic volatility, gives optimal numerical complexity and the performance is much better than for the PSOR.

On the other hand, Forsyth and Vetzal (2002) propose an implicit penalty method for valuing American option and show that when varying timestep is used, quadratic convergence is achieved. They derive sufficient conditions to guarantee monotonic convergence of the nonlinear penalty iteration and also to ensure that the solution of the penalty problem is an approximate solution to the discrete linear complementary problem. They compare the efficiency and the accuracy of the method with the commonly used technique of handling the American constraint explicitly in the lattice methodologies. Convergence rates as the timestep and the mesh size tend to zero for the standard CRR tree are compared with convergence rates for an implicit finite volume method with Crank-Nicolson timestepping and the penalty method for handling the American constraint. They find that the PDE method is asymptotically superior to the binomial lattice method, even if the solution is desired at only one point.

Bermúdez and Moreno (1981) introduce a Lagrange multiplier method to solve variational inequalities in a general abstract framework. It consists of an iterative algorithm in which the solution of the variational inequality is approximated by a sequence of solutions of variational equalities. It has not been applied in finance before but has been used extensively in other fields such as computational fluid mechanics. The algorithm enjoys great generality in the sense that it allows for any type of free-boundary that may be a function of the space variables and time.

Chapter 2
Two-factor Pricing for a Class of Contingent Claims.

We saw in Chapter 1 that the fair price of many financial derivatives can be obtained by solving a final-value problem for a degenerate parabolic partial differential equation eventually involving inequality constraints. These constraints appear, for instance, in options that can be exercised at any time before expiry (the so called American style options) or in convertible bonds. In those cases the weak formulation of the problem is a parabolic variational inequality, a well known functional tool for unilateral problems in Continuum Mechanics (Duvaut and Lions (1972)), free boundary problems (Elliot and Ockendon (1982)) and many others (Glowinski, Lions and Tremolieres (1973), Bensoussan and Lions (1978)). We recall that writing a weak formulation is of great interest for theoretical analysis of the model but it is also an unavoidable step in order to use finite element methods for numerical solution.

The contribution of this Chapter is the development of a new approach for solving two-factor option-pricing problems that are written as partial differential inequalities (PDIs). The method of characteristics and finite elements is proposed for time and space discretization respectively, together with a Lagrange multiplier method to deal with inequality constraints in the solution. The combination of these three numerical methods has not been used before in finance. Our work is published in Bermúdez and Nogueiras (2004).

There are three main issues when using numerical methods in contingent claim valuation. (1) How to account for early-exercise features; (2) how to discretize the model; and (3) how to deal with the convection dominance problem. We discussed those topics in detail in Chapter 1, and recall only the most relevant ideas below.

The numerical solution of free boundary problems is difficult and has not been done rigorously in the finance literature. In order to deal with free boundary problems we first reformulate them in a weak sense as variational inequalities. Then we propose an iterative algorithm in which the solution of the variational inequality is approximated by a sequence of solutions of variational equalities. This algorithm is a particular application of the one introduced by Bermúdez and Moreno (1981) and has been used extensively in other fields. As pointed out in Chapter 1 the algorithm is very general in the sense that allows for any type of early exercise provision that may be function of time and/or state variable. Besides, the algorithm allows keeping track of the free boundary surfaces at every time step. Therefore we can solve not just for the security value at any time during its life but also we can determine ex-ante for which levels of the state variables the free boundaries will be hit, or equivalently the embedded options will become in the money. This latter feature is very important when pricing convertible bonds.

While most of papers and books on financial derivatives employ finite differences (FD) for the numerical solution (see for instance Wilmott, Dewynne and Howison (1993)), the use of finite elements (FE) has several advantages: firstly, unstructured meshes can be convenient to make refinements at some particular parts of the domain (e.g. near free boundaries) (see Pironneau and Hetch (2000)) and, secondly, it provides great flexibility in terms of changing final or boundary conditions and incorporating inequality constraints.

When the diffusion is small relative to the convection traditional numerical schemes will present difficulties. Many different ideas and approaches have been proposed in different contexts to resolve the difficulties. Exponential fitting, compact differencing, upwinding, artificial viscosity, streamline diffusion and Petrov Galerkin methods are some examples from the main fields of FD and FE. As highlighted in Chapter 1 the method of characteristics is a possible upwinding scheme that leads to symmetrical and stable approximations of the transport PDE, reducing temporal errors and allowing for large timesteps without loss of accuracy.

Let us remark that the proposed methodology allows us to price a wide range of exotic options by doing just slight changes in the computer code. Thus, the characteristics/finite element discretization combined with the Lagrange multiplier method may lead to the development of fast and accurate package that obviates the need for a separate numerical technique for the pricing of each class of exotic option, as it is often the case.

This Chapter is organized as follows. In Section 2.1 we introduce the general modelling framework and in Section 2.2 we describe the numerical methods propose for its solution.

2.1 Modelling Framework

In this Section we introduce the general two-factor pricing framework; we set the final-value problem to be solved, write it in divergence form and make a weak formulation.

2.1.1 Partial Differential Inequalities

Let the value of a contingent claim, V, be a function of time, t, and two stochastic variables x_1, x_2, the evolution of which is given by the system of stochastic differential equations

$$dx_j = \mu_j (x_1, x_2, t)\, dt + \sigma_j (x_1, x_2, t)\, dZ_j, \;\; j = 1, 2, \tag{2.5}$$

where Z_1, Z_2 are two Wiener processes with correlation coefficient ρ.

Following the standard dynamic hedging and no-arbitrage arguments by Black and Scholes (1973) and Merton (1973), it can be shown that the value V of the contingent claim is a solution of a partial differential equation of the following form

$$\frac{\partial V}{\partial t} + \sum_{i,j=1}^{2} A_{ij} \frac{\partial^2 V}{\partial x_i \partial x_j} + \sum_{j=1}^{2} B_j \frac{\partial V}{\partial x_j} + A_0 V + f = 0 \quad \text{in}\, \Omega \times (0, T), \tag{2.6}$$

where Ω is the spatial domain, A_{ij}, B_i, A_0 and f are given measurable functions of x_1, x_2, t. Typically, x_1, x_2 represent quantities such as the value of an underlying asset or a stochastic interest rate. Therefore they run either in the interval $[0, \infty)$ or in the whole real line $\Re$.

Early-exercise features, in American options, or convertibility features, in convertible bonds, may be included in the above model by means of unilateral constraints applied to V. Hence, partial differential inequalities, rather than partial differential equations, have to be considered. Precisely, in those cases, the valuation problem consists of finding two functions V and P such that

$$\frac{\partial V}{\partial t} + \sum_{i,j=1}^{2} A_{ij} \frac{\partial^2 V}{\partial x_i \partial x_j} + \sum_{j=1}^{2} B_j \frac{\partial V}{\partial x_j} + A_0 V + f = P, \text{ in } \Omega \times (0, T) \tag{2.7}$$

and furthermore

$$R_1 \leq V \leq R_2, \tag{2.8}$$

with

$$R_1 < V < R_2 \implies P = 0, \tag{2.9}$$

$$V = R_1 \implies P \leq 0, \tag{2.10}$$

$$V = R_2 \implies P \geq 0, \tag{2.11}$$

where $R_1(x_1, x_2, t)$, $R_2(x_1, x_2, t)$ are given functions. We also have to include final conditions which depend upon the specific derivative product. The function P is a Lagrange multiplier which adds or subtracts value in order to ensure that constraints in the solution are being met. Certainly, in the region where $P = 0$, the equality in (2.6) holds. The surfaces separating the regions where $P < 0$, $P = 0$ and $P > 0$ are the so called free-boundaries.

2.1.2 Variational Formulation

In order to discretize in space using the finite element method we have to rewrite the problem in a variational (weak) form. Variational inequalities are not only the starting point for the finite element discretization, but also constitute a powerful tool to deal with theoretical issues, such as existence and uniqueness of the solution as well as numerical analysis.

We first reverse the direction of time by introducing a new variable, $T - t$, which we shall still denote by t, such that the valuation Cauchy problem is an initial value problem. Also, equation (2.7) needs to be written in divergence form

$$\frac{\partial V}{\partial t} - \sum_{i,j=1}^{2} \frac{\partial}{\partial x_i}\left(a_{ij}\frac{\partial V}{\partial x_j}\right) + \sum_{j=1}^{2} b_j \frac{\partial V}{\partial x_j} + a_0 V - f + P = 0, \tag{2.12}$$

where the new coefficients a_{ij}, b_i, a_0 are given by

$$a_{11} = A_{11},\ a_{22} = A_{22},\ a_{12} = a_{21} = \frac{1}{2}\left(A_{12} + A_{21}\right), \tag{2.13}$$

$$b_1 = \sum_{i=1}^{i=2} \frac{\partial a_{i1}}{\partial x_i} - B_1 = \frac{\partial A_{11}}{\partial x_1} + \frac{1}{2}\frac{\partial\left(A_{12} + A_{21}\right)}{\partial x_2} - B_1, \tag{2.14}$$

$$b_2 = \sum_{i=1}^{i=2} \frac{\partial a_{i2}}{\partial x_i} - B_2 = \frac{\partial A_{22}}{\partial x_2} + \frac{1}{2}\frac{\partial\left(A_{12} + A_{21}\right)}{\partial x_1} - B_2, \tag{2.15}$$

$$a_0 = -A_0. \tag{2.16}$$

Notice that we have imposed symmetry to a_{ij} matrix.

Equation (2.12) is simply a two-dimensional linear convection-diffusion-reaction equation, with diffusion tensor $a = (a_{ij})$, velocity vector $\vec{b} = (b_1, b_2)$ (convection), and reaction term, a_0.

It will be useful, for the following Sections, to formulate the model using the material or total derivative of V with respect to (inverse) time t and the velocity field $\vec{b}$, namely

$$V = \frac{\partial V}{\partial t} + \vec{b} \cdot gradV. \tag{2.17}$$

With this notation, equation (2.12) becomes

$$\dot{V} - \sum_{i,j=1}^{2} \frac{\partial}{\partial x_i}\left(a_{ij}\frac{\partial V}{\partial x_j}\right) + a_0 V - f + P = 0, \tag{2.18}$$

In principle, the problem to be solved is a pure Cauchy problem. Hence, only an initial condition needs to be prescribed. However, numerical discretization by using either finite-difference, finite-elements, or finite-volume methods makes it necessary to cut the domain at finite distance and to introduce there "artificial" boundary conditions. Those are generally obtained by financial arguments, but also by pure mathematical reasoning, and have to be included in the weak formulation. This process, called "localization", often arises in numerical finance, and introduces a model error which has been studied, for instance, by Kangro and Nicolaides (2000) or by Barles, Daher and Romano (1995). The discussion on boundary conditions, sometimes ignored in financial literature, is often a complicated task and depends on the particular financial product. We will discuss boundary conditions in detail for convertible bonds.

Let us still call Ω the bounded domain, and Γ its boundary. We denote Γ_D (respectively Γ_R) the subset of Γ where Dirichlet (respectively Robin) boundary conditions are imposed. More specifically,

$$\frac{\partial V}{\partial n_A} + \alpha V = g \text{ on } \Gamma_R, \tag{2.19}$$

$$V = h \text{ on } \Gamma_D, \tag{2.20}$$

where

$$\frac{\partial V}{\partial n_A} = \sum_{i,j=1}^{2} a_{ij} \frac{\partial V}{\partial x_j} n_i, \tag{2.21}$$

and $\vec{n} = (n_1, n_2)$ denotes a unit outward normal vector to Γ. In (2.19) and (2.20), functions α, g and h are data.

In general, $\Gamma \backslash \{\Gamma_D \cup \Gamma_R\}$ is a non-empty set where no boundary conditions are needed because the natural condition for the weak formulation is identically satisfied.

Remark 1 *The computational or "localized" domain Ω is frequently a rectangle $[x_1^{\min}, x_1^{\max}] \times [x_2^{\min}, x_2^{\max}]$. In such case, the boundary Γ may be decomposed as:*

$$\Gamma = \bigcup_{i=1}^{4} \Gamma_i, \tag{2.22}$$

with

- $\Gamma_1 = \Gamma \cap \{x_2 = x_2^{\min}\}$,
- $\Gamma_2 = \Gamma \cap \{x_1 = x_1^{\max}\}$,
- $\Gamma_3 = \Gamma \cap \{x_2 = x_2^{\max}\}$,
- $\Gamma_4 = \Gamma \cap \{x_1 = x_1^{\min}\}$.

Let $a(t; \cdot, \cdot)$ be the family of bilinear symmetric forms

$$\begin{aligned} a(t; V, W) &= \sum_{i,j=1}^{2} \int_\Omega a_{ij}(x_1, x_2, t) \frac{\partial V}{\partial x_j} \frac{\partial W}{\partial x_i} dx_1 dx_2 \\ &+ \int_\Omega a_0(x_1, x_2, t) V W dx_1 dx_2 + \int_{\Gamma_R} \alpha(x_1, x_2, t) V W d\Gamma, \quad (2.23) \end{aligned}$$

and $L(t, \cdot)$ be the family of linear forms

$$L(t; W) = \int_\Omega f(x_1, x_2, t) W dx_1 dx_2 + \int_{\Gamma_R} g(x_1, x_2, t) W d\Gamma. \quad (2.24)$$

In order to write a weak formulation of the valuation problem we multiply equation (2.18) by a test function, integrate in Ω and use a Green's formula. Then the following two equivalent weak formulations can be obtained.

- Primal Formulation, in which the Lagrange multiplier P is eliminated leading to a variational inequality of the first kind:

  Find $V(t) \in \mathcal{K}(t)$ such that

$$\int_\Omega \dot{V}(t)(W - V(t)) dx_1 dx_2 + a(t; V(t), W - V(t)) \geq L(t, W - V(t)) \;\; \forall W \in \mathcal{K}(t). \quad (2.25)$$

- Mixed Formulation, which involves the two unknowns V and P :

Find $V(t) \in \mathcal{X}$ and $P(t) \in \mathcal{M}$ satisfying conditions $(2.8) - (2.11)$ and (2.20) such that

$$\int_\Omega V(t)\,W dx_1 dx_2 + a(t; V(t), W) + \int_\Omega PW dx_1 dx_2 = L(t, W) \;\; \forall W \in \mathcal{X}_0, \quad (2.26)$$

where $\mathcal{X}$ and $\mathcal{M}$ are suitable functional spaces for V and P, respectively,

$$\mathcal{X}_0 = \left\{ W \in \mathcal{X} : \; U_{/\Gamma_D} = 0 \right\}, \quad (2.27)$$

and $\mathcal{K}(t)$ is the family of convex sets of functions defined, for each t in $[0, T]$, by

$$\begin{aligned} \mathcal{K}(t) &= \{W(x_1, x_2) \in \mathcal{X} : \; R_1(x_1, x_2, t) \le W(x_1, x_2) \le R_2(x_1, x_2, t), \text{ a.e. in } \Omega, \\ & W(x_1, x_2) = h(x_1, x_2, t) \;\; \Gamma_D\}. \end{aligned} \quad (2.28)$$

As mentioned in Chapter 1 there exist some studies on existence of solution to PDEs arising in finance. Unfortunately, rigorous theory of existence for the general two-factor model with inequality constraints above is still an open problem.

2.2 Numerical Solution of the Mixed Problem

In this Section we propose a general methodology to solve the variational inequality (2.26). We first show how the solution of a variational inequality can be approximated by a sequence of solutions of variational equalities through a duality method. Then, we describe the discretization in time with characteristics and discretization in space with finite elements. Finally the Lagrange multiplier method is applied to the problem fully discretized.

2.2.1 Dealing with the Free Boundaries: Lagrange-Multiplier Method

In this Section we propose an iterative algorithm to solve the mixed weak formulation (2.26). It has been introduced in an abstract framework by Bermúdez and Moreno (1981), who also proved convergence (see Parés, Castro and Macías (2002) for further analysis).

Recall that inequalities (2.8) – (2.11) establish a relation between P and V which can be written in a more compact way by introducing the following family (indicated by x_1, x_2, t) of multi-valued maximal monotone graphs (see, for instance, Brezis (1983)) defined by

$$G(x_1, x_2, t)(Y) = \begin{cases} \emptyset & \text{if } Y < R_1(x_1, x_2, t) \\ (-\infty, 0] & \text{if } Y = R_1(x_1, x_2, t) \\ 0 & \text{if } R_1(x_1, x_2, t) < Y < R_2(x_1, x_2, t) \\ [0, \infty) & \text{if } Y = R_2(x_1, x_2, t) \\ \emptyset & \text{if } Y > R_2(x_1, x_2, t)\,. \end{cases} \tag{2.29}$$

Then it is straightforward to show that inequalities (2.8) – (2.11) are equivalent to the relation

$$P(x_1, x_2, t) \in G(x_1, x_2, t)\left(V(x_1, x_2, t)\right). \tag{2.30}$$

This means that

$$P(t) \in \partial \mathcal{I}_{\mathcal{K}(t)}(V(t)) \text{ a.e. in } (0, T), \tag{2.31}$$

where $\partial \mathcal{I}_{\mathcal{K}(t)}$ denotes the sub-differential of the indicator function of the convex set $\mathcal{K}(t)$ (see Brezis (1983)).

Since $G(x_1, x_2, t)$ is a multi-valued function, equation (2.30) is not easy to implement. However we have the following result (see Bermúdez and Moreno (1981)):

Lemma 1 *The following two statements are equivalent:*

$$\bullet\, U \in G(x_1, x_2, t)(W), \tag{2.32}$$

$$\bullet\, U = G_\lambda(x_1, x_2, t)(W + \lambda U) \text{ for all } \lambda > 0, \tag{2.33}$$

where $G_\lambda(x_1, x_2, t)$ *is the* Yosida approximation *of* $G(x_1, x_2, t)$ *defined by*

$$G_\lambda(x_1, x_2, t)(Y) = \begin{cases} \frac{1}{\lambda}(Y - R_1(x_1, x_2, t)) & \text{if } Y \leq R_1(x_1, x_2, t) \\ 0 & \text{if } R_1(x_1, x_2, t) \leq Y \leq R_2(x_1, x_2, t) \\ \frac{1}{\lambda}(Y - R_2(x_1, x_2, t)) & \text{if } Y \geq R_2(x_1, x_2, t). \end{cases}$$

We notice that G_λ is a Lipschitz-continuous (univalued) function.

In view of this Lemma and the previous discussion, relations $(2.8)-(2.11)$, or (2.30), are equivalent to the following equality

$$P(x_1, x_2, t) = G_\lambda(x_1, x_2, t)(V(x_1, x_2, t) + \lambda P(x_1, x_2, t)), \tag{2.34}$$

where λ is a positive real number.

We are now in a position to introduce the following iterative algorithm:

1. At the beginning, function P_0 is given arbitrarily.

2. At iteration m an approximation of the Lagrange multiplier P_m is known and we proceed as follows:

   Firstly, we work out a new approximation of $V(t)$, V_{m+1}, by solving the variational equality

$$\int_\Omega V_{m+1} W dx_1 dx_2 + a(t; V_{m+1}, W) + \int_\Omega P_m W dx_1 dx_2 = L(t, W)\ \forall W \in \mathcal{X}_0, \tag{2.35}$$

   together with the initial condition

$$V_{m+1}(x_1, x_2, 0) = V^0(x_1, x_2). \tag{2.36}$$

Then, we update the Lagrange multiplier P by using equality (2.34). Precisely P_{m+1} is defined as

$$P_{m+1}(x_1, x_2, t) = G_\lambda(x_1, x_2, t)\left(V_{m+1}(x_1, x_2, t) + \lambda P_m(x_1, x_2, t)\right), \qquad (2.37)$$

where, in order to achieve convergence, λ has to be greater than some positive value which depends on the coefficients a_{ij}, a_0, b_i (see Bermúdez and Moreno (1981) for details).

2.2.2 Time and Space Discretization: a Characteristics/Finite Element Method.

We proceed next to solve the general valuation problem (2.26) numerically, with a semi-discretization in time using the method of characteristics, and a spatial discretization using finite elements. The same discretization may be applied to (2.25).

At each time step, the fully discretized problem consists of a discrete variational inequality, which will be solved by the iterative numerical algorithm introduced in Section 2.2.1 for its continuous counterpart.

Time Discretization: Method of Characteristics

The total derivative with respect to time and the velocity field $\vec{b} = (b_1, b_2)$ introduced in (2.17) can be equivalently defined as

$$\dot{V}(x,t) = \frac{\partial V}{\partial \tau}\left(\phi(x,t;\tau), \tau\right)|_{\tau=t}, \qquad (2.38)$$

where $\phi(x,t;\tau) = (\phi_1(x,t;\tau), \phi_2(x,t;\tau))$ represents the trajectory described by the material point that occupies position x at time t. It is solution of the ordinary differential equation

$$\dot{\phi}(\tau) := \frac{\partial \phi(x,t;\tau)}{\partial \tau} = \vec{b}\left(\phi(x,t;\tau),\tau\right), \tag{2.39}$$

with final condition

$$\phi(x,t;t) = x. \tag{2.40}$$

If $\vec{b}$ is a Lipschitz continuous function with respect to the spatial variable and continuous with respect to the time variable, the Cauchy problem (2.39) and (2.40) has a unique solution.

Let us consider a partition of the time interval $[0,T] = \bigcup_{n=0}^{N}[t_n, t_{n+1}]$. Then, definition (2.38) suggest the following first-order backward approximation of $\dot{V}$ at time t_{n+1}

$$\dot{V}(x, t_{n+1}) \approx \frac{V(x,t_{n+1}) - V\left(\phi\left(x, t_{n+1}; t_n\right), t_n\right)}{t_{n+1} - t_n}. \tag{2.41}$$

The approximation (2.41) leads to the following implicit semi-discrete scheme for equation (2.26),

$$\begin{aligned} &\int_\Omega \frac{V^{n+1}(x) - V\left(\phi\left(x, t_{n+1}; t_n\right), t_n\right)}{t_{n+1} - t_n} W(x)dx_1dx_2 + a\left(t_{n+1}; V^{n+1}, W\right) \\ &+ \int_\Omega P^{n+1}(x)W(x)dx_1dx_2 = L\left(t_{n+1}, W\right)\ \forall W \in \mathcal{X}_0, \end{aligned} \tag{2.42}$$

for $n = 0, \ldots, N-1$ and with V^0 given by the initial condition, where $V^n(x) = V(x,t_n)$ and $P^n(x) = P(x,t_n)$.

Space Discretization: Finite Element Method

As it is well known, finite elements methods are obtained by restricting both the solution and the test functions involved in the variational formulation to be in a finite dimensional space. This space is usually made up of globally continuous functions that are polynomials in each element of a polygonal mesh of the domain Ω. In the present work, we consider the finite element space consisting of continuous piecewise linear functions on a triangular mesh of the domain Ω. Let us denote by τ_h a family of triangulations of the domain Ω, where the parameter h tends to zero and represents the size of the mesh. Linked to the triangulation τ_h , we define a family of finite-dimensional spaces of functions, namely

$$X_h = \left\{ W_h \in C\left(\Omega\right) : W_{h/K} \in \wp_1, \forall K \in \tau_h \right\}, \tag{2.43}$$

where, as usual, $C(\Omega)$ denotes the space of continuous functions defined in Ω and $\wp_1$ represents the space of polynomials of degree less or equal than one in two variables. As in the continuous problem, we define

$$X_{0,h} = \{W_h \in X_h : W_h(Q) = 0, \ \forall Q \text{ vertex on } \Gamma_D\}. \tag{2.44}$$

Having chosen the space X_h we can define a discrete counterpart of problem (2.42) where the function V^n is approximated by $V_h^n \in X_h$ and the Lagrange multiplier is approximated by $P_h^n \in X_h$.

The discrete problem can be written as:

Find V_h^{n+1}, $P_h^{n+1} \in X_h$ for $n = 0, 1, ..., N-1$ such that

$$V_h^{n+1}(Q) = H(Q)\, \forall Q \in \Gamma_D \tag{2.45}$$

and

$$\begin{aligned} &\int_\Omega \frac{V_h^{n+1} - V_h^n\left(\phi\left(x, t_{n+1}; t_n\right)\right)}{t_{n+1} - t_n} W_h dx_1 dx_2 + a^{n+1}\left(V_h^{n+1}, W_h\right) + \int_\Omega P_h^{n+1} W_h dx_1 dx_2 \\ = \;& L^{n+1}\left(W_h\right), \text{ for every } W_h \in X_{0,h} \end{aligned} \tag{2.46}$$

where

$$\begin{aligned} P_h^{n+1}(Q) &= 0 \, if \; R_1(Q, t_{n+1}) < V_h^{n+1}(Q) < R_2(Q, t_{n+1}), \\ P_h^{n+1}(Q) &\leq 0 \, if \; R_1(Q, t_{n+1}) = V_h^{n+1}(Q), \\ P_h^{n+1}(Q) &\geq 0 \, if \; V_h^{n+1}(Q) = R_2(Q, t_{n+1}), \end{aligned} \tag{2.47}$$

for all Q vertex of the triangulation τ_h.

Thus, at each time step, a discrete variational equality is obtained. As discrete initial condition we choose the interpolated function of V^0 in the space X_h, i.e., V_h^0 is continuous piecewise linear and takes the same values as V^0 at vertices of τ_h. As in the continuous problem, the discrete variational equality (2.46) can be written in a primal form as a discrete variational inequality. Precisely, problem $(2.45)-(2.47)$ is equivalent to:

Find $V_h^{n+1} \in K_h(t_{n+1})$, $n = 0, 1, ..., N-1$ such that

$$\begin{aligned} &\int_\Omega \frac{V_h^{n+1} - V_h(\phi(x, t_{n+1}; t_n))}{t_{n+1} - t_n} \left(W_h - V_h^{n+1}\right) dx_1 dx_2 + a^{n+1}\left(V_h^{n+1}, W_h - V_h^{n+1}\right) \\ \geq \; & L^{n+1}\left(W_h - V_h^{n+1}\right), \text{ for every } W_h \in K_h(t_{n+1}), \end{aligned} \tag{2.48}$$

where the convex sets $K_h(t_n)$ for $n = 0, 1, ..., N$ are defined by

$$\begin{aligned} K_h(t_n) = \; & \{W_h \in X_h : R_1(Q, t_n) \leq W_h(Q) \leq R_2(Q, t_n) \text{ and } W_h(Q) = H(Q), \\ & \forall Q \text{ vertex of } \tau_h\}. \end{aligned} \tag{2.49}$$

The Iterative Algorithm

The algorithm we have introduced in Section 2.2.1 to solve the continuous variational inequalities can now be applied to the fully discretized problem $(2.45)-(2.47)$.

At time step $(n+1)$ we start with $P_{h,0}^{n+1} = P_h^n$ and calculate sequences $P_{h,m}^{n+1}$ and $V_{h,m}^{n+1}$ indexed by m and defined as follows:

1. At iteration m, we know $P_{h,m}^{n+1}$.

2. Then, we first compute $V_{h,m+1}^{n+1}$ as the solution of the linear PDE (written in weak form)

$$\int_\Omega \frac{V_{h,m+1}^{n+1} - V_h^n\left(\phi\left(x, t_{n+1}; t_n\right)\right)}{t_{n+1} - t_n} W_h dx_1 dx_2 + a^{n+1}\left(V_{h,m+1}^{n+1}, W_h\right)$$
$$+ \int_\Omega P_{h,m}^{n+1} W_h dx_1 dx_2 = L^{n+1}\left(W_h\right), \text{ for every } W_h \in X_{0,h} \tag{2.50}$$

3. Upgrade the Lagrange multiplier using the formula (2.37)

$$P_{h,m+1}^{n+1}(Q) = G_\lambda\left(Q, t_{n+1}\right)\left(V_{h,m+1}^{n+1}(Q) + \lambda P_{h,m}^{n+1}(Q)\right), \tag{2.51}$$

for all Q vertex of τ_h, where

$$G_\lambda\left(Q, t_{n+1}\right)(Y) = \begin{cases} \frac{1}{\lambda}\left(Y - R_1\left(Q, t_{n+1}\right)\right) & \text{if } Y \leq R_1\left(Q, t_{n+1}\right) \\ 0 & \text{if } R_1\left(Q, t_{n+1}\right) \leq Y \leq R_2\left(Q, t_{n+1}\right) \\ \frac{1}{\lambda}\left(Y - R_2\left(Q, t_{n+1}\right)\right) & \text{if } Y \geq R_2\left(Q, t_{n+1}\right) \end{cases} \tag{2.52}$$

By applying the results of convergence in Bermúdez-Moreno (1981) we know that, for λ large enough, the sequence $\left\{V_{h,m}^{n+1}\right\}$ converges to the solution V_h^{n+1} as m goes to infinity.

Let us define the bilinear form

$$\tilde{a}^{n+1}\left(V_h^{n+1}, W_h\right) = a^{n+1}\left(V_h^{n+1}, W_h\right) + \frac{1}{\Delta t_n}\int_\Omega V_h^{n+1} W_h dx_1 dx_2, \tag{2.53}$$

and the linear form

$$\begin{aligned} \tilde{L}_m^{n+1}\left(W_h\right) &= L^{n+1}\left(W_h\right) + \frac{1}{\Delta t_n}\int_\Omega V_h^n\left(\phi\left(x, t_{n+1}; t_n\right)\right) W_h dx_1 dx_2 \\ &\quad - \int_\Omega P_{h,m}^{n+1} W_h dx_1 dx_2, \end{aligned} \tag{2.54}$$

for $\Delta t_n = t_{n+1} - t_n$. Then step 2 above can be rewritten as:

Find $V_{h,m+1}^{n+1}$ such that

$$\overset{\sim n+1}{a}\left(V_{h,m+1}^{n+1}, W_h\right) = \overset{\sim n+1}{L_m}\left(W_h\right) \ \forall\, W_h \in X_{0,h}. \tag{2.55}$$

Let $B = \left\{\phi_1, \phi_2, ..., \phi_{N_h}\right\}$ be a basis of X_h. Then the solution of (2.55) can be written in the form (we omit indices for the sake of simplicity)

$$V_{h,m+1}^{n+1} = \sum_{j=1}^{N_h} \xi_j \phi_j, \tag{2.56}$$

so that the discrete problem is equivalent to finding N_h numbers $\left(\xi_1, \xi_2, ..., \xi_{N_h}\right)$ satisfying

$$\sum_{j=1}^{N_h} \overset{\sim n+1}{a}\left(\phi_j, \phi_i\right) \xi_j = \overset{\sim n+1}{L_m}\left(\phi_i\right),\ i = 1, 2, ..., N_h. \tag{2.57}$$

Equivalently, find $\left(\xi_1, \xi_2, ..., \xi_{N_h}\right) \in \Re^{N_h}$ such that

$$A\xi = b \tag{2.58}$$

where

$$\begin{aligned} A_{kl} &= \overset{\sim n+1}{a}\left(\phi_l, \phi_k\right) = \frac{1}{\Delta t_n}\int_\Omega \phi_l \phi_k dx + \sum_{i,j=1}^{2}\int_\Omega a_{ij}\frac{\partial \phi_l}{\partial x_j}\frac{\partial \phi_k}{\partial x_i}dx \\ &\quad + \int_\Omega a_0 \phi_l \phi_k dx + \int_{\Gamma_R} \alpha \phi_l \phi_k d\Gamma, \end{aligned} \tag{2.59}$$

$$\begin{aligned} b_k &= \overset{\sim n+1}{L_m}\left(\phi_k\right) = \int_\Omega F\phi_k dx + \int_{\Gamma_R} G\phi_k d\Gamma \\ &\quad + \frac{1}{\Delta t_n}\int_\Omega V_h^n\left(\phi\left(x, t_{n+1}; t_n\right)\right)\phi_k dx - \int_\Omega P_{h,m}^{n+1}\phi_k dx. \end{aligned} \tag{2.60}$$

If a is symmetric, clearly the matrix A is symmetric as well. If besides A is positive definite, Choleski's method can be used to solve the system (2.58). In the special case where coefficients a_{ij}, a_0, α do not depend on time, the linear system has a matrix independent of both time step and iteration; therefore it needs to be computed only once. Also, in expression (2.60) the first two terms in the right hand side are independent of time (if F and G are) and iteration, whereas the third term must be actualized at every iteration and for every time step. Consequently, in order to solve these systems it is convenient to use direct Gauss-like methods, because, since the factorization step needs to be done only once, at each iteration just two triangular systems have to be solved.

The problem arising now is how to choose the basis $B = \{\phi_1, \phi_2, ..., \phi_{N_h}\}$. This space is usually made up of globally continuous functions which are polynomials in each element of a polygonal mesh of the domain Ω. The elements of the basis are functions that become zero in big regions of Ω so that many terms of the matrix A are zero, i.e., A is a sparse matrix. We use Lagrange triangular finite elements of degree one in a two dimensional space; in this case the basis function ϕ_i takes value 1 in the vertex i of a triangular mesh of the domain and is zero in all other vertex. In Appendix A we work out the matrix of coefficients and the independent term for this particular case.

2.3 Conclusions

In this Chapter we combined the method of characteristics/finite elements with a Lagrange multiplier method to solve two-factor contingent claim pricing models with early-exercise features. This numerical methods have not been used before in finance, but have been used extensively in other areas to solve similar problems.

We first wrote the PDI in a divergence form and then wrote a weak or variational formulation of the problem. In the continuous model we showed how the solution of the variational inequality can be approximated by a sequence of solutions of variational equalities. Then we introduced a semi-discretization in time using the methods of characteristics and a discretization in space using a finite element method. Finally the iterative algorithm introduced in the continuos setting was applied to the fully discretized problem.

Chapter 3
A Two-factor Equity Based Convertible Bond Model Ignoring Credit Risk

In this Chapter, we extend the previous literature by using the numerical methods introduced in Chapter 2 to solve a two-factor convertible bond pricing model. The two factors are the stock price and the short term interest rate. We present a rigorous formulation of the problem and a full specification of how the numerical methods are implemented, including the boundary conditions.

The model fits the observed term structure of interest rates and allows for correlation between the state variables. We use a variant of Hull and White's (1990) (HW) framework for the dynamics of the stochastic interest rate process. This framework (i) incorporates deterministically mean-reverting features, (ii) allows for perfect matching of an arbitrary input yield curve via the introduction of time dependent parameters, and (iii) permits an exact matching of an arbitrary term structure of volatilities (at least as seen from the present time). To that end, model calibration to simple volatility-dependent instruments such as caps and floors is carried out in a very efficient way. Given that the other state variable in our convertibles model is the stock price as opposed to the overall firm's value, we can use implied volatilities from stock options to produce a convertible bond pricing framework which is compatible with the current market data for both equity and interest rates.[15]

[15] This Chapter is based on Bermúdez and Nogueiras (2004) and Barone-Adesi, Bermúdez and Hatgioannides (2003).

As noticed in Chapter 1, previous numerical work has focused on finite difference schemes and lattice methodologies. However, given the specifications of the financial valuation problem at hand, our approach has clear advantages:

Firstly, in the valuation of convertible bonds, a partial differential inequality has to be solved. Conversion, call and put provisions impose inequality constraints in the numerical solution that have to be taken into account in order to avoid arbitrage opportunities. This leads to a so-called free boundary problem. In previous numerical work where FD has been used, the treatment of the early exercise is most of the times explicit, and is therefore subject to inaccuracy problems. In lattice methodologies the treatment of American features is always explicit.

Second, contract specifications for convertibles are very complex and vary a lot across issues; consequently, the FE method provides greater flexibility and has some clear advantages in terms of computational practicalities over lattices and FD.

Finally, the convertibles' valuation PDE becomes convection dominated (in the sense that convection is big relatively to diffusion) in many regions of the domain. Convection dominance is further reinforced by the theoretically necessary choice of a mean reverting process for the interest rate. It is well known that in such situations traditional discretization schemes may lead to spurious oscillations. Previous work did not make explicit account for the convection dominance.

To validate the numerical methods we price some simple products for which there are analytical solution. We also provide theoretical convertible bond prices for more complex contract specifications an we do an empirical investigation into the pricing of an actual market issue.

The remainder of this Chapter is organized as follows. In Section 3.1 we present the two-factor valuation model for convertible bonds. In Section 3.2 we solve numerically the theoretical model. In Section 3.3 we show the numerical results. Section 3.4 concludes.

3.1 The Model

In this Section we present the valuation model: the governing equation, the inequality constraints, the auxiliary conditions and the interest rate model.

3.1.1 The Governing Equation

Let $V(r,S,t;T)$ be the price of a convertible bond with maturity date $T > t$, which is a measurable function of the underlying stock price S, the spot interest rate r and time t. The dynamics for equity and term structure are given by the following diffusion processes

$$\begin{aligned} dS &= [\mu S - D(S,t)]\,dt + \sigma S dZ_S, && (3.61)\\ dr &= u(r,t)\,dt + w(r,t)\,dZ_r, && (3.62)\\ E(dZ_r dZ_S) &= \rho(r,S,t)\,dt, \text{ with } -1 \leq \rho(r,S,t) \leq +1, && (3.63) \end{aligned}$$

where μ and σ are the expected rate of return and volatility of the underlying stock, $D(S,t)$ is the dividend yield , and u and w are the drift and volatility of the spot interest rate which may be time-dependent. This latter feature of the interest rate process ensures that the bond valuation can be made consistent with the market time value of money. The two Wiener processes dZ_S and dZ_r have correlation coefficient ρ.

Following the no-arbitrage arguments by Brennan and Schwartz (1977), for instance, it can be shown (see Kwok (1998) or Wilmott (1998)) that the fair value of the convertible bond satisfies the following PDE (in order to keep the notation light ,we suppress functional dependencies)

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \rho S \sigma w \frac{\partial^2 V}{\partial S \partial r} + \frac{1}{2} w^2 \frac{\partial^2 V}{\partial r^2} + (rS - D(S,t)) \frac{\partial V}{\partial S}$$
$$+ (u - \lambda_r w) \frac{\partial V}{\partial r} - rV = 0, \tag{3.64}$$

where $\lambda_r(r,t)$ is the market price of interest rate risk (see Vasicek (1977)) and appears in the valuation equation because the state variable r is not a traded asset itself.

3.1.2 Convertible Bond Valuation as a Free Boundary Problem

A rational investor seeks to maximize the value of the convertible bond at any point in time. Following McConnell and Schwartz (1986), the value of a convertible bond must be greater than or equal to its conversion value

$$V(r,S,t) \geq nS, \tag{3.65}$$

where n is the number of shares of the issuer's common stock into which the convertible can be converted (also known as the conversion ratio).

The optimal conversion condition implies that at each point in time t and each level of the interest rate r there is a particular value of $S = S_f(r,t)$ which marks the boundary between the holding region and the conversion region. We assume that this value is unique and we refer to it as optimal exercise price. This is what is known in the literature as a free boundary problem, similar to the valuation of American-style vanilla options, which gives rise to the following partial differential inequality

$$\frac{\partial V}{\partial t}+\frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2}+\rho S\sigma w\frac{\partial^2 V}{\partial S\partial r}+\frac{1}{2}w^2\frac{\partial^2 V}{\partial r^2}+\left(rS-D\left(S,t\right)\right)\frac{\partial V}{\partial S}$$
$$+\left(u-\lambda_r w\right)\frac{\partial V}{\partial r}-rV\leq 0. \tag{3.66}$$

When it is optimal to hold the convertible bond, the equality in (3.64) is valid and the strict inequality in (3.65) must be satisfied. Otherwise, it is optimal to convert the bond and only the inequality in (3.66) holds and the equality in (3.65) is satisfied.

In the special case where there are no coupons paid on the bond and no dividend paid on the underlying stock, the conversion is not optimal till expiry and the convertible bond can be value explicitly as a combination of cash and a European call option. An increase in the dividend yield makes early exercise more likely, whereas an increase in the coupon payment makes conversion less probable. If the underlying stock pays dividends, before expiry there may be a large range of asset values for which the solution of the governing valuation equation (3.64) is less than the conversion value nS.

The free-boundary problem also arises from extra provisions in the convertible bond's indenture agreement. A call feature, which gives the issuing company the right to buy back the convertible issue at any time (or during specified periods, known as intermittent calls) for a specified cash amount (which can be time-varying as well) , say M_C, places an upper bound to the convertible's no-arbitrage price

$$V\left(r,S,t\right)\leq M_C. \tag{3.67}$$

In practice however, the call policy followed by managers to induce conversion is not consistent with the theoretical work of Ingersoll (1977a) and Brennan and Schwartz (1977), (1980). Ingersoll (1977b) and Constantinides and Grundy (1987) provide evidence that firms delay calling convertible bonds till long after the market price has exceeded the call price. Jalan and Barone-Adesi (1995) demonstrate that the unequal tax treatment of debt and equity and the need to tap the financial markets, justifies firms to delay calling in order to induce conversion. This allows for a formal linkage between the ex-ante need to issue callable convertible bonds, as a way to increase the residual equity value of the firm, and the observed reluctance to call ex-post. Hence, we will modify the above call condition by writing

$$V(r, S, t) \leq kM_C, \tag{3.68}$$

where k is a conveniently chosen factor bigger then one.[16]

Similarly, a put feature which gives the right to the holder of the convertible to sell it back to the issuer for a cash amount, say M_P (which can be time dependent), at any time (or again, during intermittent periods) places a lower bound to the convertibles' no-arbitrage price

$$V(r, S, t) \geq M_P. \tag{3.69}$$

Clearly, convertible bonds with call features worth less than convertibles without. On the contrary, put features increase the value of the convertible to the holder.

Unilateral conditions such as (3.65), (3.67), (3.69) suggest that at each time there are in general two stock prices where downside and upside constraints start becoming binding.

[16] In view of usual corporate policy to call back convertibles when its price exceeds by 30% the set call price, in the empirical implementation of our model we will choose $k = 1.3$.

These limiting stock prices are unknown and are part of the problem's desired solution; in other words, they are free boundaries beyond which the governing equation (3.64) does not apply. When the value of the bond is strictly between the upper and lower bounds, the equality in (3.64) holds. If the upper bound is reached, the equal sign is replaced by "greater than" and if the lower bound is reached the "less than" sign becomes into place. Precisely, the valuation problem to be solved consists of finding two functions V and P such that

$$\frac{\partial V}{\partial t}+\frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2}+\rho S\sigma w\frac{\partial^2 V}{\partial S\partial r}+\frac{1}{2}w^2\frac{\partial^2 V}{\partial r^2}+\left(rS-D\left(S,t\right)\right)\frac{\partial V}{\partial S}$$
$$+\left(u-\lambda_r w\right)\frac{\partial V}{\partial r}-rV=P, \tag{3.70}$$

and

$$\max\left\{nS, M_P\right\}\leq V\leq M_C \tag{3.71}$$

together with final and boundary conditions. P is the Lagrange multiplier which adds or substracts value in order to ensure that the constraints in the solution are being met.

3.1.3 Auxiliary Conditions

If coupons are paid discretely (typically every year or half-year), no-arbitrage arguments lead to the jump condition

$$V\left(r,S,t_c^-\right)=V\left(r,S,t_c^+\right)+K\left(r,t_c\right), \tag{3.72}$$

where $K\left(r,t_c\right)$ is the amount of discrete coupon paid on date t_c. Such discrete cashflows may be incorporated in the governing valuation equation (3.70) by adding the Dirac delta function term $-K\delta\left(t-t_c\right)$.

The final condition for the convertible bond is

$$V(r,S,T) = \max(nS, F), \tag{3.73}$$

where F is the par value of the bond. If we take into account the embedded options and the possibility of coupons payments, it becomes

$$V(r,S,T) = \min\{\max\{nS, M_P, F + K(T)\}, M_C\}. \tag{3.74}$$

Although in (3.74) we have taken into account call and put provisions, convertible contracts in the market do not allow the holder to put back the bond at expiration. Furthermore, upon call at expiry the issuer pays to the holder not the agreed call price but the redemption value (Red) plus the coupon; the same holds if the holder chooses to redeem the bond at its final date to get the principal. Redemption value and face value are not necessarily equal. Therefore, in the numerical implementation we will use

$$V(r,S,T) = \max\{nS, Red + K(T)\}. \tag{3.75}$$

3.1.4 The Interest Rate Model

The most important criticism of the two-factor models by Brennan and Schwartz (1980) and Longstaff and Schwartz (1995) is that they fail to ensure that the convertible bond valuation is consistent with the time value of money observed in the market. To overcome this shortcoming, Ho and Pfeffer (1996) proposed a two-dimensional binomial lattice which takes as inputs both the observed Treasury and stock prices. They constructed their quadro-tree (the terminology is due to Cheung and Nelken (1994) who suggested a similar approach) so that when the stock movement is ignored, the two-dimensional lattice is identical to the one-factor, arbitrage-free, term structure model described by Ho and Lee (1986) (HL) and Black, Derman and Toy (1990) (BDT). Besides computational difficulties that arise from the explosive increase in the number of node points in each discrete time step, convergence problems and the choice of a meaningful time step, since convertibles have long life spans, there is a very important drawback of the quadro-tree methodology from a financial perspective: the handling of mean reversion of the interest rate process.

There are two distinct ways of imparting to the spot interest rate process the mean reverting feature which is needed in order to bring about a realistic description of the dynamics of the observed term structure. The first way is to impose a decaying behaviour to the diffusive component of the process; this is the approach taken by HL and BDT in their algorithmically constructed lattices. The second is to assign an explicitly mean-reverting component to the deterministic part of the spot interest rate process; this is the approach taken by Hull and White (1990) who extend previous work by Vasicek (1977) and Cox, Ingersoll and Ross (1985). As Rebonato (1998) points out, it is always possible to choose the parameters of the volatility-decaying process and of the deterministically mean-reverting model in such a way that, as seen from the present time, both distributions will appear identical. The same is no longer true, however, if one considers the distributions obtainable, using the same parameters, from a later time. The volatility-decaying process will produce a new distribution (as seen from the later time) with much lower variance per unit time than it was obtained initially. If the future time step is considerably apart from the present point, in order to obtain a stationary distribution, a distribution whose variance does not grow as time goes to infinity, the forward rate process for the short interest rate would have to be almost deterministic. Clearly this can have serious implications for pricing long-dated American-type options, as they appear in three guises in convertible structures (i.e., convert, call and put).

In order to overcome this important shortcoming of the HL-BDT models evident in quadro-tree approaches, we use the Hull and White (1990) framework in our empirical parametrization of the interest rate process (3.62) which (i) incorporates deterministically mean reverting features for the spot interest rate process, (ii) allows for perfect matching of an arbitrary input yield curve via an introduction of time dependent parameters, and (iii) permits for an exact conditional calibration to an arbitrary term structure of volatilities.

One might correctly argue that the Heath, Jarrow and Morton (1992) (HJM) interest rate framework is more general than the HW's. However, given the potential complexity of the calibration and, especially for American-options, evaluation procedure within the HJM framework, the latter's comparative advantage over our adopted HW's can be profitably split between one and multi-factor (for the interest rate process alone) implementations. Rebonato (1998) (Chapters 13 and 17) shows that the benefits of the HJM approach for one factor interest rate models, are indistinguishable from the HW approach. He carries on by demonstrating that this picture changes radically in moving to multi-factor interest rate approaches, where the HJM approach has a very strong appeal, especially for those users who feel that the options they have to price and risk manage require explicit accountability of the imperfect correlation among interest rates. We believe that the imperfect local correlation among interest rates is of secondary importance to the price of a convertible bond with American-style exercise features, and in any case, it would have required a three-factor model (one stochastic process for equity, two correlated interest rate processes) which would have induced further, perhaps unnecessary for the problem at hand, complexities.

By setting the risk-neutral interest rate drift $u(r,t) - \lambda_r(r,t)\,w(r,t)$ in (3.64) equal to $\beta(t) - \gamma r$ we obtain the Hull and White model

$$dr = (\beta(t) - \gamma r)\,dt + w dZ_r, \tag{3.76}$$

where w determines the overall volatility of the short rate process and γ determines the relative volatility of long and short rates.

Both γ and w can be inferred from market prices of actively traded interest rate options. Suppose we have a set of M interest rate options, the market price of which we denote by $market_i\,(i = 1, ..., M)$. Also assume that there is an interest rate option valuation model that admits closed form solution under the Hull and White specification. Let us write *model*$_i\,(\gamma, w)$ for the theoretical option values. One way to calibrate is to solve the following minimization problem

$$\underset{\gamma,w}{\mathit{minimize}}\sqrt{\sum_{i=1}^{M}\left(\frac{model_i\,(\gamma,w) - market_i}{M}\right)^2}. \tag{3.77}$$

Then, once γ and w have been estimated, we choose $\beta = \beta^*(t)$ at a reference time t^* so that theoretical model prices and market prices of an array of input discount bonds coincide.

Under the risk neutral process (3.76), the value at time t of a pure discount bond with face value equal to 1, maturing at time T, conditional on $r_t = r$ is given by

$$Z(r,t;T) = e^{A(t;T) - rB(t;T)}, \tag{3.78}$$

where

$$A(t;T) = -\int_t^T \beta^*(s) B(s;T)\,ds + \frac{w^2}{2\gamma^2}\left(T - t + \frac{2}{\gamma}e^{-\gamma(T-t)} - \frac{1}{2\gamma}e^{-2\gamma(T-t)} - \frac{3}{2\gamma}\right), \tag{3.79}$$

$$B(t;T) = \frac{1}{\gamma}\left(1 - e^{-\gamma(T-t)}\right). \tag{3.80}$$

In order to fit the yield curve at a reference time t^*, $\beta^*(t)$ has to satisfy

$$\begin{aligned} A(t^*;T) &= -\int_{t^*}^T \beta^*(s) B(s;T)\,ds + \frac{w^2}{2\gamma^2}\left(T - t^* + \frac{2}{\gamma}e^{-\gamma(T-t^*)} - \frac{1}{2\gamma}e^{-2\gamma(T-t^*)} - \frac{3}{2\gamma}\right) \\ &= \log\left(Z_M(t^*;T)\right) + r^* B(t^*;T), \end{aligned} \tag{3.81}$$

for $Z_M(t^*;T)$ the market price of discount bond expiring at T as of time t^*.

Expression (3.81) is an integral equation which can be solved by differentiating twice with respect to time T,

$$\beta^*(t) = -\frac{\partial^2}{\partial t^2}\log\left(Z_M(t^*;t)\right) - \gamma\frac{\partial}{\partial t}\log\left(Z_M(t^*;t)\right) + \frac{w^2}{2\gamma}\left(1 - e^{-2\gamma(t-t^*)}\right). \tag{3.82}$$

If the drift in (3.76) is independent of time, we obtain the Vasicek (1977) interest rate model. In that case, the discount factor is given by

$$Z(r,t;T) = e^{A(t;T) - rB(t;T)}, \tag{3.83}$$

where

$$B(t;T) = \frac{1}{\gamma}\left(1 - e^{-\gamma(T-t)}\right), \tag{3.84}$$

and

$$A(t;T) = \frac{1}{\gamma^2}\left[B(t;T) - (T-t)\right]\left(\beta\gamma - \frac{1}{2}w^2\right) - \frac{w^2 B(t;T)^2}{4\gamma}. \tag{3.85}$$

3.2 The Numerical Solution

The numerical solution of the valuation PDE for convertible bonds can be considered as a special case of the more general two-colored option pricing problem discussed in Chapter 2. Precisely, the governing valuation equation for convertibles in (3.70) is a special case of equation (2.7) for the choices:

$$x_1 = r, \tag{3.86}$$

$$x_2 = S, \tag{3.87}$$

$$A_{11} = \frac{1}{2}w^2, \; A_{12} = A_{21} = \frac{1}{2}\rho\sigma S w, \; A_{22} = \frac{1}{2}\sigma^2 S^2, \tag{3.88}$$

$$B_1 = u - \lambda_r w, \; B_2 = rS - D(S,t). \tag{3.89}$$

Moreover, unilateral conditions such as the conversion provision (3.65), the call provision (3.67) and the put provision (3.69) fit into the general form of conditions (2.8) – (2.11) for

$$R_1(r,S,t) = \max\{nS, M_P\}, \tag{3.90}$$

$$R_2(r,S,t) = M_C. \tag{3.91}$$

Indeed,

$$\max\{nS, M_P\} \leq V \leq M_C, \tag{3.92}$$

$$\max\{nS, M_P\} < V < M_C \Longrightarrow P = 0, \tag{3.93}$$

$$V = \max\{nS, M_P\} \Longrightarrow P \leq 0, \tag{3.94}$$

$$V = M_C \Longrightarrow P \geq 0. \tag{3.95}$$

3.2.1 A Note Regarding the Unilateral Conditions

Notice that the restriction

$$nS \leq V \leq M_C, \tag{3.96}$$

implies

$$S \leq \frac{M_C}{n}. \tag{3.97}$$

If

$$S = \frac{M_C}{n}, \tag{3.98}$$

then

$$nS = V = M_C, \tag{3.99}$$

and according to the discussion in Section 3.1.2, the holder of the convertible will either convert into shares or, upon call, will give the bond back to the issuer to get M_C in cash. In any case, we do not have a convertible product any more. Therefore, if the bond is callable we need to solve just for

$$S \in \left[0, \frac{M_C}{n}\right]. \tag{3.100}$$

Thus, if the bond is callable, the spatial domain is the rectangle

$$\Omega = (0, \infty) \times \left(0, \frac{M_C}{n}\right). \tag{3.101}$$

However, since M_C may depend on time (by definition or when for example we consider accrued interest), to work with $0 \leq S \leq M_C/n$ would obey us to change the domain in each time step. In order to avoid that, we extend the solution by nS for $M_C/n \leq S \leq \infty$. This may be achieved by setting

$$R_2(r, S, t) = \max\{nS, M_C\}. \tag{3.102}$$

The solution could also be extended by M_C for $M_C/n \leq S \leq \infty$. The problem is that, in such a case, the lower unilateral restriction, $V(r,S,t) \geq nS$, would not be satisfied. Notice that for $nS > M_C$ also $nS > M_P$ and therefore

$$R_1(r,S,t) = nS. \tag{3.103}$$

This implies that

$$nS \leq V(r,S,t) \leq nS \Rightarrow V(r,S,t) = nS. \tag{3.104}$$

If there is no call we set $M_C = \infty$. The domain becomes

$$\Omega = (0,\infty) \times (0,\infty), \tag{3.105}$$

and

$$R_2(r,S,t) = \max\{nS, M_C\} = \infty, \tag{3.106}$$

i.e., there is no upper restriction.

If there is no put we set $M_P = 0$. In that way

$$R_1(r,S,t) = \max\{nS, M_P\} = nS, \tag{3.107}$$

i.e., there is no lower restriction due to put features.

3.2.2 Finite Truncation of the Domain and Boundary Conditions

In order to solve the problem numerically, a weak formulation is written in a rectangular finite spatial domain, (see Remark 1). In the present case $x_1^{\min} = 0$, $x_1^{\max} = r_\infty$, $x_2^{\min} = 0$, and $x_2^{\max} = S_\infty$, where r_∞ and S_∞ are "large enough"fi xed numbers. We will use the notation given in Remark 1 for the different parts of the boundary.

As explained in Chapter 2, the boundary of the computational domain is made up of three parts, depending on the given boundary conditions. We assume coefficient α (in equation (2.19)) is zero, so that we will impose a pure Neumann condition on Γ_R. In order to determine the part of the boundary on which there is no need to impose any condition, we develop the boundary integral that appears in the bilinear form (2.23), namely

$$\begin{aligned}&\int_{\Gamma_R} \frac{\partial V}{\partial n_A} U d\Gamma \\ = &\int_{\Gamma_R} \left(\frac{1}{2}w^2\frac{\partial V}{\partial r}n_1 + \frac{1}{2}\rho\sigma Sw\frac{\partial V}{\partial S}n_1 + \frac{1}{2}\rho\sigma Sw\frac{\partial V}{\partial r}n_2 + \frac{1}{2}\sigma^2 S^2\frac{\partial V}{\partial S}n_2\right) U d\Gamma\end{aligned} \tag{3.108}$$

This integral vanishes on Γ_1 because $S = 0$ and $\overrightarrow{n} = (0, -1)$ and therefore

$$\left.\frac{\partial V}{\partial n_A}\right|_{\Gamma_1} = -\frac{1}{2}\rho\sigma Sw\frac{\partial V}{\partial r} - \frac{1}{2}\sigma^2 S^2 \left.\frac{\partial V}{\partial S}\right|_{S=0} = 0. \tag{3.109}$$

Thus

$$\Gamma_1 \subset \Gamma \backslash \{\Gamma_D \cup \Gamma_R\}. \tag{3.110}$$

It will not vanish in general on any of the other boundaries. Hence, we will need to specify boundary conditions on the remaining boundaries.

- At a high share price, it is almost certain that the bond will be converted. Hence the following boundary condition is considered for $S \to \infty$

$$V(r, S, t) \sim nS \text{ as } S \to \infty. \tag{3.111}$$

This condition can also be implemented as a Neumann boundary condition, namely

$$\frac{\partial V}{\partial S} = n \text{ as } S \to \infty. \tag{3.112}$$

- At an infinite interest rate, the straight bond component tends to zero and we are left just with the call, the put and the conversion feature. Therefore we should have

$$V(r,S,t) = \min\left\{\max\left\{nS, M_P\right\}, M_C\right\} \text{ for } r \to \infty. \tag{3.113}$$

However this definition is not consistent with the extension $V = nS$ for $S > M_C/n$ (it would be appropriate if we extend instead by $V = M_C$). Therefore we have to define

$$\begin{aligned} V(r,S,t) &= \min\left\{\max\left\{nS, M_P\right\}, \max\left\{nS, M_C\right\}\right\} \\ &= \max\left\{nS, M_P\right\}, \text{ for } r \to \infty. \end{aligned} \tag{3.114}$$

- It appears quite difficult to specify the boundary condition for very small interest rate. The boundary condition at zero interest rate is not clearly specified in the literature. Zvan, Forsyth and Vetzal (1998a), (1999) proposes a PDE on this boundary but, besides this complicates the numerical solution, no financial justification is given. Wilmott (1998) states that this condition depends on the interest rate model specification and suggests assuming a finite partial derivative,

$$\lim_{r \to 0^+} \frac{\partial V}{\partial r}(r,S,t) < \infty. \tag{3.115}$$

However this information is not enough when coming to the implementation. Moreover the compatibility between boundary conditions and inequality constraints, due to optimal call, put and conversion (as defined by Brennan and Schwartz (1977) and Ingersoll (1977a)) is not straightforward and yet unspecified in previous work.

We have decided to use the homogeneous Neumann condition,

$$\frac{\partial V}{\partial n_A} = 0 \text{ on } r = 0, \tag{3.116}$$

which is a "natural" condition for the weak formulation (2.25). Notice that the above in particular implies a finite partial derivative with respect to the interest rate.

Remark 2 *As we have shown, no condition is required on Γ_1. However, a Dirichlet condition on this boundary could be provided. At zero share price, the convertible behaves like an ordinary bond:*

$$V(r, 0, t; T) = Y(r, t; T), \tag{3.117}$$

where $Y(r, t; T)$ is the value of the corresponding bond without the convertibility feature. In the abscence of call and put features there is a close form solution for Y. In the general case, Y must be found as the solution of a PDE with the instantaneous interest rate as the only spatial variable, and subject to appropriate auxiliary conditions. As Pironneau and Hetch (2000) point out, the above Dirichlet condition for $S = 0$ is implicitly defined in the PDE (3.64). In fact, by setting $S = 0$ in equation (3.64) the one-factor valuation for the ordinary bond is obtained.

3.2.3 The System of Characteristics

The method for time discretization described in Chapter 2 requires the solution of the system of characteristics (2.39) subject to conditions (2.40). For some special models of the spot rate these equations may be solved explicitly. For example for the Vasiceck (1977) parameterization

$$u(r,t) - \lambda_r(r,t)\, w(r,t) = \beta - \gamma r \;\; \beta, \gamma \in \Re, \tag{3.118}$$

$$w(r,t) = w, \; w \in \Re, \tag{3.119}$$

and if we assume $\rho(r, S, \tau) = \rho$ and $D(S, \tau) = D_0 S$ (i.e. a constant dividend yield D_0) the system of characteristics becomes

$$\begin{cases} \dot{\phi}_1(\tau) = \frac{1}{2}\rho\sigma w - \beta + \gamma\phi_1(\tau) \\ \dot{\phi}_2(\tau) = \left(\sigma^2 - \phi_1(\tau) + D_0\right)\phi_2(\tau) \\ \phi_1(t_{n+1}) = r \\ \phi_2(t_{n+1}) = S \end{cases} \tag{3.120}$$

The solution of this can be found analytically and is given by[17]

$$\begin{cases} \phi_1(t_n) = -\delta + [r + \delta]\, e^{-\gamma \Delta t_n} = rc + \delta(c-1) \\ \phi_2(t_n) = S \exp\left[-\left(\sigma^2 + D_0 + \delta\right)\Delta t_n\right] \exp\left[\frac{1}{\gamma}(r+\delta)(1-c)\right] \end{cases} \tag{3.121}$$

where

$$\delta = \frac{1}{\gamma}\left[\frac{1}{2}\rho\sigma w - \beta\right], \tag{3.122}$$

$$c = e^{-\gamma \Delta t_n}. \tag{3.123}$$

Notice that expressions of $\phi_1(t_n)$ and $\phi_2(t_n)$ do not depend on t_n, just on the time step $t_{n+1} - t_n = \Delta t_n$. This property allows calculations to be done just once for all time steps, in the case where the time step is constant.

If instead we consider the Hull and White (1990) parameterization

$$u(r,t) - \lambda_r(r,t)\, w(r,t) = \beta(t) - \gamma r \;\; \gamma \in \Re, \tag{3.124}$$

$$w(r,t) = w \;\; w \in \Re, \tag{3.125}$$

---

[17] See Appendix B for detail calculations.

and if we assume a constant correlation, ρ, and a constant dividend yield, D_0, the system of characteristics becomes

$$\begin{cases} \dot{\phi}_1(\tau) = \frac{1}{2}\rho\sigma w - \beta(\tau) + \gamma\phi_1(\tau) \\ \dot{\phi}_2(\tau) = (\sigma^2 - \phi_1(\tau) + D_0)\,\phi_2(\tau) \\ \phi_1(t_{n+1}) = r \\ \phi_2(t_{n+1}) = S \end{cases} \tag{3.126}$$

The solution is given by[18]

$$\begin{cases} \phi_1(t_n) = rc + \delta(c-1) + e^{\gamma t_n}\int_{t_n}^{t_{n+1}} e^{-\gamma\tau}\beta(\tau)\,d\tau \\ \phi_2(t_n) = S\exp\left[-(\sigma^2 + D_0)\Delta t_n\right]\exp\left[\int_{t_n}^{t_{n+1}}\phi_1(\tau)\,d\tau\right] \end{cases} \tag{3.127}$$

where

$$\delta = \frac{1}{\gamma}\left[\frac{1}{2}\rho\sigma w - \beta\right], \tag{3.128}$$

$$c = e^{-\gamma\Delta t_n}. \tag{3.129}$$

In order to compute this numerically we approximate the integral to get

$$\begin{cases} \phi_1(t_n) = \delta(c-1) + c\left[r + e^{\gamma\Delta t_n}\beta(\tau_n) + \beta(t_{n+1})\right]\Delta t_n/2 \\ \phi_2(t_n) = S\exp\left[-(\sigma^2 + D_0)\Delta t_n\right]\exp\left[\left[\phi_1(t_n) + \phi_1(t_{n+1})\right]\Delta t_n/2\right], \end{cases} \tag{3.130}$$

Notice that expressions of $\phi_1(t_n)$ and $\phi_2(t_n)$ depend on t_n, besides of the time step $t_{n+1} - t_n = \Delta t_n$. This property obeys calculations to be done for all time steps.

Starting with the payoff function at the initial time, the solution at each time step is calculated using the solution at the previous time step. But a problem may appear if the material point given by (3.121) and (3.127) at the time t_n lays out of the computational domain. Let us study the "artificial" flows on Γ driven by the velocity fields,

$$b_1 = \frac{1}{2}\rho\sigma w - \beta(t) + \gamma r, \tag{3.131}$$

$$b_2 = \left(\sigma^2 - r + D_0\right)S, \tag{3.132}$$

---

18 See Appendix B for detail of the calculations.

where in Vasicek model, $\beta(t) = \beta$ is a constant.

- $\Gamma_1 = \Gamma \cap \{S = 0\}$. Since the second component of the velocity always vanishes, the flow is tangent on this boundary; it is a streamline.

- $\Gamma_2 = \Gamma \cap \{r = r_\infty\}$. For sufficiently large r_∞ the first component of the velocity is positive and therefore the flow crosses the boundary outwards.

- $\Gamma_3 = \Gamma \cap \{S = S_{\max}\}$. For $r > \sigma^2 + D_0$ an inflow is encountered, since the second component of the velocity becomes negative.

- $\Gamma_4 = \Gamma \cap \{r = 0\}$. On this boundary the first component of the velocity may be positive or negative depending on the parameters of the model, hence inflow or outflow may be encountered.

Therefore, by choosing r_∞ large enough, difficulties may only appear on Γ_3 if $\sigma^2 - r + D_0 < 0$ and on Γ_4 if $\frac{1}{2}\rho\sigma w - \beta(t) < 0$. Because on Γ_3 the actual solution is known and may be extended above the boundary outside the domain, we just evaluate Dirichlet condition (3.111) on the points above this boundary. On the other hand, because on Γ_4 neither Dirichlet nor Neumann conditions are known, we just approximate the solution at points that fall outside the domain by the solution on the respective nearest points on the boundary.

3.3 Numerical Results

In this Section we present some numerical results. Firstly we benchmark the numerical method using three simple "test" contracts; secondly we price CBs with more complex contract specifications and finally we do an empirical investigation into the pricing of an actual market issue.

3.3.1 Benchmarking

In order to test the numerical method we have considered three particular cases for which an analytical solution is available; a straight zero coupon bond with no embedded options, a bond which is convertible just at expiration under constant deterministic interest rate, and a bond which is convertible just at expiration with stochastic interest rate but zero correlation between the state variables. In all the tests, we assume constant parameters for the interest rate process, i.e. we use Vasicek parametrization. If we were to use Hull and White model, we would need the function $\beta(t)$ depending on first and second derivatives of the yield curve (see (3.82)); therefore, we would have an extra source of error coming from the interpolation used to approximate the market yield curve, and the approximation used to estimate its first and second derivatives. In order to isolate the error coming from the discretization scheme, we use Vasicek model.

Test 1: Straight bond

Firstly we price a straight bond by setting

$$R_1(r,S,t) = 0, \quad (3.133)$$

$$R_2(r,S,t) = \infty, \quad (3.134)$$

and

$$V(r,S,T) = F. \quad (3.135)$$

Under the Vasicek interest rate model the value of a zero coupon bond is given by

$$V(r,t;T) = FZ(r,t;T) \quad (3.136)$$

where $Z(r,t;T)$ is the Vasicek discount factor given by $(3.83)-(3.85)$.

Test 2: Bond convertible just at expiration with deterministic interest rates

Secondly we assume a constant deterministic interest rate by setting

$$\beta = \gamma = w = \rho = 0, \quad (3.137)$$

and we price a bond convertible just at expiration by doing again

$$R_1(r,S,t) = 0, \quad (3.138)$$

$$R_2(r,S,t) = \infty, \quad (3.139)$$

but

$$V(r,S,T) = \max(nS,F). \quad (3.140)$$

Notice that

$$\begin{aligned} V(r,S,T) &= \max(nS,F) \\ &= \max(nS-F,0)+F \\ &= n\max(S-\tfrac{F}{n},0)+F. \end{aligned} \tag{3.141}$$

Therefore, the value of the convertible may be written as the sum of the value of the straight bond plus n call options on the underlying stock with strike price $X = F/n$. On the other hand, under deterministic interest rates the value of the call is given by the Black-Scholes formula. Therefore the value of the convertible may be written as

$$V(r,S,t;T) = FZ(r,t;T) + nC(S,t;r,T,X), \tag{3.142}$$

where

$$Z(r,t;T) = e^{-r(T-t)}, \tag{3.143}$$

and

$$C(S,t;r,T,X) = Se^{-D_0(T-t)}N(d_1) - XZ(r,t;T)N(d_2), \tag{3.144}$$

with

$$d_1 = \frac{1}{\sigma\sqrt{T-t}}\ln\left(\frac{Se^{-D_0(T-t)}}{XZ(r,t;T)}\right) + \frac{1}{2}\sigma\sqrt{T-t}, \tag{3.145}$$

and

$$d_2 = d_1 - \sigma\sqrt{T-t}. \tag{3.146}$$

Test 3: Bond convertible just at expiration with stochastic interest rate but zero correlation between the state variables.

We price a bond convertible just at expiration by doing again

$$R_1(r,S,t) = 0, \tag{3.147}$$

$$R_2(r,S,t) = \infty, \tag{3.148}$$

and

$$V(r,S,T) = \max(nS,F). \tag{3.149}$$

It can be shown that, if we assume zero correlation between the state variables, the value of the convertible may still be written as the sum of the value of the straight bond plus n call options on the underlying stock with strike price $X = F/n$. Precisely,

$$V(r,S,t;T) = FZ(r,t;T) + n \times \widetilde{C}(S,t;r,T,X), \tag{3.150}$$

where $Z(r,t;T)$ is the Vasiceck discount factor given in $(3.83)-(3.85)$ and $\widetilde{C}$ is a modified Black-Scholes formula where the discounting is done using the Vasiceck discount factors $Z(r,t;T)$. Specifically

$$\widetilde{C}(S,t;r,T,X) = Se^{-D_0(T-t)}N\left(\widetilde{d_1}\right) - XZ(r,t;T)\,N\left(\widetilde{d_2}\right), \tag{3.151}$$

with

$$\widetilde{d_1} = \frac{1}{\sigma\sqrt{T-t}}\ln\left(\frac{Se^{-D_0(T-t)}}{XZ(r,t;T)}\right) + \frac{1}{2}\sigma\sqrt{T-t}, \tag{3.152}$$

$$\widetilde{d_2} = \widetilde{d_1} - \sigma\sqrt{T-t}. \tag{3.153}$$

We assume the volatility of the underlying stock is $\sigma = 15\%$ and its continuous dividend yield is $D_0 = 4\%$. We value a convertible bond with face value of $F = 100$

currency unit, $T = 3.5$ years to maturity. For Test 2 and Test 3 the bond can be converted into $n = 1$ unit of the asset. For Test 1 and Test 3 the interest rate parameters are $\beta = 0.007$, $\gamma = 0.1$ and $w = 0.02$. The instantaneous interest rate is $r = 0.07$ and the stock price $S = 100$.

Domain bounds are set to be $\Omega^r = [0, 1.5]$ and $\Omega^S = [0, 400]$. Ω^S corresponds to roughly a 99.9% confidence interval on S_T. We give L^2 errors over both the entire domain Ω and also over a narrower region of interest $\widehat{\Omega} = \widehat{\Omega}^r \times \widehat{\Omega}^S$, where $\widehat{\Omega}^r = [0, 0.15]$ and $\widehat{\Omega}^S = [50, 200]$. $\widehat{\Omega}^S$ is roughly a 99% confidence interval on S_T. $\widehat{\Omega}$ reflects a range of values of r and S likely to be observed in practice and so the error on $\widehat{\Omega}$ is likely to be more representative.

We present results obtained for successive grid refinements for the relative error in L^2. Mesh 1 is the coarsest with just 15 space steps in the interest rate dimension, 40 in the stock dimension, and 35 time steps up to time $T = 3.5$. Each successive mesh doubles both the number of space steps in each dimension and the number of time steps so that the finest mesh, mesh 5, has 240 interest rate steps, 640 equity steps, and 560 times steps up to three and a half years. For Test1 we double the number of time steps for all meshes, because better convergence was achieved in that way. We use as benchmarking measure the total relative error define as

$$\frac{\left[\int_0^T \|error_t\|_{\ell^2}^2 \, dt\right]^{\frac{1}{2}}}{\left[\int_0^T \|solution_t\|_{\ell^2}^2 \, dt\right]^{\frac{1}{2}}}, \tag{3.154}$$

where

$$\|f\|_{\ell^2} =: \left(\int_\Omega f^2 d\Omega\right)^{\frac{1}{2}}, \tag{3.155}$$

Test1	Test2	Test3
78.44	92.45	92.52

Table 3.2. Analytical solution for $r = 0.07$ and $S = 100$

and

$$error_t = exact\ solution_t - num\ solution_t. \tag{3.156}$$

The exact solution for the current interest rate and stock price levels it is given in Table 3.2.

The numerical results are presented in Tables 3.3-3.5. On the boundaries we use the analytical solution. In each case two of the boundaries are Dirichlet and two are Neumann. 'Error TD' is the error on the entire domain Ω; 'Error RI' is the error on the region of interest, $\widehat{\Omega}$. 'Factor' is progressive error reduction factor in moving to a finer mesh level from the preceding mesh level. Times are in seconds.[19]

The characteristics/finite element method was analyzed by Pironneau (1982) for convection-diffusion equations. Unconditional stability and convergence order of $O(h) + O\left(\frac{h^2}{\Delta t}\right) + O(\Delta t)$ have been proved under suitable conditions for the coefficients of the equation. Although our models do not satisfy the required assumptions, the same error estimate has been obtained empirically. We see that the ratio between two consecutive errors tends to 2, which is consistent with the order of convergence given above. Errors are significantly less on the region of interest compared to the total domain.

All specifications lie within the region of interest so, in line with the errors reported, CB values are given to 2 decimal places.

---

[19] The implementation was in Fortran 77 run on a 2.4 Mhz Pentium IV PC.

Test1					
Mesh	Error TD	Factor	Error RI	Factor	Time
1	$7.38E-03$	–	$1.29E-03$		1
2	$3.77E-03$	1.96	$4.95E-04$	2.61	4
3	$1.91E-03$	1.98	$2.46E-04$	2.01	33
4	$9.59E-04$	1.99	$1.27E-04$	1.93	338
5	$4.81E-04$	1.99	$6.48E-05$	1.96	4600

Table 3.3. Error and convergence for Test 1

Test2					
Mesh	Error TD	Factor	Error RI	Factor	Time
1	$1.69E-02$		$4.54E-03$		0
2	$8.86E-03$	1.91	$1.94E-03$	2.34	3
3	$4.55E-03$	1.95	$1.08E-03$	1.79	25
4	$2.32E-03$	1.97	$5.90E-04$	1.83	266
5	$1.17E-03$	1.98	$3.09E-04$	1.91	3137

Table 3.4. Error and convergence for Test 2

Test3					
Mesh	Error TD	Factor	Error RI	Factor	Time
1	$1.48E-02$		$5.62E-03$		1
2	$7.81E-03$	1.89	$2.47E-03$	2.27	4
3	$4.03E-03$	1.94	$1.51E-03$	1.64	27
4	$2.06E-03$	1.96	$1.00E-03$	1.50	274
5	$1.05E-03$	1.97	$7.35E-04$	1.36	3214

Table 3.5. Error and convergence for Test 3

MATURITY	YIELD
$1m$	0.07
$6m$	0.07447
$1yr$	0.07016
$2yr$	0.06631
$5yr$	0.06224
$7yr$	0.06121
$10yr$	0.06037
$30yr$	0.05990

Table 3.6. Term structure of interest rates

3.3.2 Contract Specifications and Theoretical Convertible Bond Prices

We consider here the pricing of a set of theoretical convertible bonds with different contract specifications using our numerical approach. The Hull and White interest rate model (3.76) is fitted to the term structure[20] given in Table 3.6 with

$$\gamma = 0.1, w = 0.02.$$

We assume a constant correlation between the spot interest rate and the underlying stock $\rho = 0.1$. The volatility of the underlying stock is 15% and its continuous dividend yield is 4%. We value a convertible bond with face value of 1 currency unit, 3.5 years to maturity, which can be continuously converted into 1 unit of the asset and pays a semi-annual coupon of 3%. The bond can be called back at any time for 1.15 and it is continuously putable for 0.95. The following results were obtained for asset value $S = 1$ and spot rate $r = 7\%$, using 100 steps with spatial domain $[0, 2] \times [0, 4]$ for non-callable bonds and $[0, 2] \times [0, 1.15]$ for callable ones, and a regular mesh that takes 40 points not equally spaced on each axis. Results are in Table 3.7.

[20] This theoretical term structure as well as other input values are taken from Epstein, Haber and Wilmott (2000).

CONTRACT CHARACTERISTICS	VALUE
Zero-Coupon CB on Non-Dividend Paying Stock	1.0243
Zero-Coupon CB on Dividend Paying Stock	1.0000
Coupon-Bearing CB on Dividend Paying Stock	1.0869
Coupon-Bearing, Callable CB on Dividend Paying Stock	1.0498
Coupon-Bearing, Callable, Putable CB on Dividend Paying Stock	1.0810
Coupon-Bearing, Putable CB on Dividend Paying Stock	1.1296

Table 3.7. Convertible bond prices for different contract specifications

The value of the convertible declines for dividend paying stocks; this occurs because a higher dividend yield implies a lower expected rate of stock price appreciation and because the value of dividends is not impounded in the bond's price since the convertibles' investor does not receive dividend payments. Adding coupons to the bond increases, as expected, the value of the contract, and makes the probability of conversion lower. The callable feature is valuable to the issuer, hence the convertible decreases in value, whereas when the redemption option is added the contract's value increases; the two effects are however non-symmetric.

3.3.3 Empirical Results

In order to validate our two-factor convertible bond price model we carry out a comparison of our numerical solution against market quotes for the National Grid $4-\frac{1}{4}\%$ convertible issue (rated Aa3 by Moody's, A+ by S&P) maturing on 17/02/2008.

More specifically, we compare daily market quotes for the convertible's clean price with the projected price given by our model on the 21st of August 2000 for 215 days, i.e., from 21st of August 2000 to 15th of June 2001. As of the starting time in our sample, that

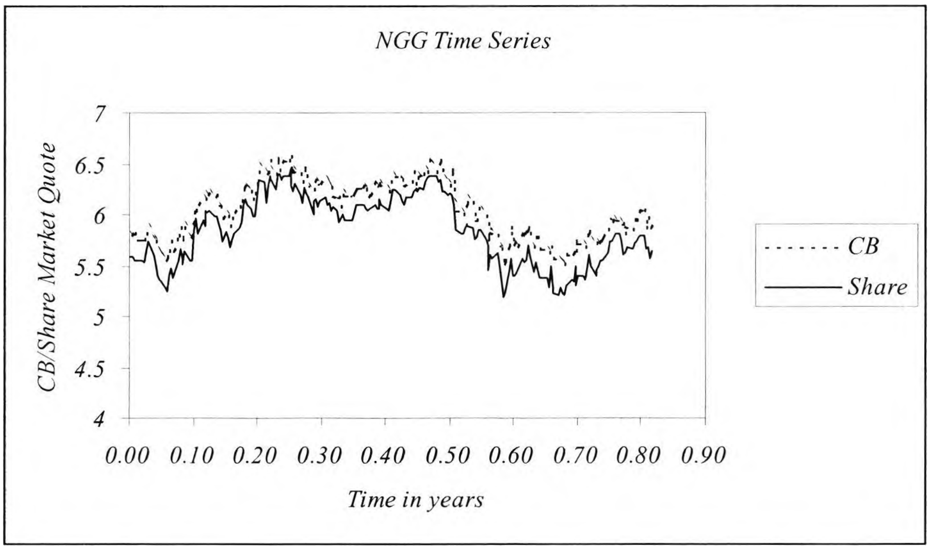

Fig. 3.1. National Grid share and clean convertible bond market prices from 21st of August 2000 (reference day) to 15th of June 2001 (end of sample). Convertible prices are expressed in a per share basis. Daily data is used throughout.

is, the 21st of August 2000, the expiration of the convertible expressed in years is

$$T = 7.49589. \tag{3.157}$$

The bond has face value

$$F = \pounds 1,000.00, \tag{3.158}$$

and can be redeemed at expiration for $\pounds 1,209.31$.

The National Grid's issue can be converted at any time at

$$n = 239.8082. \tag{3.159}$$

Figure 3.1 plots the Convertible bond- National Grid share prices from 21st August 2000 to 15th of June 2001.

The bond is continuously callable at a variety of rates shown in Table 3.8.

FROM DATE	TO DATE	CALL PRICE
$17-Feb-03$	$17-Aug-03$	108.975
$17-Aug-03$	$17-Feb-04$	110.022
$17-Feb-04$	$17-Aug-04$	111.100
$17-Aug-04$	$17-Feb-05$	112.209
$17-Feb-05$	$17-Aug-05$	113.350
$17-Aug-05$	$17-Feb-06$	114.525
$17-Feb-06$	$17-Aug-06$	115.734
$17-Aug-06$	$17-Feb-07$	116.978
$17-Feb-07$	$17-Aug-07$	118.258
$17-Aug-07$	$17-Feb-08$	119.575

Table 3.8. Call and put prices for NGG convertible bond

We have used as a proxy for the instantaneous interest rate the UK spot rate (see Duffee (1996) for an interesting discussion of alternative interest rate series). The historical correlation between the share price of the National Grid Group and the UK spot rate was calculated using daily data for the last five years[21]

$$\rho = 0.1317. \tag{3.160}$$

As input for the underlying stock's volatility, we have used at-the-money implied volatility (σ_{NG}) for the vanilla put option on National Grid Group as of the 21st August 2000

$$\sigma_{NG} = 35.05\%. \tag{3.161}$$

The share has an annual dividend yield

$$D_{NG} = 2.51\%. \tag{3.162}$$

The Hull and White interest rate model in (3.76) has been fitted and calibrated to market data as of the 21st of August 2000. Values for the overall interest rate volatility parameter w and the relative (long/short) volatility parameter γ have been chosen using actively traded

[21] Both time series were obtained from Datasteam.

T_i	σ_i	R_i
1	10.50%	0.064959
2	13.40%	0.065984
3	15.40%	0.066527
4	16.20%	0.066567
5	17.00%	0.066677
6	17.10%	0.066546
7	17.20%	0.066416
8	17.27%	0.066222
9	17.33%	0.065890
10	17.40%	0.065289

Table 3.9. Cap data

caps, with tenor of 0.25 years, and with maturities running from 1 to 10 years. Liquid cap data with expiration (T_i), rate (R_i) and at-the-money volatility (σ_i) for the 21st of August 2000 are shown in Table 3.9 . The above data set reveals one important advantage of imposing mean-revertion directly in the deterministic part of the interest rate as opposed to the volatility structure. As discussed in Section 3.1.4, algorithmically constructed lattices in the spirit of Black, Derman and Toy (1990) require a decreasing volatility structure for mean reversion of the interest rate to take place. Clearly, this pattern is not evident in the caps data above so the quadro-tree approach of Ho and Pfeffer (1996) or Cheung and Nelken (1994) for market-consistent pricing of convertible bonds fails to impose mean reversion in the interest rate process.

We have chosen as inputs of market interest rates, the zero spot curve[22] with expirations ranging from zero to ten years, equally spaced by $\tau = 0.25$ (compatible with the tenor of the caps). Figure 3.2 depicts the zero curve.

These rates have been used to approximate via cubic splines the logarithm of the market zero bond price of arbitrary maturity t as of reference time t^* (i.e., the 21st of August

[22] The zero spot interest rate curve for the 21st of August 2000 was taken from Datastream.

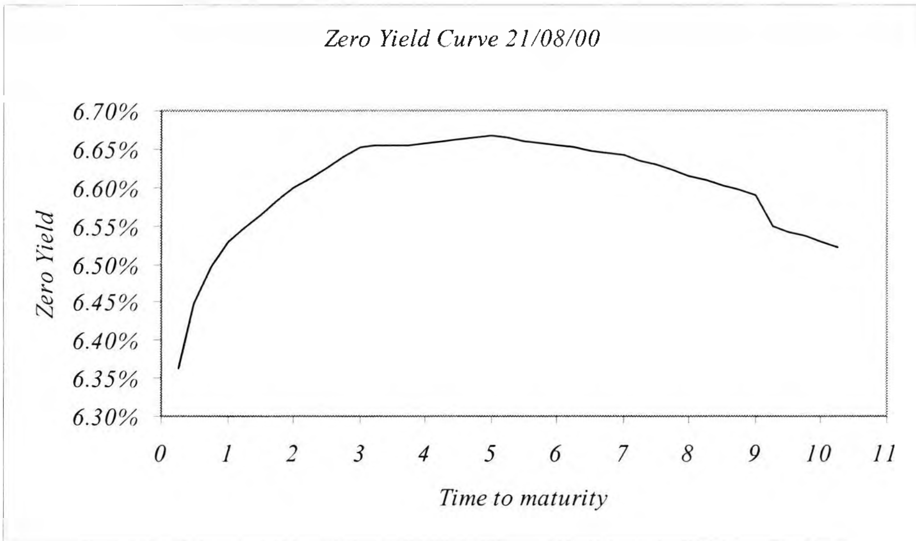

Fig. 3.2. Zero spot yield curves for 21st of August 2000 (reference day).

2000), $Z_M(t^*, t)$. Once the function $\log(Z_M(t^*, t))$ has been built, $\beta(t^*, t)$, see expression (3.82), can be evaluated and the model is guaranteed to fit observed market bond prices. Note that we do not add a constant credit spread to the riskless term structure. As discussed in Chapter 1, adding a constant option-adjusted spread or effective credit spread to the riskless interest rate penalizes the credit risk-free equity upside potential of the convertible bond. How to account optimally for the credit risk of the issuer will be discussed in later Chapters.

After calibration, the following values were obtained for the interest rate volatility parameters

$$\gamma = 0.00628, \tag{3.163}$$

$$w = 0.01025. \tag{3.164}$$

We are using 2736 daily time steps (since the convertibles' expiration is the 17th of February 2008) in our numerical solution. This provides a clear advantage of our numerical methodology compared to quadro-tree approaches which, because of the inherent explosion of the number of nodes at each time step, can only accommodate a much smaller number of time steps, thus reducing the accuracy of the calculations. We have chosen a spatial domain of $[0, 2] \times [0, 20]$. We have considered all inputs in a per share basis, i.e., we have normalized the conversion ratio to unity, and we have divided all other inputs (face value, redemption value, coupons, call price) by the given conversion ratio. On a per share basis, the historical share and bond prices, as well as all other inputs, fall in the range $[0, 10]$.

As we have seen above, the National Grid Group convertible is not callable before the 23rd of February 2003 and afterwards the call price varies with time to maturity. Therefore, we have made the up unilateral constraint time dependent.

Figure 3.3 plots our numerical valuation results as of the 21st of August 2000 against actual markets quotes for 215 successive trading days. As it can be seen in Figure 3.3, our model systematically underestimates the market. Two reasons can explain this deviation. First, it is well known that issuers of convertible bonds do not actually follow what we define as rational call policy. Instead, they wait until the share price is well above the call price in order to exercise their right. We have used a value of 30% to account for this delayed call practice by issuers, which of course, is an open matter. Second, we did not take into account in our valuations the accrued interest $(AccIR)$ which must be paid by the issuer upon call and upon put. In that case, the unilateral constraints in expression (3.92)

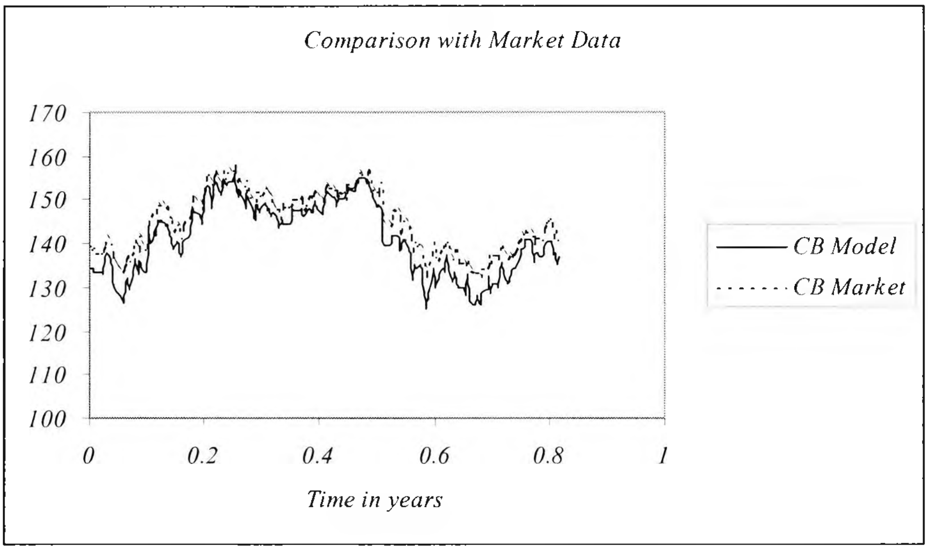

Fig. 3.3. Market quotes of the NGG convertible bond and model forecasted prices from 21st of August 2000 to 15th of June 2001. Both are clean prices and expressed in a per share basis. Daily intervals for 214 trading days.

should read as

$$\begin{aligned} \max\{nS, M_P + AccIR\} &\leq V \leq \max\{M_C + AccrIR, nS\}\ , \text{ where} \\ AccIR(t) &= K(t_{c_{i+1}})\frac{t - t_{c_i}}{t_{c_{i+1}} - t_{c_i}}, \end{aligned} \tag{3.165}$$

and $t_{c_i}, t_{c_{i+1}}$ are successive coupon payments such that $t \in [t_{c_i}, t_{c_{i+1}}]$. Omission of the accrued interest clearly underestimates the convertible bond's value.

Overall, our valuation results appear to be very promising. As it can be seen in Figure 3.4, almost all of model predictions fall within 5% of market values.

This is a considerable improvement in the accuracy of valuation results compared to the 10% average overpricing and 12.90% overpricing that King (1986) and Carayannopoulos (1996) report, respectively.

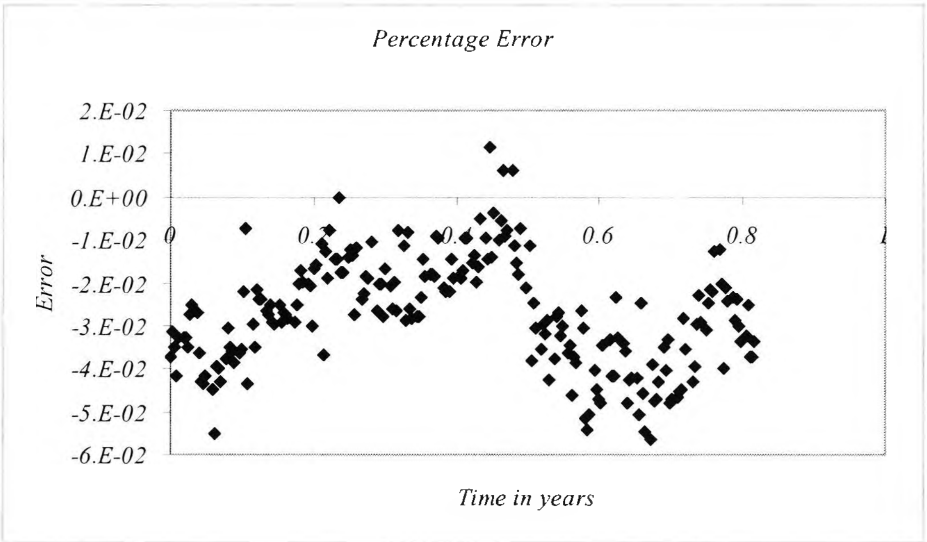

Fig. 3.4. Percentage predicted errors in the NGG convertible bond from 21st of August 2000 to 15th of June 2001. Daily intervals for 214 trading days.

3.4 Conclusion

In this Chapter we extend the previous literature on the valuation of convertible bonds by solving a two-factor model that fits the observed yield curve, imposes mean reversion in the interest rate process directly in the drift function, calibrates both interest rate and underlying equity volatilities to market data and allows for correlation between the state variables.

We have applied the method of characteristics/finite elements for time and space discretization together with a Lagrange multiplier method to deal with the early exercise features. There are clear advantages of our numerical scheme compared with the traditionally used finite differences and lattice methodologies in terms of its (i) flexibility in incorporating final conditions (the payoff function of the contingent claim), boundary conditions (at zero or infinity) and jump conditions arising from discrete intermediate payoffs of the state variables (discrete dividends and coupons), (ii) generality to pricing a wide array of exotic options and (iii) accuracy, especially for two-dimensional problems . Since our algorithm allows keeping track of the free-boundary surfaces for every discrete time step, it provides not just the solution for the price of a convertible bond at any time but also determines ex-ante for which levels of the underlying asset and the short-term interest rate the embedded conversion, call and put option will become in-the-money.

Empirical investigation into the pricing of National Grid Group's convertible issue produced prediction errors of less than 5% for 215 successive trading days, a substantial improvement compared to the 10-12% biases reported in the empirical studies of King (1986) and Carayannopoulos (1996).

Chapter 4
A Unified Framework for Pricing CBs with Interest Rate and Credit Risk

This Chapter introduces a unified intensity-based framework for pricing convertible bonds in a two and a half factor setting. The two factors are the stock price and the interest rate so that this model builds on the one presented in Chapter 3. We model the hazard rate (our half factor) as a deterministic function of the stock, the interest rate and time. We account explicitly for the stock price behaviour and holder's rights in the event of default as well as the recovery value on the bond. Most comparable existing models are special cases of our general setting. Based on this unified framework we also introduce new models. We find that different models lead to significantly different convertible bond values.

To fully specify a model it is necessary to specify how default is triggered and what happens upon default to the stock price, the CB value and holder's rights.

We recall from Chapter 1 that the first authors to have modelled default exogenously in the spirit of reduced form models, were Davis and Lischka (2002) (DL) and later Takahashi, Kobayahashi and Nakagawa (2001) (TKN). They assume that default occurs at the first jump of a Poisson process and they model the intensity of the jump as a deterministic function of the stock price. They assume that upon default the stock price jumps to zero. DL model the recovery as a constant fraction R of the par value of the bond, whereas TKN model recovery as a fraction of the market value of the bond prior to default. However, it can be argued that these approaches penalize the equity upside of the CB. The value of a convertible bond has components of different default risk; the value contributed to the bond by its conversion rights should not be subject to the same risk treatment as the fixed payments. Therefore, given the hybrid debt-equity nature of convertibles it may be convenient to split its value into a bond part and an equity part. In general, the value of the debt and equity components will be linked, and the valuation problem reduces to solving a coupled system of equations. Splitting models allow one to apply a different credit regime to the debt and equity components. Moreover, they may be of interest to investors in order to identify different sources of risk and be able to hedge them. How to split the convertible value though, is an open an controversial matter.

We saw in Chapter 1 that the first authors presenting a splitting and writing the model as a coupled system of equations were Tsiveriotis and Fernandes (1998) (TF). The value of the equity component and the value of the bond component are discounted differently to reflect their different credit risk. Ayache, Forsyth and Vetzal (2002), (2003) (AFV) extend previous literature by proposing a general specification of default in which the stock price jumps by a given percentage η upon default and the issuer has the right either to convert or to recover a given fraction R of the bond part of the convertible. The way they define the bond part is different from the original definition of Tsiveriotis and Fernandes.

We propose a unified framework to price convertible bonds incorporating interest rate and credit risk. We assume a jump-diffusion process for the stock price and a mean reverting process for the interest rate. We model the intensity as a deterministic function of the stock and the interest rate, leading to an extra so called quasi-factor or half factor. Upon default, the model allows arbitrary loss rate η on the stock price, and an arbitrary default value V^* on the convertible that may be a function of the state variables.

The model contains many other models as special cases. We identify most of the previous models and we show that the main difference between them is the specification of the recovery value.

We describe and implement an algorithm to solve the model which recovery is of (1) market value (RMV), (like TKN), (2) par (RP) (like, DL), (3) an equity-part (as for instance TF) and (4) a bond part (like AFV). We also introduce new models for the recovery value.

We propose three possible decompositions of the convertible into bond and equity parts and we investigate the need to split the convertible bond value to incorporate credit risk. We also analyze the implications of using different splitting procedures in the valuation. Based on this analysis we choose "our best model". In "our best model" the bond and equity part are defined in a different way to what has been done before in the literature; also, the recovery is specified separately in the bond and equity part as a fraction of the market value prior to default. For this reason, we refer to our model as a dual recovery model.

An implicit algorithm is proposed to solve the coupled system of equations arising from the splitting procedures. The equations are solved using the numerical methods introduced in Chapter 2.

For all models we compute the analytical solution of the special case of a bond convertible only at expiration. We show how for each model, in this special case, the CB can be written as a portfolio of straight bonds, equities and/or vanilla options. We benchmark and study the convergence of the numerical method using this special case. Based on the analytical solutions and the numerical results, we compare the implications of the different model specifications, regarding assumptions about both recovery and stock price behaviour. We also provide some sensitivity analysis with respect to the hazard rate parametrization.

In the next Section we present the general valuation framework. Section 4.2 introduces three splitting procedures to decompose the convertible into debt and equity components. Section 4.3 provides a detailed specification of the model, namely the interest rate model, the hazard rate, the recovery value and the conversion rights upon default. Also in this Section, nested models in the general framework are identified, including previous models and the dual recovery model. Section 4.4 provides the analytical solution for a special bond convertible just at expiration in all nested models. Section 4.5 describes the numerical method. Section 4.6 gives the numerical results and Section 4.7 concludes.

4.1 The Governing Equation

We follow a standard procedure given, for instance, by Protter (1995). Suppose that the value S_t of the underlying asset follows a jump augmented geometric Brownian motion under the objective measure, P*,

$$dS_t = (\mu_S - d_t)\, S_{t\_}dt + \sigma_S S_{t\_} dZ^*_{S_t} - \eta_t S_{t\_} dN_t, \tag{4.166}$$

where $Z^*_{S_t}$ is a standard Brownian motion under P^* and d_t is the continuous dividend yield. N_t is a counting process with intensity p^*_t. η_t is a deterministic loss rate. N_t models exogenous default events. At a jump time τ for N_t the equity value falls by a proportion

η_τ,

$$S_\tau = S_{\tau_-} (1 - \eta_\tau) . \qquad (4.167)$$

It is well known that the process $\nu_t^* = p_t^* t$ is the P^*-compensator of N_t, i.e., the unique finite variation previsible process such that $N_t - p_t^* t$ is a martingale under P^*.

Under the equivalent martingale measure, P, associated with the accumulator numeraire $B_t = \exp\left(\int_0^t r_s ds\right)$, the relative price $\frac{S_t}{B_t}$ is a martingale so

$$dS_t = (r_t - d_t) S_{t\_} dt + \sigma_S S_{t\_} dZ_{S_t} - \eta_t S_{t\_} (dN_t - p_t dt) , \qquad (4.168)$$

where Z_{S_t} is a standard Brownian motion under P. $\nu(dt) = p_t dt$ is the $\mathrm{P}-$compensator of the jump component. Under the change of measure, the compensator for N_t changes according to Girsanov theorem (see Jacod and Shiryaev (1988))

$$\nu_t = H(t) \nu_t^*$$

hence

$$p_t = H(t) p_t^*$$

is the jump intensity under the EMM.

Notice that, the setup defined by equation (4.168) is an incomplete market, meaning that there exists at least one contingent claim that cannot be hedged. Equivalently, under the assumption of no arbitrage, there is no unique equivalent martingale measure with which to price a contingent claim. However, given that the loss rate, η_t, is deterministic, the market can be completed by adding a defaultable bond issued by the firm which value of equity is S_t.

Lets also assume that the interest rate follows the stochastic process

$$dr_t = \mu_r dt + \sigma_r dZ_{r_t}, \tag{4.169}$$

where μ_r and σ_r are the expected rate of return and volatility of the spot interest rate, which may be function of the interest rate level as well as time. dZ_{S_t} and dZ_{r_t} are both Brownian motions that may be correlated

$$< dZ_{r_t} dZ_{S_t} >= \rho_t dt, \text{ with } -1 \leq \rho_t \leq +1. \tag{4.170}$$

We suppose that the firm has issued a convertible bond with market value V_t. The bond matures at time T with face value F. At any time up to and including time T the bond may be converted to equity. Its value upon conversion at time t is $n_t S_t$, where n_t is the conversion ratio. The bond may be called by the holder for a call price M_{C_t} and also it may be redeemed by the holder for a put price M_{P_t}. We assume that call and put prices include already accrued interest, which must be paid by the issuer upon call and upon put.

By Ito's lemma (see Jacod and Shiryaev (1988)), the process followed by V_t is

$$\begin{aligned} dV_t &= \left(\frac{\partial V_t}{\partial t} + \frac{1}{2}\sigma_S^2 S_t^2 \frac{\partial^2 V_t}{\partial S_t^2} + \rho S_t \sigma_S \sigma_r \frac{\partial^2 V_t}{\partial S_t \partial r_t} + \frac{1}{2}\sigma_r^2 \frac{\partial^2 V_t}{\partial r_t^2} + (r_t - d_t + p_t \eta_t) S_t \frac{\partial V_t}{\partial S_t} \right. \\ &\quad \left. + \mu_r \frac{\partial V_t}{\partial r_t} \right) dt + \sigma_S S_t \frac{\partial V_t}{\partial S_t} dZ_{S_t} + \sigma_r \frac{\partial V_t}{\partial r_t} dZ_{r_t} + \Delta V\left(S_{t\_}\right) \end{aligned} \tag{4.171}$$

where $\Delta V\left(S_{t\_}\right) = V_t^*\left(S_t, t\right) - V_t\left(S_{t\_}, t\right)$ and $V_t^*\left(S_t, t\right)$ is the value of the convertible bond if a jump occurs at time t. This is the value of the bond if default occurs.

Under the EMM the relative price $\frac{V_t}{B_t}$ is a martingale. Imposing this condition we have,

$$
\begin{aligned}
r_t V_t &= \frac{\partial V_t}{\partial t} + \frac{1}{2}\sigma_S^2 S_t^2 \frac{\partial^2 V_t}{\partial S_t^2} + \rho S_t \sigma_S \sigma_r \frac{\partial^2 V_t}{\partial S_t \partial r_t} + \frac{1}{2}\sigma_r^2 \frac{\partial^2 V_t}{\partial r_t^2} + (r_t - d_t + p_t \eta_t) S_t \frac{\partial V_t}{\partial S_t} \\
&\quad + \mu_r \frac{\partial V_t}{\partial r_t} + p_t \mathbb{E}_{t_-}\left[V_t^*(S_t, t) - V_t(S_{t_-}, t)\right] \qquad (4.172)
\end{aligned}
$$

Notice that $p_t \mathbb{E}_{t_-}\left[V_t^*(S_t, t) - V_t(S_{t_-}, t)\right] dt$ is the compensator of the jump $\Delta V(S_{t_-})$. When V_t^* is a deterministic function of $S_t = S_{t_-}(1 - \eta_t)$, equation (4.172) reduces to

$$
\begin{aligned}
(r_t + p_t) V_t &= \frac{\partial V_t}{\partial t} + \frac{1}{2}\sigma_S^2 S_t^2 \frac{\partial^2 V_t}{\partial S_t^2} + \rho S_t \sigma_S \sigma_r \frac{\partial^2 V_t}{\partial S_t \partial r_t} + \frac{1}{2}\sigma_r^2 \frac{\partial^2 V_t}{\partial r_t^2} \\
&\quad + (r_t - d_t + p_t \eta_t) S_t \frac{\partial V_t}{\partial S_t} + \mu_r \frac{\partial V_t}{\partial r_t} + p_t V_t^*(S_t, t). \qquad (4.173)
\end{aligned}
$$

Notice that the hazard rate p_t and the loss rate η_t have two opposite effects in the value of the CB. On one hand, p_t appears on the discounting term, therefore we expect that an increase in p_t will decrease the CB value. Similarly, if the default value, V_t^*, is a function of the stock level, $S_t = S_{t_-}(1 - \eta_t)$, an increase in the loss rate η_t will decrease the recovery value, and therefore the CB value. Hence both p_t and η_t have a negative effect on the CB value. On the other hand, under the EMM, the product of the hazard rate and the loss rate, $p_t \eta_t$, appears with positive sign on the drift of the stock price (and the CB value); this is the compensator for the jump component of S_t. The higher the loss rate and/or hazard rate, the higher the return required to compensate the risk, and therefore the higher the CB value; so p_t and η_t have also a positive effect on the CB value. If we think of the convertible as made of an equity part and a bond part, we can see that an increase of the hazard rate (and maybe the loss rate) will decrease the bond part of the CB, but will increase the equity part; an option to buy a risky asset is worth more than the option to buy a risk free one.

Inequality constraints that follow from the optimal conversion, redemption and call strategies as defined by Brennan and Schwartz (1977), make the convertible bond valuation problem a free-boundary problem which can be formulated as a variational inequality. This is modelled below via the Lagrange multiplier P which adds or substracts value to ensure that the constraints are being met.

We will use the following notation

$$\mathcal{L} = \frac{\partial}{\partial t} + \frac{1}{2}\sigma_S^2 S_t^2 \frac{\partial^2}{\partial S_t^2} + \rho S_t \sigma_S \sigma_r \frac{\partial^2}{\partial S_t \partial r_t} + \frac{1}{2}\sigma_r^2 \frac{\partial^2}{\partial r_t^2} + \left(r_t - d_t + p_t \eta_t\right) S_t \frac{\partial}{\partial S_t} + \mu_r \frac{\partial}{\partial r_t}, \tag{4.174}$$

to write equation (4.173) in short as

$$P_t = \mathcal{L} V_t - \left(r_t + p_t\right) V_t + p_t V_t^* \left(S_t, t\right). \tag{4.175}$$

together with conditions (see Chapter 3)

$$\max\left\{n_t S_t, M_{P_t}\right\} \leq V_t \leq \max\left\{M_{C_t}, n_t S_t\right\}, \tag{4.176}$$

$$\max\left\{n_t S_t, M_{P_t}\right\} < V_t < \max\left\{M_{C_t}, n_t S_t\right\} \Longrightarrow P_t = 0, \tag{4.177}$$

$$V_t = \max\left\{n_t S_t, M_{P_t}\right\} \Longrightarrow P_t \leq 0, \tag{4.178}$$

$$V_t = M_{C_t} \Longrightarrow P_t \geq 0. \tag{4.179}$$

If the bond pays coupons discretely, typically every year or half-year, let $K\left(r_t, t_c\right)$ be the amount of discrete coupon paid on date t_c. Then the following condition must be imposed in order to avoid arbitrage opportunities

$$V_t\left(r_t, S_t, t_c^-\right) = V_t\left(r_t, S_t, t_c^+\right) + K\left(r_t, t_c\right). \tag{4.180}$$

Such discrete cashflows may be incorporated in the governing valuation equation by adding a Dirac delta function term $-K\delta\left(t-t_c\right)$ to the RHS of (4.175).

The final condition for the convertible bond is the exercise condition at the maturity time T,

$$V_T\left(r_T, S_T, T\right)=\max\left(n_T S_T, F+K\left(r_T, T\right)\right)=\max\left(n_T S_T, \widetilde{F}\right), \tag{4.181}$$

where we have introduced the adjusted face value $\widetilde{F}=F+K\left(r_t, T\right)$.

Solving (4.175), (4.176) – (4.179) subject to boundary, final (4.181), and jump (4.180) conditions , gives the theoretical value of the convertible bond.

4.2 Splitting Procedures

Given the hybrid nature of convertibles, it is possible and often desirable to split the value V of the convertible into a bond part W and an equity part U. Early models valued CBs by replication as a portfolio of a bond and a warrant. Unfortunately this approach is correct when the bond is convertible only at expiration and there are no other embedded options, such as call and put features. In general, the two parts are linked and the valuation problem is a coupled system of equations. Splitting models allow a different credit treatment to be applied to the debt and equity parts. This may be of interest to investors in order to identify their risks and be able to hedge them.

We split the value of the bond into an equity part U and a bond part, W. U is the part related to payments in equity, and therefore includes the conversion and call option. W is related to payments in cash, and includes the coupons and the put option. In general, both are derivatives on the underlying stock price and the instantaneous interest rate, and will follow a partial differential equations similar to (4.175) with default values given by W^* and U^* respectively. The two parts have embedded early exercise features, and therefore follow inequalities with Lagrange multipliers P^W and P^U,

$$\mathcal{L}W_t - (r_t + p_t) W_t + p_t W_t^* (S_t, t) = P_t^W, \tag{4.182}$$

$$\mathcal{L}U_t - (r_t + p_t) U_t + p_t U_t^* (S_t, t) = P_t^U. \tag{4.183}$$

We will set $P_t^W = 0$ in the equation for W, and instead of solving a free boundary problem for W, we will impose the constraints explicitly using the total value of the convertible V (whenever it hits the free boundaries). This approach was initially followed by TF. AFV, on the contrary, propose a coupled system of free boundary problems. The latter approach makes the implementation extremely complex.

If the bond pays coupons we need to consider the jump condition for W

$$W_t \left(r_t, S_t, t_c^-\right) = W_t \left(r_t, S_t, t_c^+\right) + K \left(r_t, t_c\right), \tag{4.184}$$

where $K(r_t, t_c)$ is the amount of discrete coupon paid on date t_c.

To be fully specified we need to supply inequality constraints and final conditions to (4.182) and (4.183). At the final time the payoff to the convertible is given by (4.181)

$$V_T \left(r_T, S_T, T\right) = \max \left(n_T S_T, \widetilde{F}\right). \tag{4.185}$$

The splitting determines how V_T is allocated between W_T and U_T.

The motivation of the splitting is to apply a different credit treatment to equity and debt. Originally, in the TF model, the main objective was to use a different discount factor for the debt part and the equity part. But if we model the hazard rate as a function of the stock price this is not anymore a priority. When the stock price is low, the equity part has very little value and the convertible is almost all debt; the hazard rate will be high as well as the spread over the risk-free rate. On the contrary, for high stock prices the value of the bond part is little and the convertible is all equity; the hazard rate will be low and discounting is almost at the risk-free rate. Hence, there may be no need to split the bond value in order to apply a different discounting to the bond part and the equity part. However, if we want to use a different recovery in bond and equity, a splitting is necessary in order to define the recovery value of the convertible. It will be mandatory to solve a coupled system of equations only when the default value of the convertible V^* depends explicitly on the values, either one or both, of the equity value, U and the bond value W.

How to decompose the convertible value, or equivalently how to define the bond and equity parts, is an open and controversial matter.

We adopt TF splitting and we propose two new splitting procedures. This amounts to providing three payoff functions that will act as final conditions for (4.182) and (4.183).

4.2.1 Splitting 1. U_T: asset or nothing call, W_T: cash or nothing put

This is the original splitting introduced by TF. They define a hypothetical derivative security, the cash-only part of the convertible bond (COCB) which entitles the holder to all cash flows from the bond part but no equity cash flows. The COCB is the same as W. At time t the bond part, W_t, is set to zero if it is optimal to convert,

$$W_t(r_t, S_t, t) = 0 \; if \; V_t(r_t, S_t, t) = nS_t. \tag{4.186}$$

Similarly, and because it will be optimal for the issuer to call the bond back when

$$V_t(r_t, S_t, t) = M_{C_t} = nS_t, \tag{4.187}$$

they also set[23]

$$W_t(r_t, S_t, t) = 0 \; if \; V_t(r_t, S_t, t) = M_{C_t}. \tag{4.188}$$

On the other hand, when a cash payment takes places due to coupon payments or to redemption upon a put or at maturity, the COCB takes a non-zero value. Specifically,

$$W_T(r_T, S_T, T) = \begin{cases} 0 & \text{if } S_T \geq \widetilde{F}/n \\ \widetilde{F} & \text{otherwise} \end{cases} \tag{4.189}$$

$$W_t(r_t, S_t, t) = M_{P_t} \text{ if } V_t(r_t, S_t, t) = M_{P_t}, \tag{4.190}$$

If the bond pays coupons we add the jump condition (4.184)

$$W_t\left(r_t, S_t, t_c^-\right) = W_t\left(r_t, S_t, t_c^+\right) + K\left(r_t, t_c\right). \tag{4.191}$$

Notice that the final condition for the cash component W given in (4.189) is the payoff of a cash or nothing put with strike $\widetilde{F}/n$ and payout $\widetilde{F}$. Also notice that with this splitting the value of the equity component U becomes at maturity

$$U(r_T, S_T, T) = \begin{cases} nS_T & \text{if } S_T \geq \widetilde{F}/n \\ 0 & \text{otherwise} \end{cases} \tag{4.192}$$

which is the payoff of n asset or nothing call options with strike $\widetilde{F}/n$.

We now introduce two new splitting procedures. Splitting 2 defines the equity part in a natural way as the parity value of the convertible, and then sets the bond part as the total value of the convertible minus the equity part. Splitting 3 defines the bond part in a natural

23 Recall that upon call the holder reserves the right to convert into equity.

way as the bond floor of the convertible, and sets the equity part to be the difference of that with the full CB value.

4.2.2 Splitting 2. U_T: equity, W_T: equity premium (put)

In splitting 2 the equity part is a security that at maturity pays the value of the parity, a natural definition for the equity part of the CB. Hence, the payoff of the equity component U is

$$U\left(r,S,T\right)=nS_T, \tag{4.193}$$

which is just the value of n underlying stocks.

The value of the bond part is the value above parity, basically the equity premium. With this splitting the cashflows of the bond part are the actual cashflows of the CB minus the parity; this is consistent with the idea of the bond part being the part at risk, since upon default at a cashflow we can always convert into equity, and therefore the amount at risk is the cashflow minus the parity.

Specifically, in the event of default at expiration the holder can still convert into equity, and therefore the amount he risks is $\widetilde{F}-nS_t$ rather than $\widetilde{F}$. Because the bond part is the portion subject to credit risk, we set its payoff to be

$$\begin{aligned} W_T\left(r_T,S_T,T\right) &= \begin{cases} 0 & \text{if } S_T\geq\widetilde{F}/n \\ \widetilde{F}-nS_T & \text{otherwise} \end{cases} \\ &= \max\left\{0,\widetilde{F}-nS_T\right\}. \end{aligned} \tag{4.194}$$

Notice that this is the payoff of n put options with strike $\widetilde{F}/n$.

Similarly, if the holder decides to redeem the bond by exercising the put, the amount subject to risk is not the put price M_{P_t} but the difference $M_{P_t}-nS_t$, since the conversion into equity is guaranteed at any time. We have

$$W_t(r_t, S_t, t) = M_{P_t} - nS_t \text{ if } V_t(r_t, S_t, t) = M_{P_t}. \tag{4.195}$$

Conditions (4.186) and (4.188) apply the same in this case.

It is worth mentioning that with this definition of W, the final condition becomes a continuous function of the stock price S. This property, which is not shared by the TF splitting, makes the numerical solution far easier.

4.2.3 Splitting 3. U_T: risk premium (warrant), W_T: bond floor

We define the bond part as the floor value of the convertible. The bond floor is a straight bond with the same face value and coupon payments as the convertible, adjusted up if there is an embedded put option. The payoff to W is then

$$W_T(r_T, S_T, T) = \widetilde{F}. \tag{4.196}$$

With this splitting the final condition for the cash component W given in (4.196) is the payoff of a straight bond with face value $\widetilde{F}$. With the splitting 3 the payoff of the equity component U becomes

$$U_T(r_T, S_T, T) = \max\left\{nS_T - \widetilde{F}, 0\right\}, \tag{4.197}$$

which is the payoff of n call options with strike $\widetilde{F}/n$.

The advantage of this splitting is that is consistent with the standard decomposition of bond plus warrant in the special case in which the bond is convertible just at expiration. Also it agrees with the splitting done in convertible bond asset swaps (CBAS). CBAS are the most popular way to hedge convertibles.

4.3 Detailed Specification of the Model

In this Section we discuss in detail the remaining components of the model, namely the interest rate model, the hazard rate specification, the recovery value and the conversion rights upon default.

4.3.1 The Interest Rate Model

As in the previous Chapter we assume the interest rate model is extended Vasicek. This model combines tractability with the flexibility to calibrate to a pre-specified initial term structure. We recall that the short rate process under the EMM is

$$dr_t = \alpha\left(\theta(t) - r_t\right)dt + \sigma_r dZ_{r_t}, \tag{4.198}$$

where $\theta(t)$ can be chosen so that model spot rates coincide with market spot rates. We set $\mu_r \equiv \mu(t, r) = \alpha\left(\theta(t) - r_t\right)$ to be the drift of r.

The Vasicek model allows rates to become negative. An alternative would be to use the CIR model in which rates are certain to remain non-negative. However, choosing the Vasicek model allows us to simplify the numerical procedure.

4.3.2 The Hazard Rate Process

Instead of allowing the hazard rate to be stochastic (like Davis and Lischka (2002)), we model it as a deterministic function of the state variables and time. As mentioned in Chapter 1 several parametrizations have been used in the literature. We assume that p_t decreases as both S_t and r_t increase. The negative correlation between interest rates and hazard rates agrees with some empirical studies, although evidence is mixed (see for example Kiesel, Perraudin and Taylor (2002)). For concreteness we assume

$$p_t = k \exp\left[-\left(aS_t + br_t\right)\right] \quad a, b > 0. \tag{4.199}$$

Das and Sundaram (2004) studied a default process of the form

$$p_t = k \exp\left[-a\mathrm{Ln}\left(S_t\right) + br_t + c(T-t) + d\right]. \tag{4.200}$$

From a computational point of view, our parametrization (4.199) has the advantage over (4.200) that (4.199) can be used for $S_t = 0$, which is the lower bound of our computational domain in the S direction (see Section 3.2).

4.3.3 The Recovery Value

Suppose default occurs at time τ. We define the recovery value on the CB, V^*, as the sum of the recovery values on the bond and equity parts, W^* and U^* respectively,

$$V^* = W^* + U^*. \tag{4.201}$$

We consider the following models for the recovery value, which are special cases of the general specification above

- Recovery of bond and equity part (RBE)

We define the recovery on the equity part as a fraction of its market value prior to default. We propose to use $1-\eta$ as the recovery rate, to be consistent with the assumption about recovery on the stock price upon default. For the bond part the recovery is a fraction R of its market value prior to default, leading to

$$V_\tau^* = (1-\eta)\, U_{\tau_-} + RW_{\tau_-}. \tag{4.202}$$

- Recovery of bond part (RB)

The recovery on the equity part is zero. The recovery value on the bond part is a fraction R of its market value prior to default. Therefore the recovery on the CB is

$$V_{\tau}^{*} = RW_{\tau_{-}}. \tag{4.203}$$

If the the bond part is defined as the bond floor, this recovery model is consistent with the CBAS market.

- Recovery of equity part (RE)

The recovery on the bond part is zero. The recovery on the equity part is a fraction $1 - \eta$ of its market value prior to default. Hence,

$$V_{\tau}^{*} = (1 - \eta)\, U_{\tau_{-}}. \tag{4.204}$$

- Recovery of market value (RMV)

The recovery value of the convertible is a fraction R of its market value prior to default, namely

$$V_{\tau}^{*} = RV_{\tau_{-}} = RU_{\tau_{-}} + RW_{\tau_{-}}. \tag{4.205}$$

It can be interpreted as a RBE model in which the recovery rate on the stock equals the recovery rate on the debt, $(1 - \eta) = R$. The recovery of market value has the advantage that it is very easy to implement.

- Recovery of par (RP)

The recovery value of the convertible is a fraction R of the face value of the bond,

$$V_{\tau}^{*} = RF. \tag{4.206}$$

It can be thought as a RB model, in which the recovery of the bond part is define in terms of its principal rather than market value. The recovery of par model is very popular in the markets. It is easy to implement and it is consistent with the credit default swap market. A variation of the recovery of par is to define the recovery value as a fraction of the riskless discounted par value

$$V_{\tau}^{*} = RZ\left(\tau, r_{\tau}; T\right) F, \tag{4.207}$$

where $Z\left(\tau, r_{\tau}; T\right)$ is the riskless discount factor with maturity T as of time τ.

4.3.4 Conversion Rights upon Default

Another important issue regarding the default value, is whether or not the model should allow for conversion upon default. Realdon (2003) shows it can be rational for CB holders to convert when the debtor approaches distress. In the pricing literature, only AFV allow for conversion upon default. This is consistent with the assumption that the stock price falls on default by a given fraction η and not necessarily vanish. We adopt their assumption and redefine the bond value upon default as the maximum between the conversion price and the recovery value. In this case the pricing equations can be written as

$$P_t = \mathcal{L}V_t - \left(r_t + p_t\right) V_t + p_t \max\left\{n_t S_t \left(1 - \eta\right), V_t^{*}\right\}. \tag{4.208}$$

No other models explicitly consider holder rights on default. However, given that DL and TKN assume the stock price jumps to zero upon default, the conversion option is worthless.

This is a special case in our specification. If the stock price vanishes on default, $\eta = 1$ and

$$\max\left\{nS_{t_-}(1-\eta), V_t^*\right\} = \max\left\{0, V_t^*\right\} = V_t^*. \tag{4.209}$$

In the RBE and RMV models it is redundant to allow for conversion upon default if $R > 1-\eta$. In that case, at a default time τ, the value of the CB in the RBE model becomes

$$V_\tau^{RBE*} = (1-\eta)\,U_{\tau_-} + RW_{\tau_-} > (1-\eta)\,U_{\tau_-} + (1-\eta)\,W_{\tau_-} = (1-\eta)\,V_{\tau_-} \geq (1-\eta)\,nS_{\tau_-}. \tag{4.210}$$

where the last inequality follows from (4.176). Similarly for the RMV model

$$V_\tau^{RMV*} = RV_{\tau_-} > (1-\eta)\,V_{\tau_-} \geq (1-\eta)\,nS_{\tau_-}. \tag{4.211}$$

4.3.5 Nested Previous Models

Most of the previous models fit into the general framework presented above. The particular specification of the hazard rate, the loss rate and the recovery value will determine the difference. We have summarized why is on Table 4.10.

- Davis and Lischka

Their equation is a special case of (4.175) for deterministic interest rate, loss rate η equal to 1 and recovery of par.

- Takahasi, Kobayahashi, Nakagawa

Their equation is a special case of (4.175) for deterministic interest rate, loss rate η equal to 1 and recovery of market value.

- Tsiveriotis and Fernandes

Although TF do not discuss about default, and they model credit risk via a credit spread, a posteriori we could identify their model in the more general setting of the previous Section. The equation they propose for the total value of the convertible is the one factor counterpart of (4.175) for zero loss rate, η, constant hazard rate, p, equal to the credit spread, r_c, and value upon default, V^*, equal to the equity part of the bond, U,

$$\eta_t = 0, \tag{4.212}$$

$$p_t = r_c, \tag{4.213}$$

$$V_t^* = U_t. \tag{4.214}$$

This means that in the event of default the stock price does not jump. Also the bond part vanishes, and therefore the holder is not entitled to any cashflows, but conversion is allowed at any time after default. This was pointed out by AFV.

We rather give the following interpretation. If we write the credit spread, r_c, as the product of a hazard rate, p, and a loss rate $1 - R$, where R is the recovery rate on the bond part, it can be easily shown that the default value of the convertible turns out to be $V_t^*(S_t, t) = U_t(S_t, t) + RW_t(S_t, t)$. This means that on default the total equity part is recovered, which is consistent with the fact that the stock price does not jump on default, or equivalently the recovery on equity is one. On the other hand, the recovery on the bond part is not zero, which in our understanding is more sensible. Therefore TF can be seen as a special case of (4.175) with zero loss rate η and recovery a fraction of bond and equity part.

Model	Loss rate η	Default value V_τ^*	U_τ^*	W_τ^*
TF	0	$U_\tau + RW_\tau$	U_τ	RW_τ
TKN	1	RV_τ	RU_τ	RW_τ
DL	1	RF	0	RF
AFV (not implemented)	η	$nS_\tau (1-\eta) \wedge RW_\tau$	$(nS_\tau (1-\eta) - RW_\tau) \wedge 0$	RW_τ
AFV total default	1	0	0	0
AFV partial default	0	nS_τ	nS_τ	0

Table 4.10. Comparison of previous models

4.3.6 The Dual Recovery Model

We propose to use a recovery of bond and equity model with splitting 2.

In splitting 2 the equity part is a security that at maturity pays off the value of the parity, a natural definition for the equity part of the CB. The value of the bond part is the value above parity, basically the equity premium. With this splitting the cashflows of the bond part are the actual cashflows of the CB minus the parity; this is consistent with the idea of the bond part being the part at risk, since upon default at a cashflow we can always convert into equity, and therefore the amount at risk is the cashflow minus the parity. Also, as mentioned in Section 4.2, with splitting 2 the final conditions of the bond and equity part are continuous functions of the stock price, as opposed to TF splitting 1, and this facilitates the numerical solution considerably.

Finally we believe it is reasonable to account for the recovery of equity and bond part independently.

To our knowledge this model has not been used before.

4.4 Analytical Solutions for Bond Convertible just at Expiration

We consider a special case for which an analytical solution is available: a zero coupon bond, which is convertible just at expiration, with stochastic interest rate and equity value, but zero correlation between the state variables. Also, we assume constant hazard rate, loss rate and dividend yield and we write

$$\begin{aligned} p_t &= p, \\ \eta_t &= \eta, \\ d_t &= d. \end{aligned}$$

The equation for the total value of the bond V (4.208) reduces to

$$\mathcal{L}V_t - (r_t + p)\,V_t + pV_t^* = 0. \tag{4.215}$$

We provide the analytical solution for the value of the special CB with all the models of the recovery value discussed in Section 4.3.3.

In this Section we introduce the risk-adjusted dividend yield $q = d - p\eta$ and we write X for the strike price that we set to $X = \frac{F}{n}$. Then we define d_1 and d_2 in the usual way as

$$d_1 = \frac{1}{\sigma_S\sqrt{\tau}} \ln\left(\frac{Se^{-q(T-t)}}{XZ\,(r,t;T)}\right) + \frac{1}{2}\sigma_S\sqrt{\tau}, \tag{4.216}$$

$$d_2 = d_1 - \sigma_S\sqrt{\tau}, \tag{4.217}$$

where $Z(r,t;T)$ is the value at time t of a pure discount bond with face value 1, maturing at time T. If we assume a constant deterministic interest rate by setting

$$\theta = \alpha = \sigma_r = \rho = 0, \tag{4.218}$$

we have

$$Z(r,t;T) = e^{-r(T-t)}. \tag{4.219}$$

When interest rates are stochastic, under the Vasicek interest rate model the value of the zero coupon bond is given by

$$Z(r,t;T) = e^{A(t;T)-rB(t;T)}, \tag{4.220}$$

where

$$B(t;T) = \frac{1}{\alpha}\left(1 - e^{-\alpha(T-t)}\right), \tag{4.221}$$

and

$$A(t;T) = \frac{1}{\alpha^2}\left(B(t;T) - (T-t)\right)\left(\theta\alpha^2 - \frac{1}{2}\sigma_r^2\right) - \frac{\sigma_r^2}{4\alpha}B(t;T)^2 \tag{4.222}$$

4.4.1 Analytical Solution with Recovery of Bond Part

For this special case of bond convertible only at expiry, equations (4.182), (4.183) become

$$\mathcal{L}W_t - (r_t + p(1-R))W_t = 0, \tag{4.223}$$

$$\mathcal{L}U_t - (r_t + p)U_t = 0. \tag{4.224}$$

Remark 3 *Notice that equation* (4.223) *for W is a modified Black-Scholes PDE, where the drift and the discount term are risk adjusted. If X is the solution of the original Black-*

Scholes PDE with dividend yield $q = d - p\eta$, *then*

$$W(r,S,t) = e^{-p(1-R)(T-t)} X(r,S,t) \tag{4.225}$$

solves equation (4.223) *above. Similarly, the solution of* (4.224) *is given by*

$$U(r,S,t) = e^{-p(T-t)} X(r,S,t) \,. \tag{4.226}$$

We next provide the analytical solution for this special case under the three splitting procedures described in Section 4.2.

Recovery of Bond Part with Splitting 1:

The final condition in splitting 1 for the cash component W is the payoff of a cash or nothing put with strike F/n and payout F. Therefore the solution of equation (4.223) together with the final condition (4.189) is given by

$$W(r,S,t) = e^{-p(1-R)(T-t)} F \times CONPut(S,t;r,q,T,X) \,, \tag{4.227}$$

where

$$CONPut(S,t;r,q,T,X) = Z(r,t;T)\, N(-d_2) \,. \tag{4.228}$$

Also with splitting 1 the payoff of the equity component U is that of n asset or nothing call options with strike F/n. Hence the solution of equation (4.224) together with the final condition (4.192) is given by

$$U(r,S,t) = ne^{-p(T-t)} AONCall(S,t;r,q,T,X) \,, \tag{4.229}$$

where

$$AONCall(S,t;r,T,X) = Se^{-q(T-t)} N(d_1) \,. \tag{4.230}$$

We conclude that with splitting 1 the total value of the bond, and therefore the solution of equation (4.175) subject to the final condition (4.181) is given by

$$V(r,S,t) = U(r,S,t)+W(r,S,t) = ne^{-p(T-t)}AONCall+e^{-p(1-R)(T-t)}F\times CONPut. \tag{4.231}$$

Recovery of Bond Part with Splitting 2:

With splitting 2 the final condition for the cash component W given in (4.194) is the payoff of n puts with strike F/n. Hence the solution of equation (4.223) together with the final condition (4.194) is given by

$$W(r,S,t) = e^{-p(1-R)(T-t)}nPut(S,t;r,q,T,X), \tag{4.232}$$

where

$$Put(S,t;r,q,T,X) = -\left(Se^{-q(T-t)}N(-d_1) - XZ(r,t;T)N(-d_2)\right). \tag{4.233}$$

Also the payoff of the equity component U is just the value of n underlying stocks. It is straightforward to show that the solution of (4.224) together with the final condition (4.193) is given by

$$U(r,S,t) = e^{-(p+q)(T-t)}nS. \tag{4.234}$$

We conclude that with splitting 2 the total value of the bond, and therefore the solution of equation (4.175) subject to the final condition (4.181) is given by

$$V(r,S,t) = e^{-(p+q)(T-t)}nS + e^{-p(1-R)(T-t)}Put. \tag{4.235}$$

Recovery of Bond Part with Splitting 3:

With splitting 3 the final condition for the cash component W given in (4.196) is the payoff of a straight bond with face value F. Therefore the solution of equation (4.223) together with the final condition (4.196) is given by

$$W(r,S,t) = e^{-p(1-R)(T-t)} FZ(r,t;T). \tag{4.236}$$

The payoff of the equity component U becomes the payoff of n call options with strike F/n. Therefore the solution of equation (4.224) together with the final condition (4.197) is given by

$$U(r,S,t) = e^{-p(T-t)} nCall(S,t;r,q,T,X), \tag{4.237}$$

where

$$Call(S,t;r,T,X) = Se^{-(d-p\eta)(T-t)} N(d_1) - XZ(r,t;T) N(d_2). \tag{4.238}$$

We conclude that with splitting 3 the total value of the bond, and therefore the solution of equation (4.175) subject to the final condition (4.181), is given by

$$V(r,S,t) = e^{-p(T-t)} nCall(S,t;r,q,T,X) + e^{-p(1-R)(T-t)} FZ(r,t;T). \tag{4.239}$$

4.4.2 Analytical Solution with Recovery of Par

Equation (4.175) for the recovery of par case is

$$\mathcal{L}V_t - (r_t + p) V_t + pRF = 0. \tag{4.240}$$

The solution to this PDE subject to the final condition (4.181) can be shown to be

$$V_t = FZ\left(r,t;T\right)e^{-p(T-t)} + \frac{pRF}{r+p}\left(1 - Z\left(r,t;T\right)e^{-p(T-t)}\right) + e^{-p(T-t)}nCall\left(S,t;r,q,T,X\right). \tag{4.241}$$

4.4.3 Analytical Solution with Recovery of Market Value

Equation (4.175) becomes for the recovery of market value case,

$$\mathcal{L}V_t - \left(r_t + p\left(1-R\right)\right)V_t = 0. \tag{4.242}$$

It is straightforward to show that the following function

$$V_t = FZ\left(r,t;T\right)e^{-(p(1-R)(T-t))} + e^{-p(1-R)(T-t)}nCall\left(S,t;r,q,T,X\right), \tag{4.243}$$

satisfies the above PDE together with the final condition (4.181). Notice that we have written the convertible value as the value of a portfolio of a risky bond and n risky call options with strike $\frac{F}{n}$.

4.4.4 Analytical Solution with Recovery of Bond and Equity Part

For the recovery of bond and equity part, the analytical solution under the three splitting procedures can be easily worked out, proceeding as in the previous Section. Results are shown in Table 4.11, where we use the following notations

$$Z = Z\left(r,t;T\right), \qquad C_1 = e^{-p(T-t)}, \qquad C_2 = e^{-p(1-R)(T-t)},$$

$$C_3 = e^{-(p+q)(T-t)}, \qquad C_4 = e^{-d(T-t)}, \qquad C_5 = e^{-p\eta(T-t)}.$$

Recovery		Convertible value V_t
RP		$C_1\,(nCall + FZ) + \frac{pRF}{r+p}\,(1 - C_1)$
RMV		$C_2\,(nCall + FZ)$
	spl1	$C_1 nAONCall + C_2 FCONPut$
RB	spl2	$C_3 nS + C_2 Put$
	spl3	$C_1 nCall + C_2 FZ$
	spl1	$C_5 nAONCall + C_2 FCONPut$
RBE	spl2	$C_4 nS + C_2 Put$
	spl3	$C_5 nCall + C_2 FZ$

Table 4.11. Analytical solution for a zero coupon CB convertible only at expiry in nested models

Remark 4 *Table 4.11 shows that all models lead to different CB values. Only if the probability of default p equals zero, it can be shown using put call parity that all solutions are equal. Differences are due to*

1. *different recovery assumptions in the event of default*
2. *different allocations of the payoff to the CB at the terminal time between U_T and W_T. Once the terminal payoffs are assigned the main difference of value is the recovery assumption.*

4.4.5 Comparison of Nested Models Based on Analytical Solutions

In this Section we show a graphical representation of the analytical solution for the value of a zero coupon bond, convertible only at maturity in all nested model (see Table 4.11). The CB value is plotted against the stock price and interest rates. For simplicity, we assume a constant deterministic interest rate. Parameters are base case parameters as provided in Section 4.6.1. Stock and interest rates vary in what we define as domain of interest. Table

Model	TF	DL	TKN	RB			RBE		
				spl1	spl2	spl3	spl1	spl2	spl3
CB value	89.27	94.42	95.13	86.11	85.09	89.68	89.80	89.53	90.71

Table 4.12. Analytical solution for bond convertible only at expiration with deterministic interest rates in all nested models, for $r = 0.07$ and $S = 100$

4.12 gives the analytical value of a zero coupon bond, convertible only at maturity and with deterministic interest rates, for $r = 0.07$ and $S = 100$ in all models.

Comparison between previous models

Figures 4.5 and 4.6 compare the analytical solution of the simple CB in the TF, DL and TKN models with respect to the stock price and the interest rate respectively. Both Figures show that TKN prices are always above DL prices and TF prices. This can be explained as follows. TKN and DL models differ only on the recovery value assumption. We expect the higher the recovery value the higher the price. TKN assumes recovery of market value and DL recovery of par. Whether the market value of the CB is at a discount or at a premium with respect to the face value depends on the parameters. Although in general, for a convertible, we would expect the market value to be higher than the face value, and therefore prices in TKN to be above prices in DL. This is observed in Figures 4.5 and 4.6. The difference increases as the stock price increases, since the market value will increase, and also for higher interest rates since the present value of the face value will decrease.

Figures 4.5 and 4.6 show that TF model prices are always below DL and TKN. The justification is as follows. TF model differs from DL and TKN in the recovery value but also in the drift term. Although the recovery value in TF is higher than the recovery value

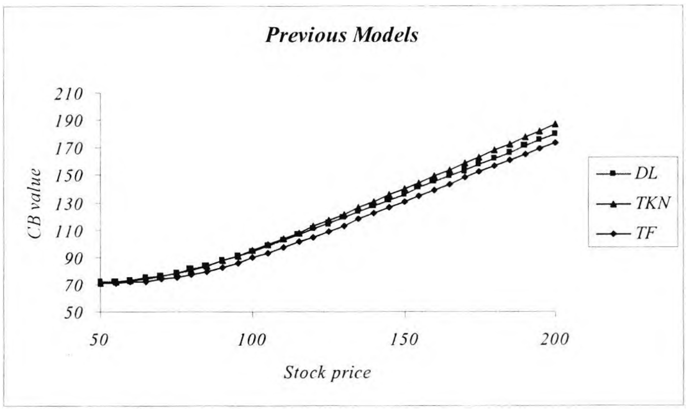

Fig. 4.5. Comparison between previous models. CB value versus stock price.

in TKN and DL,

$$V_t^{TF^*} = U_t + RW_t > RU_t + RW_t = RV_t = V_t^{TKN^*}, \tag{4.244}$$

the instantaneous rate of return in TF is lower

$$\mu_{V_t}^{TF} = r_t - d_t < \mu_{V_t}^{TKN} = \mu_{V_t}^{DL} = r_t - d_t + p_t. \tag{4.245}$$

Because in TF the stock does not jump upon default, we do not need an extra compensation for risk when we hold an option to buy the stock, and this makes the CB value lower than in TKN and DL, where an extra reward for bearing the risk is allowed. The difference is higher for higher levels of the stock and the interest rate.

Figure 4.7 shows the difference between the value of the riskless CB (obtained by setting $p = 0$) and the risky CB value, plotted against the stock price, for TF, DL and TKN models. Figure 4.7 shows that TF risky prices are always less than riskfree prices, and the difference decreases as the stock price increases. On the contrary, for our choice of parameters, Figure 4.7 shows that TKN and DL produce risky prices that are above

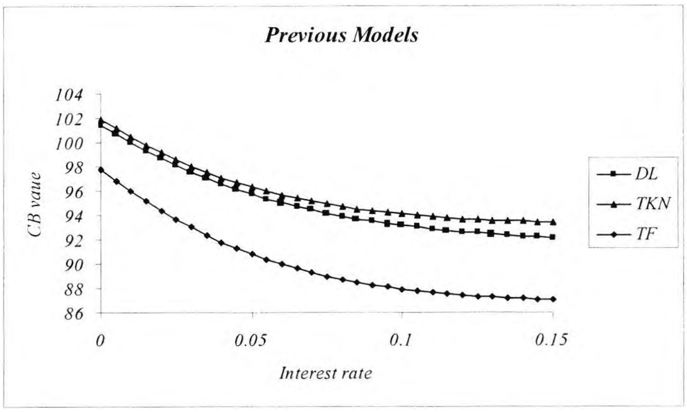

Fig. 4.6. Comparison between previous models. CB value versus interest rate.

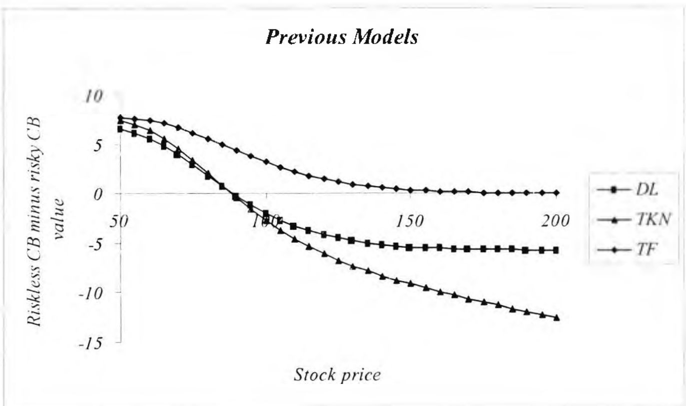

Fig. 4.7. Comparison between previous models. Riskless CB minus risky CB values.

their riskless counterparts. The reason is that in their models the stock price is subject to default risk and an option to buy a risky stock has a higher value than an option to buy a riskless one. If we think of the convertible as a portfolio of debt and equity, we can see that increasing the hazard rate will decrease the value of the debt part but will increase the value of the equity part.

Comparison between splitting procedures

Figures 4.8 and 4.9 compare the different splitting procedures in the recovery of bond part plotted against the stock price and the interest rate respectively. Figures 4.11 and 4.12 do the same for the recovery of bond and equity part model. In all 4 Figures, prices in splitting 3 are the highest, followed by splitting 1. Splitting 2 always produces the lowest prices.

In splitting 3 at maturity the bond pays the face value in all states of the world, in splitting 1 only pays the face value when it is not optimal to convert and in splitting 2 the bond always pays less than the face value. If we denote by W^i the bond part value under splitting i, we have at maturity

$$W_T^3 \geq W_T^1 \geq W_T^2. \tag{4.246}$$

Given that in this special case early exercise is not allowed, it follows that at any time $t < T$ $W_t^3 \geq W_t^1 \geq W_t^2$.

In the recovery of bond part model, when conversion is allowed only at maturity, the default value is given by $V_t^{i^*} = RW_t^i$. Therefore we would expect the following relationship between the CB values

$$V_t^3 \geq V_t^1 \geq V_t^2. \tag{4.247}$$

This relation can be observed on the Figures 4.8 and 4.9.

Figure 4.10 shows the difference between riskless and risky prices in the recovery of bond part model. Risky prices are always below riskfree prices. When the stock price increases, risky price increase towards riskless prices, and then decrease again. This be-

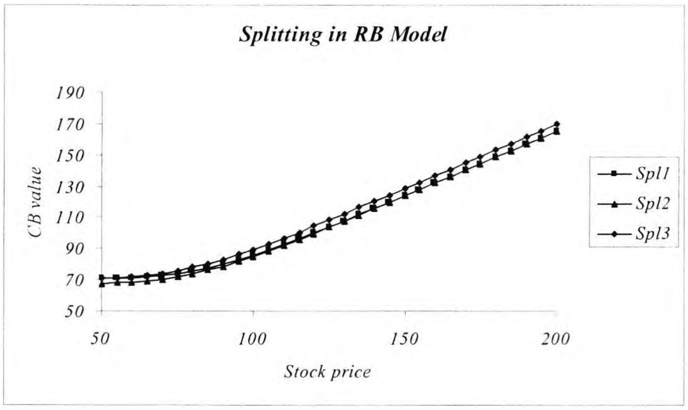

Fig. 4.8. Comparison between splitting procedures in the recovery of bond part model. CB value versus stock price.

haviour is different in the recovery of equity and bond part model (Figure 4.13) in which risky prices converge monotonically to riskless prices as the stock price increases.

Finally comparing Figures 4.8 and 4.9, for RB, with Figures 4.11 and 4.12, for RBE, we can see that the values in the recovery of bond part, if conversion is not allowed, are always below values in the recovery of bond and equity part model. This is easily explained, given that both models differ only in the default values, and the default value in the recovery of bond part, RW, is always below the default value in the recovery of bond and equity part model, $(1 - \eta) U + RW$.

4.5 Numerical Solution

DL and TKN implement their model in a lattice. TF use explicit finite difference and an explicit algorithm to solve the coupled system of equations. AFV use a modified Crank-

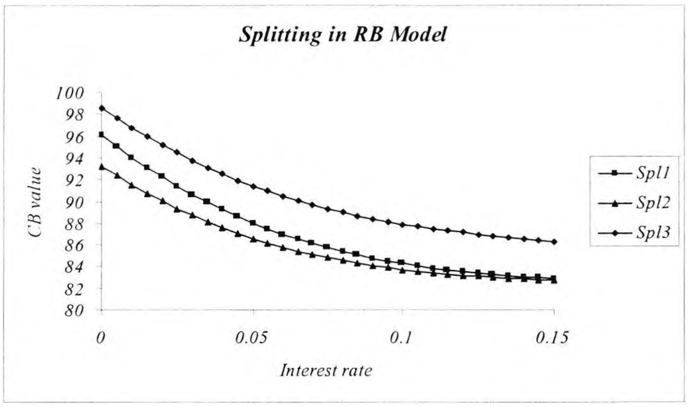

Fig. 4.9. Comparison between splitting procedures in the recovery of bond part model. CB value versus interest rate.

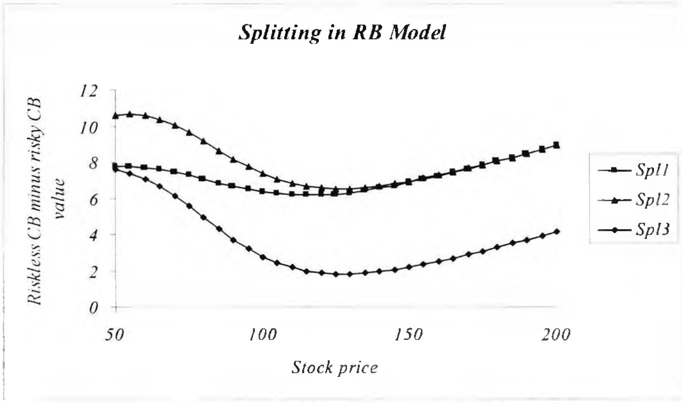

Fig. 4.10. Comparison between splitting procedures in the recovery of bond part model. Riskfree CB minus risky CB value.

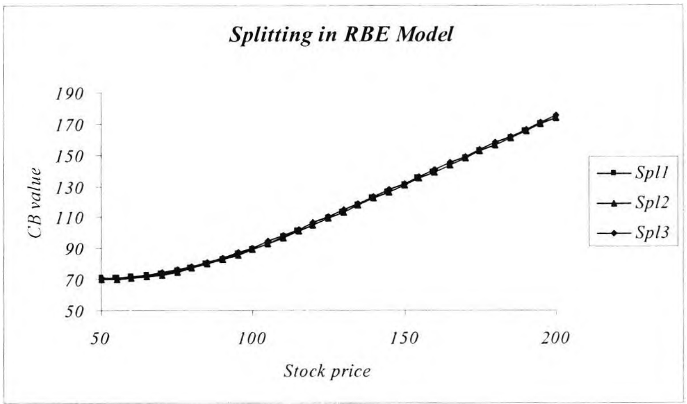

Fig. 4.11. Comparison between splitting procedures in the recovery of bond and equity part model. CB value versus stock price.

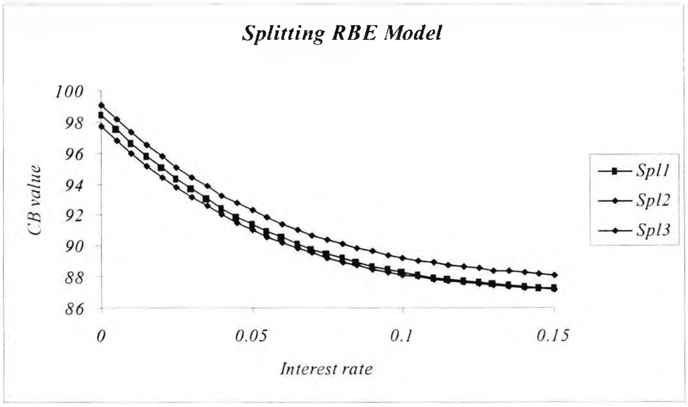

Fig. 4.12. Comparison between splitting procedures in the recovery of bond and equity part model. CB value versus interest rate.

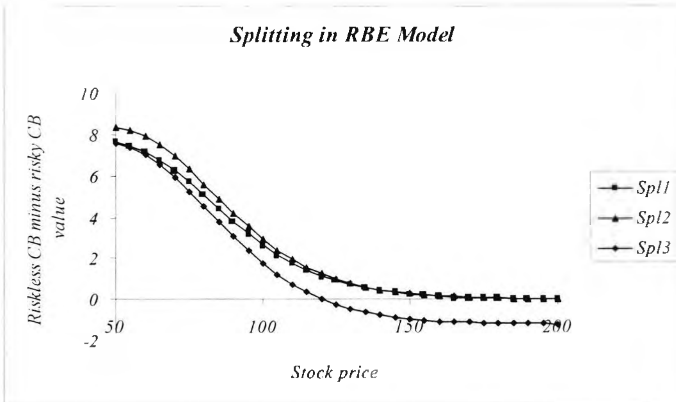

Fig. 4.13. Comparison between splitting procedures in the recovery of bond and equity part model. Riskless CB minus risky CB value.

Nicolson method combined with a penalty method for the free boundaries and an implicit algorithm to solve the coupled system of equations.

We discretize using a Lagrange-Galerkin method, and we use a duality method to deal with the free-boundaries. When it is mandatory to solve the coupled system of equations, we do so for the total value of the CB and the bond part, rather than for the equity and bond components. In order to solve the coupled system of equations (4.175), (4.182) we use an implicit algorithm.

The valuation of convertible bonds can be considered as a special case of the more general two-factor option problem presented in Chapter 2. In Chapter 3 the method was implemented to price convertible bonds ignoring credit risk. The method described in Chapter 2 was for pure diffusion processes. However, as implied from the results in Section 4.1, we can easily extend the solution procedure to jump diffusion processes. Indeed, equation (4.175) differs from equation (3.70), where credit risk was ignored, only by the term $p_t \max\{nS_t(1-\eta_t), RW_t\}$, the drift coefficient and the discount term. The term

$p_t \max\{nS_t(1-\eta_t), RW_t\}$ may be considered as a time and W–depending right hand side term in the equation, and can easily be incorporated in the procedure. Moreover the model in Section 4.1 is a special case of the problem (2.7) – (2.11) for the choices:

$$x_1 = r_t, \tag{4.248}$$

$$x_2 = S_t, \tag{4.249}$$

$$A_{11} = \frac{1}{2}\sigma_r^2,\ A_{12} = A_{21} = \frac{1}{2}\rho S_t \sigma_S \sigma_r,\ A_{22} = \frac{1}{2}\sigma_S^2 S_t^2, \tag{4.250}$$

$$B_1 = \mu_r,\ B_2 = (r_t - d_t + p_t\eta_t), \tag{4.251}$$

$$A_0 = r_t + p_t,\ F = p_t V_t^*(S_t, t), \tag{4.252}$$

and

$$R_1(r_t, S_t, t) = \max\{nS_t, M_{P_t}\}, \tag{4.253}$$

$$R_2(r_t, S_t, t) = \max\{nS_t, M_{C_t}\}. \tag{4.254}$$

Also notice that the problem for W, (4.182), (4.184) together with any of the final conditions (4.189), (4.194) or (4.196), falls also into the general problem (2.7) – (2.11) and can be solved numerically using the characteristics/finite element approach. In fact the problem for W is similar to the one for V but with

$$P_t^W = 0, \tag{4.255}$$

and

$$R_1(r_t, S_t, t) = 0, \tag{4.256}$$

$$R_2(r_t, S_t, t) = \infty. \tag{4.257}$$

4.5.1 Solution of Coupled System of Equations

In the recovery of bond part and in the recovery of bond and equity part it is mandatory to solve a coupled system of equations. We will solve the system consisting of the equation for V (4.175) and the equation for W (4.182) with $P^W = 0$.

In order to solve the coupled system of equations (4.175), (4.182) we use an implicit algorithm. Specifically, at each timestep we carry out the following iterative procedure:

- At the beginning, a function W_0 approximating the solution at that timestep, is given arbitrarily.
- At iteration k an approximation of the bond part value, W_k, is known and we proceed to find W_{k+1}.
- The algorithm stops when two consecutive values of the total value of the bond, V_k and V_{k+1}, differ less than a given tolerance.

The iteration step is as follows:

First, we work out a new approximation of V, V_{k+1}, by solving equation (4.175) subject to unilateral conditions (4.176) – (4.179), the final condition (4.181) and the jump condition (4.180).

Then, we update W_k by solving equation (4.182) subject to the final condition (4.189), (4.194) or (4.196) and jump condition (4.184). Then we "block" the values of the solution, W_{k+1}, according to conditions (4.186), (4.188) and (4.190) (or (4.195)) for the splitting 1 or 2 respectively by means of the Lagrange multiplier values. Specifically we set:

Splitting 1

$$W_{k+1}(r_t, S_t, t) = M_{P_t} \quad \text{if } P_{k+1}(r_t, S_t, t) < 0 \text{ and } nS_t < M_{P_t}, \tag{4.258}$$

$$W_{k+1}(r_t, S_t, t) = 0 \qquad \text{if } P_{k+1}(r_t, S_t, t) < 0 \text{ and } nS_t > M_{P_t}, \tag{4.259}$$

and

$$W_{k+1}(r_t, S_t, t) = 0 \quad \text{if } P_{k+1}(r_t, S_t, t) > 0. \tag{4.260}$$

Also in the region $nS_t \geq M_c$ where the solution was extended by nS_t we set

$$W_{k+1}(r_t, S_t, t) = 0. \tag{4.261}$$

Splitting 2

As for splitting 1 except for condition

$$W_{k+1}(r_t, S_t, t) = M_{P_t} - nS_t \quad \text{if } P_{k+1}(r_t, S_t, t) < 0 \text{ and } nS_t < M_{P_t}. \tag{4.262}$$

4.6 Numerical Results

In this Section we first benchmark the model, investigating the convergence properties of the numerical method. We then compare the consequences of different splitting procedures under the recovery of bond part model. Then we compare convertible bond values obtained solving numerically previous models for CB valuation. Finally we provide some sensitivity analysis for our model, the dual recovery model.

4.6.1 Benchmarking

In order to test the numerical method we consider a particular case for which an analytical solution for the CB value is available: a bond convertible just at expiration with stochastic interest rate but zero correlation between the state variables. This was investigated in detail in Section 4.4.

We set

$$R_1(r_t, S_t, t) = 0, \tag{4.263}$$

$$R_2(r_t, S_t, t) = \infty, \tag{4.264}$$

given that there is no early-exercise embedded options.

We provide benchmarking results for the TF, TKN, and DL[24] and the recovery of bond part (RB) and recovery of bond and equity part (RBE) models. We use for the default specification $p = 0.05$, $\eta = 0.7$ and $R = 0.4$.

The volatility of the underlying stock is $\sigma_S = 15\%$ and its continuous dividend yield is $d = 4\%$. We value a convertible bond with face value of $F = 100$ currency unit and $T = 3.5$ years to maturity. The bond can be converted into $n = 1$ units of the stock. Interest rate parameters are $\theta_t = \theta = 0.07$, $\alpha = 0.1$ and $\sigma_r = 0.02$. We assume zero correlation between the spot interest rate and the underlying stock. The instantaneous interest rate is $r = 0.07$ and the stock price $S = 100$.

Domain bounds are set to be $\Omega^r = [0, 1.5]$ and $\Omega^S = [0, 400]$. Ω^S corresponds to roughly a 99.9% confidence interval on S_T. We give L^2 errors over both the entire domain

[24] For DL we assume deterministic interest rates.

Ω and also over a narrower region of interest $\widehat{\Omega} = \widehat{\Omega}^r \times \widehat{\Omega}^S$, where $\widehat{\Omega}^r = [0, 0.15]$ and $\widehat{\Omega}^S = [50, 200]$. $\widehat{\Omega}^S$ is roughly a 99% confidence interval on S_T. $\widehat{\Omega}$ reflects a range of values of r and S likely to be observed in practice and so the error on $\widehat{\Omega}$ is likely to be more representative.

We present results obtained for successive grid refinements for the relative error in L^2. Mesh 1 is the coarsest with just 15 space steps in the interest rate dimension, 40 in the stock dimension, and 35 time steps up to time $T = 3.5$. Each successive mesh doubles both the number of space steps in each dimension and the number of time steps so that the finest mesh, mesh 4, has 120 interest rate steps, 320 equity steps, and 280 times steps up to three years and a half. For DL and TKN models we double the number of time steps for all meshes, because better convergence was achieved in that way; we also show an extra level of refinement (mesh 5) for these two models. We use as benchmarking measure the total relative error define as

$$\frac{\left[\int_0^T \|error_t\|_{\ell^2}^2 \, dt\right]^{\frac{1}{2}}}{\left[\int_0^T \|solution_t\|_{\ell^2}^2 \, dt\right]^{\frac{1}{2}}},$$

where

$$\|f\|_{\ell^2} =: \left(\int_\Omega f^2 d\Omega\right)^{\frac{1}{2}},$$

and

$$error_t = exact\ solution_t - num\ solution_t.$$

The analytical formulae for the "exact solution" in all nested models is summarized in Table 4.11; the value of the "exact solution" for the current level of the interest rate and the stock price is given in Table 4.13.

Model	TF	DL	TKN	RB			RBE		
				spl1	spl2	spl3	spl1	spl2	spl3
CB value	89.31	94.41	95.16	86.16	85.13	89.72	89.84	89.57	90.75

Table 4.13. Analytical solution for bond convertible only at expiration with stochastic interest rates in all nested models, for $r = 0.07$ and $S = 100$

The numerical results are presented in Tables 4.14-4.22. On the boundaries we use the analytical solution. In each case two of the boundaries are Dirichlet and two are Neumann. 'Error TD' is the error on the entire domain Ω; 'Error RI' is the error on the region of interest, $\widehat{\Omega}$. 'Factor' is progressive error reduction factor in moving to a finer mesh level from the preceding mesh level. Times are in seconds.[25]

The characteristics/finite element method was analyzed by Pironneau (1982) for convection-diffusion equations. Unconditional stability and convergence order of $O(h) + O\left(\frac{h^2}{\Delta t}\right) + O(\Delta t)$ have been proved under suitable conditions for the coefficients of the equation. Although our models do not satisfy the required assumptions, the same error estimate has been obtained empirically. In all Tables we see that the ratio between two consecutive errors tends to 2, which is consistent with the order of convergence given above. Also on the region of interest the convergence rate is faster than on the whole domain. Errors are significantly less on the region of interest compared to the total domain.

In the following Sections, Tables are computed using mesh 4 and asymptotic boundary conditions. All specifications lie within the region of interest so, in line with the errors reported, CB values are reported to 2 decimal places. With early exercise possible and semiannual coupons, a typical execution time is around 2500 seconds if a splitting is specified and 800 seconds otherwise.

[25] The implementation was in Fortran 77 run on a 2.4 Mhz Pentium IV PC.

TF					
$Mesh$	$ErrorTD$	$Factor$	$ErrorRI$	$Factor$	$Time$
1	$1.49E-02$	–	$5.93E-03$	–	1
2	$7.89E-03$	1.90	$2.63E-03$	2.26	8
3	$4.07E-03$	1.94	$1.60E-03$	1.64	119
4	$2.08E-03$	1.96	$1.06E-03$	1.51	1547

Table 4.14. Error and convergence in Tsiveriotis and Fernandes model

TKN					
$Mesh$	$ErrorTD$	$Factor$	$ErrorRI$	$Factor$	$Time$
1	$8.36E-03$	–	$5.25E-03$	–	0
2	$4.59E-03$	1.82	$2.99E-03$	1.76	6
3	$2.45E-03$	1.88	$1.92E-03$	1.56	50
4	$1.28E-03$	1.91	$1.19E-03$	1.61	459
5	$6.66E-04$	1.92	$7.91E-04$	1.51	4811

Table 4.15. Error and convergence in Takahashi, Kobayashi and Nakagawa model

DL					
$Mesh$	$ErrorTD$	$Factor$	$ErrorRI$	$Factor$	$Time$
1	$9.36E-03$		$4.34E-03$		0
2	$5.07E-03$	1.85	$2.28E-03$	1.90	140
3	$2.67E-03$	1.90	$1.45E-03$	1.58	280
4	$1.38E-03$	1.93	$8.10E-04$	1.79	560
5	$7.05E-04$	1.96	$4.28E-04$	1.89	4521

Table 4.16. Error and convergence in Davis and Lischka model

RB with splitting 1					
$Mesh$	$ErrorTD$	$Factor$	$ErrorRI$	$Factor$	$Time$
1	$1.59E-02$	–	$4.10E-03$	–	1
2	$8.36E-03$	1.90	$1.94E-03$	2.11	8
3	$4.31E-03$	1.94	$1.27E-03$	1.53	122
4	$2.20E-03$	1.96	$8.72E-04$	1.46	1813

Table 4.17. Error and convergence in recovery of bond part model: splitting 1

RB with splitting 2					
$Mesh$	$ErrorTD$	$Factor$	$ErrorRI$	$Factor$	$Time$
1	$1.59E-02$	–	$5.51E-03$	–	0
2	$8.39E-03$	1.90	$2.76E-03$	2	8
3	$4.33E-03$	1.94	$1.68E-03$	1.64	119
4	$2.21E-03$	1.96	$1.08E-03$	1.55	1530

Table 4.18. Error and convergence in recovery of bond part model: splitting 2

RB with splitting 3					
$Mesh$	$ErrorTD$	$Factor$	$ErrorRI$	$Factor$	$Time$
1	$1.58E-02$	–	$5.08E-03$	–	1
2	$8.31E-03$	1.90	$2.53E-03$	2.01	8
3	$4.28E-03$	1.94	$1.55E-03$	1.63	119
4	$2.19E-03$	1.96	$9.97E-04$	1.55	1538

Table 4.19. Error and convergence in recovery of bond part model: splitting 3

RBE with splitting 1					
$Mesh$	$ErrorTD$	$Factor$	$ErrorRI$	$Factor$	$Time$
1	$1.55E-02$	–	$4.95E-03$	–	1
2	$8.16E-03$	1.90	$2.48E-03$	2	8
3	$4.21E-03$	1.94	$1.55E-03$	1.6	121
4	$2.15E-03$	1.96	$1.03E-03$	1.5	1585

Table 4.20. Error and convergence in recovery of bond and equity part model: splitting 1

RBE with splitting 2					
$Mesh$	$ErrorTD$	$Factor$	$ErrorRI$	$Factor$	$Time$
1	$1.55E-02$	–	$5.28E-03$	–	1
2	$8.16E-03$	1.90	$2.64E-03$	2	8
3	$4.21E-03$	1.94	$1.62E-03$	1.63	119
4	$2.15E-03$	1.96	$1.05E-03$	1.55	1528

Table 4.21. Error and convergence in recovery of bond and equity part model: splitting 2

RBE with splitting 3					
$Mesh$	$ErrorTD$	$Factor$	$ErrorRI$	$Factor$	$Time$
1	$1.55E-02$	–	$5.21E-03$	–	1
2	$8.15E-03$	1.90	$2.62E-03$	1.99	7
3	$4.20E-03$	1.94	$1.61E-03$	1.63	120
4	$2.14E-03$	1.96	$1.05E-03$	1.54	1535

Table 4.22. Error and convergence in recovery of bond and equity part model: splitting 3

Contract parameters				
Parameter	F	T	n	K
Base	100	3.5	1	3%

Table 4.23. Contract parameter values

Process parameters							
For r				For S			Corr.
r	θ	α	σ_r	S	σ_S	d	ρ
0.07	0.07	0.1	0.02	100	0.15	0.04	0.1

Table 4.24. Process parameter values

4.6.2 Comparison of Nested Models based on Numerical Solutions

In this Section we extend the comparative analysis done in Section 4.4 by allowing stochastic interest rates, correlated with the stock price, continuous conversion and coupon payments. In this case there is no analytical solution for the value of the CB and the comparison is based on the numerical results. First we compare the three splitting procedures in the recovery of bond part model. Then we compare CB values in previous models. The bond can be converted at any time into $n = 1$ unit of the stock. The bond pays a coupon $K = 3$ semiannually. The numerical solution is shown at fixed points of the domain for different moneyness of the option and different interest rate levels. The instantaneous interest rate is $r = 0.07$ and the stock price $S = 100$ for the ATM case, $S = 95$ for the OTM case and $S = 105$ for the ITM case. We show the sensitivity to the credit spread $s = p(1 - R)$. We fix the loss rate, η, and the recovery rate, R, and we vary the hazard rate, p. All parameters are given in Tables 4.23, 4.24 and 4.25. Results have been calculated using mesh 4.

Recovery and loss parameters	
R	η
0.4	0.7

Table 4.25. Recovery and loss parameter values

Comparison Between Splitting Procedures in the Recovery of Bond Part Model

First we show numerical results for the three splitting procedures in the recovery of bond part model.

In the recovery of bond part model the default value is the maximum between the value if converted and a fraction of the bond part. It is mandatory to solve a coupled system of equations, given that we need the value of the bond part W to compute default value of the convertible. We expect that different definitions of the bond part (equivalently different splittings) will lead to different default values and therefore different convertible bond prices.

Table 4.26 shows the numerical solution for the recovery of bond model with splitting 1, Table 4.27 for splitting 2 and Table 4.28 for splitting 3. For this coupon bearing, continuously convertible CB the monotonic relation among the splitting procedures shown in Section 4.4 still holds. Splitting 3 leads to the highest values, followed by splitting 1 and 2. The three Tables show that in all three models CB values decrease as the credit spread, s, increases. The sensitivity to the credit spread is higher for lower stock prices and lower interest rates. Splitting 2 is the most sensitive to movements in the hazard rate, followed by splitting 1 and 2.

Splitting 1									
r	0.06			0.07			0.08		
s	OTM	ATM	ITM	OTM	ATM	ITM	OTM	ATM	ITM
0	109.59	112.32	115.33	108.13	111.02	114.20	106.80	109.85	113.17
0.01	107.59	110.52	113.75	106.30	109.41	112.79	105.14	108.40	111.92
0.02	105.81	108.98	112.41	104.71	108.03	111.60	103.71	107.17	110.87
0.03	104.28	107.66	111.29	103.33	106.85	110.60	102.48	106.13	109.99
0.04	102.96	106.54	110.34	102.15	105.86	109.77	101.43	105.25	109.26
0.05	101.84	105.59	109.55	101.15	105.02	109.07	100.54	104.51	108.64

Table 4.26. Sensitivity to the hazard rate in the recovery of bond part model: splitting 1

Splitting 2									
r	0.06			0.07			0.08		
s	OTM	ATM	ITM	OTM	ATM	ITM	OTM	ATM	ITM
0	109.59	112.32	115.33	108.13	111.02	114.20	106.80	109.85	113.17
0.01	107.41	110.40	113.66	106.17	109.31	112.72	105.03	108.33	111.87
0.02	105.56	108.80	112.28	104.50	107.88	111.50	103.54	107.06	110.79
0.03	104.00	107.46	111.15	103.11	106.69	110.49	102.30	106.00	109.90
0.04	102.69	106.35	110.21	101.93	105.70	109.66	101.25	105.13	109.17
0.05	101.59	105.42	109.43	100.95	104.88	108.97	100.38	104.40	108.57

Table 4.27. Sensitivity to the hazard rate in the recovery of bond part model: splitting 2

	Splitting 3								
r	0.06			0.07			0.08		
s	OTM	ATM	ITM	OTM	ATM	ITM	OTM	ATM	ITM
0	109.59	112.32	115.33	108.13	111.02	114.20	106.80	109.85	113.17
0.01	107.91	110.84	114.03	106.62	109.70	113.05	105.44	108.66	112.14
0.02	106.43	109.54	112.90	105.28	108.53	112.03	104.23	107.62	111.24
0.03	105.12	108.39	111.90	104.09	107.49	111.13	103.16	106.69	110.44
0.04	103.94	107.37	111.01	103.03	106.57	110.34	102.20	105.86	109.73
0.05	102.89	106.46	110.23	102.08	105.75	109.63	101.34	105.13	109.11

Table 4.28. Sensitivity to the hazard rate in the recovery of bond part model: splitting 3

TF									
r	0.06			0.07			0.08		
s	OTM	ATM	ITM	OTM	ATM	ITM	OTM	ATM	ITM
0	109.59	112.32	115.33	108.13	111.02	114.20	106.80	109.85	113.17
0.01	107.81	110.74	113.95	106.51	109.60	112.96	105.33	108.57	112.07
0.02	106.10	109.23	112.63	104.95	108.23	111.77	103.91	107.34	111.01
0.03	104.46	107.78	111.36	103.45	106.92	110.63	102.56	106.16	110
0.04	102.88	106.39	110.14	102.01	105.66	109.54	101.25	105.04	109.03
0.05	101.35	105.04	108.97	100.63	104.45	108.50	100	103.95	108.10

Table 4.29. Sensitivity to the hazard rate in Tsiveriotis and Fernandes model

TKN									
r	0.06			0.07			0.08		
s	OTM	ATM	ITM	OTM	ATM	ITM	OTM	ATM	ITM
0	109.59	112.32	115.33	108.13	111.02	114.20	106.80	109.85	113.17
0.01	108.97	112.05	115.40	107.73	110.96	114.45	106.60	109.98	113.60
0.02	108.78	112.21	115.88	107.74	111.30	115.09	106.79	110.49	114.39
0.03	108.97	112.73	116.69	108.10	111.97	116.04	107.31	111.30	115.45
0.04	109.49	113.55	117.78	108.76	112.92	117.23	108.10	112.35	116.74
0.05	110.29	114.62	119.09	109.68	114.08	118.62	109.13	113.61	118.20

Table 4.30. Sensitivity to the hazard rate in Takahashi, Kobayashi and Nakagawa model

Comparison of Default Specification in Previous Models (TF, TKN, DL)

In this Section we compare convertible prices based on different previous models. We present the numerical solution at fixed levels of stock and interest rate obtained for TF, TKN and DL models in Tables 4.29, 4.30 and 4.31 respectively. The numerical results in this Section are also consistent with the results based on the analytical solution for the special CB with deterministic interest rates in Section 4.4. TKN gives the highest CB values, followed by DL and TF.

Table 4.29 shows that in TF model the CB values decrease as the credit spread increases. On the contrary, in TKN (see Table 4.30) and DL (see Table 4.31) models, prices

DL									
r	0.06			0.07			0.08		
s	OTM	ATM	ITM	OTM	ATM	ITM	OTM	ATM	ITM
0	109.59	112.32	115.33	108.13	111.02	114.20	106.80	109.85	113.17
0.01	108.79	111.82	115.11	107.56	110.73	114.16	106.43	109.75	113.31
0.02	108.38	111.68	115.21	107.33	110.76	114.41	106.38	109.94	113.70
0.03	108.28	111.82	115.56	107.39	111.04	114.88	106.58	110.34	114.26
0.04	108.43	112.18	116.09	107.67	111.51	115.50	106.97	110.90	114.96
0.05	108.78	112.69	116.74	108.12	112.10	116.21	107.51	111.57	115.73

Table 4.31. Sensitivity to the hazard rate in Davis and Lischka model

TKN/DL/AFV Total Default with $R = 0$									
r	0.06			0.07			0.08		
s	OTM	ATM	ITM	OTM	ATM	ITM	OTM	ATM	ITM
0	109.59	112.32	115.33	108.13	111.02	114.20	106.80	109.85	113.17
0.01	107.86	110.79	114.00	106.57	109.66	113.01	105.39	108.63	112.13
0.02	106.34	109.46	112.85	105.20	108.47	111.99	104.16	107.58	111.23
0.03	105.00	108.31	111.86	104.00	107.44	111.12	103.09	106.66	110.46
0.04	103.83	107.30	111.00	102.95	106.55	110.36	102.15	105.87	109.79
0.05	102.80	106.43	110.27	102.03	105.78	109.72	101.34	105.19	109.23

Table 4.32. Sensitivity to the hazard rate in in Ayache, Forsyth and Vetzal Total Default

decrease when the credit spread increases only for low credit spreads. For high credit spreads prices increase above their riskfree counterpart.

For zero recovery rate TKN and DL models are identical and they also agree with AFV Total Default model ($\eta = 1, R = 0$). Table 4.32 shows the numerical solution for this case. Prices are below TKN and DL with no zero recovery rate as it would be expected. But prices are higher than in TF with non zero recovery rate.

For completeness we also show results for AFV Partial Default model with zero recovery rate ($\eta = 0, R = 0$) in Table 4.33. AFV Partial Default model is similar, but not identical to TF. In TF the full market value of the equity part is recovered on default. In AFV Total Default model, the value upon conversion nS is recovered instead. In TF model the equity part U is a security that pays at maturity either 0 or the equity value upon

AFV Partial Default with $R = 0$									
r	0.06			0.07			0.08		
s	OTM	ATM	ITM	OTM	ATM	ITM	OTM	ATM	ITM
0	109.59	112.32	115.33	108.13	111.02	114.20	106.80	109.85	113.17
0.01	109.16	111.95	115.03	107.75	110.70	113.93	106.45	109.56	112.94
0.02	108.74	111.60	114.73	107.38	110.39	113.67	106.12	109.29	112.71
0.03	108.34	111.26	114.44	107.01	110.09	113.41	105.80	109.02	112.48
0.04	107.95	110.93	114.17	106.67	109.79	113.17	105.49	108.76	112.27
0.05	107.57	110.61	113.90	106.33	109.51	112.93	105.19	108.51	112.06

Table 4.33. Sensitivity to the hazard rate in in Ayache, Forsyth and Vetzal Partial Default

conversion nS (see 4.192). Therefore, if default occurs at maturity, in TF model we may recover $U_T = 0 < nS_T$ when $S_T < F$; a similar argument applies to any other default time $\tau < T$. That is why AFV Partial Default model produces prices above TF.

Notice that the total default and partial default models reduce to one PDE only, and therefore there is no need to solve the couple system of equations.

4.6.3 Pricing in the Dual Recovery Model

In this Section we explore the effect upon CB values of altering parameter values within our dual recovery model, looking particularly at the exercise conditions, equity and interest rate values and parameters, the recovery parameters R and η, and the default parameters k, a, b.

Each parameter has a base case value, and a high and a low value. These are given in Tables 4.34, 4.35, 4.36 and 4.37. For the base case we suppose that the CB has $T = 3.5$ years to maturity with face value $F = 100$. The CB may be converted at any time with conversion ratio $n_t \equiv n = 1$. The CB pays a coupon of 3% every half year, an annual dividend yield of 4%. It is callable and redeemable at any time with call and redemption

Contract parameters			
Parameter	$Face\ value$	$Maturity$	K
Base	100	3.5	3%
High	105	5	4%
Low	95	1	2%

Table 4.34. Contract parameter values

Exercise parameter values			
Parameter	$Conversion\ ratio$	$Call\ Price$	$Put\ Price$
Base	1	110	95
High	1.1	115	100
Low	0.9	105	90

Table 4.35. Exercise parameter values

prices given by $M_C = 100$ and $M_P = 95$. The initial stock value is $S_0 = 100$ and initial interest rate is $r_0 = 0.07$. For the default intensity function we set $k = 0.15$, $a = 0.015$, $b = 1.5$, giving $p_t \in [0, 0.7]$ over the computational domain, with $p_t = 0.03$ in the base case. For middling values of S_t and r_t, p_t has about the same sensitivity to changes in each. Other parameter values are given in the Tables.

Exercise Conditions

We investigate the effect of the presence or absence of the various exercise conditions. We first consider a riskless coupon bond and then, we add default and various combinations of exercise conditions, ending with the full specification of the base case CB. We also

Process parameters								
	For r				For S			Corr.
Parameter	r	θ	α	σ_r	S	σ_S	d	ρ
Base	0.07	0.07	0.1	0.02	100	0.15	0.04	0.1
High	0.08	0.08	0.15	0.03	105	0.20	0.05	0.15
Low	0.06	0.06	0.05	0.01	95	0.10	0.03	0.05

Table 4.36. Process parameter values

Default, recovery and loss parameters					
	Default			Recovery	Loss
	k	a	b	R	η
Base	0.15	0.015	1.5	0.4	0.7
High	0.2	0.03	3	0.5	0.8
Low	0.1	0.003	0.3	0.3	0.6

Table 4.37. Default and recovery parameter values

give an approximation to the value of $\partial V_t/\partial r$ found by central difference from CB values computed at different initial values of the interest rate.

Tables 4.38 and 4.39 show the results. 'Def' is defaultable (with 'dual' recovery), 'Con' is convertible, 'Red' is redeemable and 'Call' is callable. Δ and Γ are the CB delta and gamma respectively. The riskless bond values are Vasicek values computed analytically and shown for comparison. The base case value of the CB is 104.14, shown in bold. Table 4.38 shows the standard case. For comparison, Table 4.39 shows the effect of reducing the conversion factor from 1 to 0.9. We first discuss Table 4.38.

With our specification and model parameters, the presence of default reduces the value of the corresponding riskless bond by a little over 6%. The bond has a high credit risk, higher for low interest rate and low stock prices.

Adding the conversion feature increases the value of the CB by around 18.5% in the base case. The effect is greater at higher values of S_0 as the CB becomes more in the money, and at higher values of r. The introduction of the call feature reduces the value of the CB by around 4%. The reduction is greater at higher values of S_0 as the CB is more likely to be called. Adding the redemption feature has very little effect on the CB value, with our specification.

Equity deltas are significant. They vary only a little as the initial interest rate changes, increasing slightly as the interest rate increases. Introducing conversion to a straight defaultable bond increases the delta by a factor of 10. The call feature does affect the stock delta, reducing it by 17% in the basecase. Adding a redemption feature also affects the equity delta, reducing it by a bit less than adding a call feature. When added together, call and put features reduce the equity delta by around 20%. The basecase CB has delta 0.5 as it would be expected from an at the money issue.

$\partial V_t/\partial r$, the CB's rho, indicates the sensitivity of the CB to changes in the initial value of the interest rate. The conversion feature reduces the absolute size of the CB's rho by around a third, depending on the initial stock value. The redemption feature has little effect, but the call feature reduces rho at higher stock values.

Allowing the riskless bond to become defaultable reduces rho very little but adding additional optionality reduces it further, by about 80% in the base case. For this CB, additional optionality effectively decreases the interest rate exposure of the CB. Conversion significantly increases its equity value exposure, but callability and putability smooth this effect.

Table 4.39 shows how the situation changes if the conversion ratio is significantly reduced, to $n = 0.9$. Now the CB is out of the money.

Adding convertibility increases the value of the defaultable bond by about 12%. Adding the call increases the value by around 2%, half the effect of the $n = 1$ case. The redemption features has little effect on the CB value, although the effect doubles with respect to the $n = 1$ case. Sock deltas are affected, but by less that in the $n = 1$ case when adding

Exercise Conditions				S_0	r			$\partial V_t/\partial r$
Def	Con	Red	Call		0.06	0.07	0.08	
Riskless Bond				100	99.43	96.76	94.16	-263.1
√	-	-	-	95	93.32	90.91	88.57	-237.9
(Defaultable bond)				100	93.75	91.32	88.96	-239.6
				105	94.15	91.70	89.32	-241.3
				Δ	0.08	0.08	0.08	
				Γ	0.00	0.00	0.00	
√	√	-	-	95	105.77	104.75	103.82	-97.7
				100	109.11	108.21	107.39	-85.8
				105	112.66	111.87	111.16	-74.8
				Δ	0.69	0.71	0.73	
				Γ	0.01	0.01	0.01	
√	√	√	-	95	106.22	105.31	104.50	-85.9
				100	109.39	108.57	107.82	-78.3
				105	112.84	112.09	111.43	-70.2
				Δ	0.66	0.68	0.69	
				Γ	0.01	0.01	0.01	
√	√	-	√	95	102.19	101.30	100.49	-85.3
				100	104.45	103.81	103.23	-61.1
				105	107.05	106.69	106.37	-33.9
				Δ	0.49	0.54	0.59	
				Γ	0.01	0.01	0.02	
√	√	√	√	95	102.65	101.88	101.20	-72.4
(Base case)				100	104.71	**104.14**	103.63	-53.8
				105	107.16	106.84	106.55	-30.7
				Δ	0.45	0.50	0.53	
				Γ	0.02	0.02	0.02	

Table 4.38. Effect of Exercise Features, base case, $n = 1$

callability, and by more when adding the put feature. Rhos are reduced, but by much less than in the $n = 1$ case.

The CB of Table 4.39 is relatively 'bond-like' whereas that of Table 4.38 is much more 'equity-like'.

Table 4.40 shows the affect of changing exercise conditions. Since the CB is at the money changing the conversion ratio has a significant effect. Changing the redemption level has little effect but changing the call level has a large effect for this at the money CB.

Exercise Conditions				S_0	r			$\partial V_t/\partial r$
Def	Con	Red	Call		0.06	0.07	0.08	
Riskless Bond				100	99.43	96.76	94.16	−263.1
				95	93.32	90.91	88.57	−237.9
√	-	-	-	100	93.75	91.32	88.96	−239.6
(Defaultable bond)				105	94.15	91.70	89.32	−241.3
				Δ	0.08	0.08	0.08	
				Γ	0.00	0.00	0.00	
				95	100.71	99.38	98.15	−127.9
√	√	-	-	100	103.19	101.98	100.87	−115.7
				105	105.91	104.83	103.83	−103.7
				Δ	0.52	0.54	0.57	
				Γ	0.01	0.01	0.01	
				95	101.59	100.52	99.56	−101.7
√	√	√	-	100	103.79	102.77	101.84	−97.5
				105	106.32	105.36	104.49	−91.5
				Δ	0.47	0.48	0.49	
				Γ	0.01	0.01	0.01	
				95	98.82	97.49	96.26	−128.2
√	√	-	√	100	100.57	99.41	98.34	−111.4
				105	102.39	101.44	100.55	−91.6
				Δ	0.36	0.39	0.43	
				Γ	0.00	0.00	0.01	
				95	99.77	98.73	97.80	−98.5
√	√	√	√	100	101.21	**100.26**	99.40	−90.8
(Base case)				105	102.81	101.99	101.24	−78.2
				Δ	0.30	0.33	0.34	
				Γ	0.01	0.01	0.01	

Table 4.39. Effect of Exercise Features, low conversion ratio $n = 0.9$

Exercise parameter values			
Parameter	*Conversion ratio*	*Call Price*	*Put Price*
High	110.20	105.58	105.03
Low	100.26	102.28	103.84

Table 4.40. Sensitivities to changes in exercise conditions

	r-parameters				S-parameters			Corr.
Parameter:	r	θ	α	σ_r	S	σ_S	d	ρ
High:	103.63	104.09	104.12	104.27	106.84	104.84	103.76	104.17
Low:	104.71	104.19	104.16	104.00	101.88	103.31	104.50	104.11
Delta:	−53.8	−5	−0.4	13.5	0.5	15.3	−36.7	0.7
Gamma:	632	50	1	−137	0	−54	−186	0

Table 4.41. Sensitivities to changes in parameter values

Parameter Deltas and Gammas

We investigate the sensitivity of the base case CB to changes in parameter values. We value the CB at the higher value and lower value of each parameter. The delta and gamma are then computed by central differences. Results are given in Table 4.41. r is the initial value of the stochastic Vasicek interest rate.

Deltas are significant. The stock volatility σ_S has similar delta to the interest rate volatility σ_r. Increasing the correlation ρ slightly increases the bond value. θ, the level to which r_t reverts, has a slightly smaller delta, but still significant since it reflects the longer term value of r_t. The sensitivity to the dividend yield is very high, as it could be expected for an ATM issue.

The Default Specification

We explore the consequences of changing the default specifications. Tables 4.42, 4.43 and 4.44 summarizes the results.

Care must be taken in interpreting these Tables. They assume that the initial stock value S_0 and the default and loss parameters (R, k, a, b, η) may be determined independently, so that, for instance, the default rate might increase while S_0 remains fixed. In practice an increase to p might be expected to cause S_0 to fall, so that the CB value would be computed for a reduced value of S_0.

Note that these considerations do not affect the practical implementation of a model. Calibrating to market data fits to mutually determined values of S_0 and the default and loss parameters, so CB values are correctly determined.

Table 4.42 shows sensitivities to the recovery rate, R, on the bond part and base, high and low values of p, conditional on S_0 remaining fixed. The effect of the recovery rate is very little, specially for low probability of default. Increasing the probability of default decreases the bond value by almost 3%, whereas decreasing the probability of default increases the convertible value by much less.

Table 4.43 shows the effect on CB values of k, a and b individually taking high or low values, conditional on fixed S_0. The movements are in the expected direction. Changes in a, the coefficient of the stock price in the hazard rate function, have the bigger effect, followed by k, the hazard rate for zero interest rate and stock price, and b, the coefficient of the interest rate.

Table 4.44 shows the effect upon the CB value of varying the loss rate, η, for different initial stock values S_0. Changing η has two opposite effects in the value of the CB. On one hand, $1 - \eta$ is the recovery on the equity part of the bond. Therefore, we expect that an increase in η will lead to a lower recovery value and consequently lower value of the convertible. On the other hand η appears in the drift of the stock price (and the CB value). The higher the loss rate on the stock, the higher the return required to compensate the risk, and therefore the higher the CB value. With our parameter choice the two effects

Recovery Rate	Hazard rate, p_t		
	Base	High	Low
Base	104.14	101.34	105.58
High	104.19	101.38	105.59
Low	104.10	101.29	105.57

Table 4.42. Sensitivities to changes in default parameters

Parameter	Default Parameters		
	k	a	b
High	103.70	105.38	104.27
Low	104.65	102.09	104.03

Table 4.43. Sensitivities to changes in default parameters

almost cancel each other, although the effect on the recovery value seems more important; increasing the loss rate decreases slightly the value of the CB for all levels of the stock price.

Contract Parameters

Table 4.45 shows the effect on the CB value of changing the face value, the maturity and the cash value of the coupon payments.[26] Increasing the face value or the maturity of the contract has very little effect on the value of the CB; decreasing any of them has a significant effect. Changing the coupon has some effect, as it would be expected from an ATM issue.

[26] When T increases or decreases, the time step Δt is held constant and the number of time steps is varied.

S_0	Loss rate, η		
	Base	High	Low
Base	104.14	103.98	104.31
High	106.84	106.75	106.93
Low	101.88	101.66	102.12

Table 4.44. Effect of Different Loss Rates

Contract parameters			
Parameter	*Face value*	*Maturity*	*Coupon*
High	104.77	104.22	105.71
Low	101.88	101.66	102.11

Table 4.45. Effect of contract parameters

Initial Conditions		Constant r			$\partial V_t/\partial r$
		r			
		0.06	0.07	0.08	
S_0	95	102.40	101.61	100.93	−73.2
	100	104.52	103.93	103.43	−54.4
	105	107.05	106.72	106.43	−30.8
Δ		0.47	0.51	0.55	
Γ		0.02	0.02	0.02	

Table 4.46. Effect of a stochastic interest rate

The Effect of a Stochastic Interest Rate

We have seen the effect upon the bond value of changes in the parameters of the interest rate process. We can also test the effect of a stochastic interest rate. By setting $\sigma_r = 0$ and $r = \theta$ we effectively make r non-stochastic. We investigate the presence of a stochastic interest rate in more detail. Table 4.46 gives the results, looking at several sets of initial conditions for the basecase convertible.

Comparing to Table 4.38, we see that making r constant at its initial value has the effect of decreasing the value of the CB although the effect is very little. Making r constant slightly increases the rho and the delta.

4.7 Conclusions

The main contribution of this Chapter is to provide a valuation framework for contingent claims with equity, interest rate and default risk. It allows for early-exercise embedded options, modelled via variational inequalities. The model allows calibration of interest rates to the actual yield curve, and hazard rates to market prices of derivatives as well as time series data. We identify most of the previous models for CB valuation as special cases of our general framework and we extend them to two factors. We also propose three possible splitting of the CB value into bond and equity parts and we introduce new models for the recovery value. We analyze the different models and we make a choice among the splittings and recovery specification, leading to our dual recovery model. In the dual recovery model the bond and equity part are defined as the equity premium and the parity respectively. Moreover, the recovery is specified separately in the bond and equity part as a fraction of the market value prior to default.

We propose an implicit algorithm to solve the coupled system of equations arising from the splitting procedures. A variational formulation of the problem is the starting point to carry out discretization using characteristics and finite element methods. An iterative algorithm is applied over the discretized problem to deal with the free boundaries arising from the embedded call, put and conversion options. We benchmark and study the convergence.

Based on numerical results and analytical solutions, we compare consequences of different model specifications, regarding assumptions about hazard rate function, recovery value and stock price behaviour.

We find that different model specifications lead to significantly different CB prices. In particular some models produce risky prices that are below their risk-free counterpart. We provide some justification for the differences found.

We conclude that it is necessary to choose the recovery value consistently with the hazard rate and the loss rate. In view of our analysis, this seems easier to do if the convertible value is split and default is specified independently in bond and equity parts. In particular, our dual recovery model provides reasonable results for a great number of parameter specifications. On the other hand, it has enough degrees of freedom to produce many different default-recovery scenarios.

Chapter 5
An Asset Based Model of Defaultable CBs with Endogenised Recovery

In the previous Chapter we presented an equity based model. This Chapter proposes another two-factor model for CBs in which the state variables are the firm's asset value, V_t, and the short interest rate, r_t. As discussed in Chapter 1, firm value models have the disadvantage that they require the estimation of the volatility of the firm's asset, which is not observed. Besides, they lead to very complicated models for complex capital structures. However, they provide a natural link among the equity and debt, which is ideal to account for the hybrid nature of CBs. Also, as we will demonstrate in this Chapter, default risk is more easily incorporated in an asset based model.

We assume that the firm has a single debt class composed of convertible bonds. The CB defaults either at the unpredictable jump time of a counting process or, potentially, when the firm is required to make a cashflow to the CB. We endogenise recovery upon default into our model by assuming that the firm can invoke temporary protection against its creditors, leading to a quantification of the recovered value of the claim against the firm.[27]

We recall from Chapter 1 that credit risk models fall into two main categories, structural and reduced form. In structural models default occurs when a state variable, usually V_t, breaches a barrier level. It is necessary to specify the process for V_t, the location of the

[27] This Chapter is based on Bermúdez and Webber (2004).

barrier, and the form and amount of recovery upon default. In reduced form models default is exogenous, occurring at the jump time of a counting process, N_t, with jump intensity λ_t. The main issues in reduced form models are the specification of processes for the riskless short rate r_t, the hazard rate λ_t, and the loss rate w_t. Our model has both reduced form and structural features.

The early models of convertible bonds (Ingersoll (1977a) and Brennan and Schwartz (1977)) follow Merton (1973) in using V_t with geometric Brownian motion as the sole state variable. Brennan and Schwartz (1980) and more recently Nyborg (1996) and Carayannopoulos (1996) include in addition a stochastic interest rate. Default risk is usually incorporated structurally by capping payouts to the bond by the value of the firm.

The majority of the recent literature uses the stock price, S_t, with geometric Brownian motion as the main state variable, incorporating either an interest rate variable, or default. or both.

Dimensional problems restrict the use of more than two factors in a model. Single factor models that incorporate default risk and do so in the reduced form framework include Andersen and Buffum (2003), Takahashi, Kobayahashi and Nakagawa (2001) and Ayache, Forsyth and Vetzal (2002), (2003).[28] Davis and Lischka (2002) present a reduced form modelling framework with several graphical comparisons. Tseveriotis and Fernandes (1998) and Yigitbasioglu (2002), extended by Ayache, Forsyth and Vetzal (2002), (2003), split the CB value into a bond part and an equity part each with its own discount rate. Tseveriotis and Fernandes and Yigitbasioglu impose a fixed credit spread between the discount

[28] Takahashi, Kobayahashi and Nakagawa also discuss a structural model of default.

rates. Ayache, Forsyth and Vetzal and Takahashi, Kobayahashi and Nakagawa determine the credit spread via a hazard rate in a reduced form framework. All of these models are equity based, with S_t as the main state variable.

Care is required to specify correctly what happens to the convertible bond when default occurs. When the underlying state variable is the asset value V_t, it is relatively easy to so do in a logically consistent way. When the state variable is the equity value S_t, considerable difficulties may arise. For instance, boundary conditions are hard to specify in a financially consistent manner; some models may not require that when $S_t \to 0$ the bond value goes to zero.

To avoid specification problems inherent in the models based upon S_t, we choose V_t as our primary state variable, taking care to impose financially consistent boundary conditions. We are not aware of other reduced form specifications that model default when V_t is the state variable.

We obtain values by solving numerically a partial differential inequality (PDI) using the method described in Chapter 2.

We find that our recovery specification allows a wide range of behaviour upon default. For illustrative examples, we find that the sensitivity of the value of the convertible bond to changes in the asset value and interest rates depends crucially on the specification of a conversion feature. Its sensitivity to the call and redemption values also depends on the conversion feature specification.

The first Section of the Chapter describes the CB valuation model. The second Section describes the numerical method. The third Section presents numerical results, and the final Section concludes.

5.1 An Asset Based Model for Convertible Bond Valuation

Let V_t denote the value of a firm's assets, S_t the value of its equity and D_t the value of the firm's debt, so that $V_t = S_t + D_t$.

Suppose that the value V_t of the firm's assets follows a jump-augmented geometric Brownian motion under the objective measure,

$$\mathrm{d}V_t = \mu_t V_{t\_}\mathrm{d}t + \sigma_V V_{t\_}\mathrm{d}z_t^V - w_t V_{t\_}\mathrm{d}N_t, \tag{5.265}$$

where z_t^V is a standard Brownian motion and N_t is a counting process with intensity λ_t. w_t is a proportional loss. N_t models exogenous default events. At a jump time τ for N_t the asset value falls by a proportion w_τ,

$$V_\tau = V_{\tau\_}\left(1 - w_\tau\right). \tag{5.266}$$

Since we focus on asset risk and interest rate risk we assume that w_t is non-stochastic. Under the equivalent martingale measure (EMM) associated with the accumulator numeraire $B_t = \exp\left(\int_0^t r_s \mathrm{d}s\right)$ the relative price V_t/B_t is a martingale so

$$\mathrm{d}V_t = \left(r_t + \bar{\lambda}_t w_t\right) V_{t\_}\mathrm{d}t + \sigma_V V_{t\_}\mathrm{d}z_t^V - w_t V_{t\_}\mathrm{d}N_t, \tag{5.267}$$

where r_t is the instantaneous short rate, $\bar{\lambda}_t = \lambda_t \gamma(t)$ is the jump intensity under the EMM and $\bar{\lambda}_t w_t dt$ is the compensator for the jump component of V_t. As discussed in the previous

Chapter, $\gamma(t)$ is the measure adjustment of the hazard rate as given by Girsanov's theorem (see Jacod and Shiryaev (1988)).

We suppose that the firm has issued a convertible bond with market value D_t at time t. The bond matures at time T with face value F. At certain times t_i, $i = 1, \ldots, N$, $t_N = T$, it pays coupons of size P_{t_i}, and $V_{t_i} = V_{t_i-} - P_{t_i}$. At certain times up to and including time T the bond may be converted to equity. Its value upon conversion at time t is $\kappa_t V_t$ where κ_t is the proportion of the firms asset value acquired by the debt holders.[29] Dilution effects are absorbed into κ_t.

We assume that the CB may be both callable and redeemable with call price C_t and redemption price R_t at certain times t. On any particular date the CB need be neither callable nor redeemable but we assume that if it is callable on some date then it is also convertible on that date. In the sequel we suppose that the call price and redemption price are imputed to accrue interest on coupons and that if $t_i \leq t < t_{i+1}$ for coupon payment dates t_i the call price and redemption price are set to be

$$C_t = C + \frac{t - t_i}{t_{i+1} - t_i} P_{t_{i+1}}, \tag{5.268}$$

$$R_t = R + \frac{t - t_i}{t_{i+1} - t_i} P_{t_{i+1}}, \tag{5.269}$$

for constants C and R.

If the bond is redeemed, or if a coupon or principal is to be paid, we suppose that the firm may choose to default. We assume that if the firm defaults, whether exogenously or endogenously, the CB holders may choose to convert.

---

[29] Unlike Ayache, Forsyth and Vetzal (2002) and Tsiveriotis and Fernandes (1998) we account for the effect of conversion upon the value of the firm's equity. At conversion, the firm's total value is unchanged but it becomes all equity.

Our default specification has both reduced form and structural elements. Summarizing, a default event may occur in one of two ways. Firstly, when the counting process N_t jumps, the firm is supposed to have been hit by an unexpected exogenous default. Secondly, when a claim is made against the firm, specifically when the CB is redeemed or when a coupon or principal payment is due to be made, the firm may choose to default.

We first give a detailed specification of the components of the model. Then we display the PDI obeyed by the convertible bond value in this framework, and its boundary conditions.

5.1.1 Detailed Specification of the Model

To specify a model we need to define what happens to the CB value when default occurs, define the hazard rate process $\bar{\lambda}_t$, and provide an interest rate model. We consider each of these in turn. Finally we bring together the separate components into a fully specified model with a consistent set of boundary conditions and inequality constraints.

The Default Event and Recovery Values

So far no assumptions have been made about what happens upon default. We now assume that at the time τ of a default event the firm loses the right to call the debt, and that CB holders may no longer redeem the debt, but upon default the CB holders have the option to convert.[30] We write D^*_τ for the value of the CB at a default time τ and F^*_τ for the recovery value of the CB at time τ. Since bondholders have the option to convert in the

---

[30] We see below it may indeed be optimal for the CB holders to do so.

event of default we have

$$D^*_\tau = \max\{F^*_\tau,\ \kappa_\tau V_\tau\}. \tag{5.270}$$

Now we consider the recovery value F^*_τ of the convertible bond upon default.

As reviewed in Chapter 1, many different assumptions are made in the credit literature about the recovery value F^*_τ of a defaulted bond. The most common models are the recovery of market value (RMV) and the recovery of treasury (RT). The RMV model assumes that the ratio $l_\tau = \left(D_{\tau_-} - F^*_\tau\right)/D_{\tau_-}$, the loss in the event of default, may be modelled and so determines F^*_τ from D_{τ_-}. The RT model supposes that F^*_τ is a function of the riskless present value to time τ of the face value F.

In Section 4.3.3 we analyzed models of recovery in the context of CB valuation. Most models of convertibles assume that the recovery value is a fraction of either the bond principal F (for example Andersen and Buffum (2003), Davis and Lischka (2002)), or the market price of the CB just prior to default D_{τ_-} (for example Takahashi, Kobayahashi and Nakagawa (2001)), or, in splitting schemes, some proportion of the bond part of the CB (for example Ayache, Forsyth and Vetzal (2003)).[31] Each of these assumptions has some attractions, but neither attempts to model the recovery process, regarding recovery values as exogenously determined and separately estimated.

We endogenise recovery into our model.

In practice default may occur when the firm value is significantly greater than the value of its obligations, a feature allowed in our model. The outcome of default is to put the firm into reorganization during which time it receives protection against the claims of

[31] Ayache, Forsyth and Vetzal (2002) discuss in detail default issues in equity based models.

its creditors. The effect is that even though theoretically the firm may have the capacity to fulfill the claims against it, in practice the values of the claims may be considerably less than their face values.

We operationalise this as follows. We interpret a default event simply as a trigger that puts the firm into reorganization, giving the firm protection against its creditors. Upon default at time τ the bondholders have a claim of value F_τ against the firm where

$$F_\tau = \begin{cases} F + P_T, & \text{if } \tau = T \text{ is at the maturity time } T, \\ F + P_{t_i}, & \left\{ \begin{array}{l} \text{if default is at a coupon payment time, } \tau = t_i, \\ \text{or a redemption date coinciding with a coupon date,} \end{array} \right. \\ F, & \left\{ \begin{array}{l} \text{if default is exogenous,} \\ \text{or at a redemption date not coinciding with a coupon date.} \end{array} \right. \end{cases} \tag{5.271}$$

Alternatively we could assume, for instance, that F_τ contains accrued interest, or that on a redemption date $F_\tau = \max\{F, R_\tau\}$.

We suppose that the protection offered by reorganization grants the firm a grace period of length s after default, such that during this period the bondholders no longer have the right to enforce default, but the firm has the option at any time during this period to choose to default (handing the assets over to the holders and not paying the debt). At a put time $t \geq \tau$, the recovery value of the CB is $F_t^* = \min\{F_\tau, V_t\} = F_\tau - (F_\tau - V_t)_+$, where we suppose that the bond holders' claim does not earn interest. Hence, given no disbursements or refinancing, at a default time τ the value of debt is

$$F_\tau^* = \mathrm{Pv}(F_\tau) - p(V_\tau, F_\tau), \text{ where} \tag{5.272}$$

$$V_\tau = V_{\tau_-}(1 - w_\tau), \tag{5.273}$$

Pv denotes present value and p is the value of a put option.[32] Effectively the bondholders are forced to give a put option to the firm allowing the firm to annul the bondholders' claim of F_τ by transferring the firm to the bondholders.

For simplicity we suppose that default is a unique event. Once default has occurred we suppose the firm value follows a geometric Brownian motion but that the event of default influences the growth of firm's future asset value. If we suppose that both the CB and the firm's equity continue to trade during reorganization, then under the EMM the instantaneous return to the firm's assets is still the short rate r_t. However, it seems reasonable to suppose that the volatility of V_t may change, perhaps increasing, as a consequence of default. We denote the post-default volatility by σ^*. The recovery value $F_\tau^*(V_\tau)$ at default time τ is thus determined by two parameters, s and σ^*, each with a natural interpretation.

In the exposition that follows we employ a simplifying assumption. We suppose the firm may exercise the put only at the time $\tau + s$, so p becomes a European put. Effectively, after a default event, the bondholders are obliged to wait a period s before receiving a payment of $\min\{F_\tau, V_{\tau+s}\}$. This assumption enormously decreases the complexity and cost of finding numerical solutions to (5.277).

A feature of our formulation is that at a default time τ the CB holders never receive the amount of their claim, F_τ. For instance at the maturity time T there will be a range of asset values where the firm will not default but where bondholders will not convert, receiving instead an amount equal to F_T. In our model, if the bondholders do not convert they recover F_T^*, which can be significantly less that F_T. This behaviour could be inappropriate

[32] We could also assume that the claim does earn interest, in which case $F_t^* = \min\{\mathrm{Fv}(F_\tau), V_t\} = \mathrm{Fv}(F_\tau) - (\mathrm{Fv}(F_\tau) - V_t)_+$, where $\mathrm{Fv}(F)$ stands for the future value of F, so that $F_\tau^* = F_\tau - p(V_\tau, \mathrm{Fv}(F_\tau))$.

at moderate levels of the firm's assets. We can overcome this problem by allowing s to depend on $\frac{V_T}{F_T}$ so that $s \sim 0$ when $V_T \gg F_T$. However, this would complicate the model and in any case the effect is likely to be small.

When at a default time τ the asset value $V_\tau \gg F_\tau$ is high we refer to 'technical default', since the CB will be converted.

The Interest Rate Model

As in previous Chapters we assume an extended Vasicek interest rate model, which process under the EMM is given by

$$\mathrm{d}r_t = \alpha\left(\theta(t) - r_t\right)\mathrm{d}t + \sigma_r \mathrm{d}z_t^r, \tag{5.274}$$

where $\theta(t)$ can be chosen so that model spot rates coincide with market spot rates. We set $\mu_r \equiv \mu(t, r) = \alpha\left(\theta(t) - r_t\right)$ for the drift of r and write ρ for the correlation between z_t^r and z_t^V, $\mathrm{d}z_t^r \mathrm{d}z_t^V = \rho \mathrm{d}t$.

In the Vasicek model when $\rho = 0$ there is a simple explicit solution for $p(V, r)$.[33]

The Hazard Rate Process

We do not model the risk-adjusted hazard rate $\bar{\lambda}_t$ with its own specific risk. Instead we suppose that $\bar{\lambda}_t \equiv \bar{\lambda}(V_t, r_t)$ is a deterministic function of V_t and r_t. We assume that $\bar{\lambda}_t$ decreases as both V_t and r_t increase. In principle a credit spread model implicitly determines an intensity function. In Chapter 1 we discuss several functional forms for $\bar{\lambda}_t$.

---

[33] In fact in our numerical work we use this formula even when $\rho \neq 0$. The error introduced is small (over our range of values of ρ) and the numerical burden is considerably reduced.

We allow $\bar{\lambda}_t$ to depend on V_t and r_t. For concreteness we choose the functional form

$$\bar{\lambda}(r_t, V_t) = \lambda_0 \exp\left(-\left(aV_t + br_t\right)\right), \qquad a, b \geq 0. \tag{5.275}$$

The coefficients λ_0, a and b control the background default rate and the sensitivity of $\bar{\lambda}_t$ to V_t and r_t. Default risk decreases as V_t increases. As $r_t \to -\infty$, $\bar{\lambda}_t \to \infty$ so that default becomes inevitable.

Note that (5.275) does not require $\bar{\lambda}_t$ to go to infinity when V_t goes to zero. However a consequence of our formulation is that $D_t < V_t$ for all t, so that D_t goes to zero as V_t goes to zero without any constraint on $\bar{\lambda}_t$.

5.1.2 A PDI for a Convertible Bond

We need to specify both the PDI, its boundary conditions and inequality constraints.

The PDI

By Itō's lemma (Protter (1995)) the process followed by D_t is

$$\begin{aligned} \mathrm{d}D_t \;=\; & \left(\frac{\partial D}{\partial t} + \left(r_t + \bar{\lambda}_t w_t\right) V_t \frac{\partial D}{\partial V} + \frac{1}{2}\sigma_V^2 V_t^2 \frac{\partial^2 D}{\partial V^2}\right. \\ & \left. + \mu_r \frac{\partial D}{\partial r} + \frac{1}{2}\sigma_r^2 \frac{\partial^2 D}{\partial r^2} + \rho\sigma_r\sigma_V V_t \frac{\partial^2 D}{\partial V \partial r}\right) \mathrm{d}t \\ & + \sigma_V V_t \frac{\partial D}{\partial V} \mathrm{d}z_t^V + \sigma_r \frac{\partial D}{\partial r} \mathrm{d}z_t^r + \Delta D_t\left(V_{t\_}\right), \end{aligned} \tag{5.276}$$

where $\Delta D_t\left(V_{t\_}\right) = D_t^*\left(V_t\right) - D_{t\_}\left(V_{t\_}\right)$ is the change in the value of the convertible bond if a jump, hence a default, occurs at time t.

Under the EMM the relative price D_t/B_t is a martingale. Imposing this condition we find,

$$\begin{aligned} r_t D_t &= \frac{\partial D}{\partial t} + \left(r_t + \bar{\lambda}_t w_t\right) V_t \frac{\partial D}{\partial V} + \frac{1}{2}\sigma_V^2 V_t^2 \frac{\partial^2 D}{\partial V^2} + \mu_r \frac{\partial D}{\partial r} + \frac{1}{2}\sigma_r^2 \frac{\partial^2 D}{\partial r^2} + \rho\sigma_r\sigma_V V_t \frac{\partial^2 D}{\partial V \partial r} \\ &\quad + \bar{\lambda}_t \mathbb{E}_{t_-}\left[D_t^*\left(V_t\right) - D_{t_-}\left(V_{t_-}\right)\right] \end{aligned} \tag{5.277}$$

Since we assume deterministic loss and recovery conditions this becomes

$$\begin{aligned} \left(r_t + \bar{\lambda}_t\right) D_t &= \bar{\lambda}_t D_t^* + \frac{\partial D}{\partial t} + \left(r_t + \bar{\lambda}_t w_t\right) V_t \frac{\partial D}{\partial V} + \frac{1}{2}\sigma_V^2 V_t^2 \frac{\partial^2 D}{\partial V^2} \\ &\quad + \mu_r \frac{\partial D}{\partial r} + \frac{1}{2}\sigma_r^2 \frac{\partial^2 D}{\partial r^2} + \rho\sigma_r\sigma_V V_t \frac{\partial^2 D}{\partial V \partial r} \end{aligned} \tag{5.278}$$

where in our formulation $D_t^* = \max\left\{F_t^*, \kappa_t V_t\right\}$ and $F_t^* = \mathrm{Pv}\left(F_t\right) - p$ for a put $p \equiv p\left(V_t, F_t\right)$ where $V_t = V_{t-}\left(1 - w_t\right)$.

If V_t is the sole state variable this becomes

$$\left(r_t + \bar{\lambda}_t\right) D_t = \bar{\lambda}_t D_t^* + \frac{\partial D}{\partial t} + \left(r_t + \bar{\lambda}_t w_t\right) V_t \frac{\partial D}{\partial V} + \frac{1}{2}\sigma_V^2 V_t^2 \frac{\partial^2 D}{\partial V^2}, \tag{5.279}$$

which is the form of equation (43) in Ayache, Forsyth and Vetzal (2003).

Inequality Constraints and Auxiliary Conditions for the PDI

We need to specify the final payoff to the convertible bond at time T and payoffs at intermediate times $0 \leq t < T$. We also specify inequality constraints and other conditions on the bond's value.

At the Final Exercise Time T

At times when a cash payment has to be made to the bond the firm has the option to default. At the final time T the firm will default if $V_{T_-} < F_T$. If $V_{T_-} \geq F_T$ we suppose the firm acts to maximize the value of equity by minimizing the value of the CB. Since the bond may convert upon default, we have

$$D_T = D_T^* = \max\{F_T^*,\ \kappa_T V_T\}. \tag{5.280}$$

There is a critical asset value $V_T^* > F_T$ such that $F_T^* = \mathrm{Pv}(F_T) - p(V_T, F_T) = \kappa_T V_T^*$. The CB holders will convert if $V_T > V_T^*$, whether or not the firm elects to default.

At Redemption and Call dates

Consider a redemption date t at which the CB is not callable and which no coupon is paid. Redemption is at the option of the bondholders. If the bond is redeemed the firm has the option to default. If the firm defaults the CB holders have the option to convert. Hence

$$\max\{\min\{F_t^*,\ R_t\},\ \kappa_t V_t\} \leq D_t. \tag{5.281}$$

At a call date t there is an upper bound on the value of the bond. If it is called at time t the CB holders have the option to convert so

$$\begin{array}{ll} \kappa_t V_t \leq D_t \leq \min\{V_t, C_t\}, & V_t < C_t/\kappa_t, \\ D_t = \kappa_t V_t, & V_t \geq C_t/\kappa_t. \end{array} \tag{5.282}$$

We can combine (5.281) and (5.282) into a single expression

$$\max\{\min\{F_t^*,\ R_t\},\ \kappa_t V_t\} \leq D_t \leq \max\{\min\{V_t, C_t\},\ \kappa_t V_t\}, \tag{5.283}$$

where R_t is set to zero on a non-redemption date, C_t is set to $+\infty$ on a non-call date, and κ_t is set to zero on a non-conversion date.

At a Coupon Date

Suppose a coupon of size P_t is due to be paid at time t and that no exercise conditions are invoked so that $D_t^* = F_t^*$. We suppose that the firm acts to maximize the value of equity.

The firm has the choice of paying the coupon or defaulting so the equity value is

$$S_t = \begin{cases} \left(V_{t\_} - P_t\right)_+ - D_t, & \text{if the coupon is paid,} \\ V_{t\_} - F_t^* & \text{if the firm defaults,} \end{cases} \tag{5.284}$$

so $S_t = \max\left\{\left(V_{t\_} - P_t\right)_+ - D_t,\ V_{t\_} - F_t^*\right\} > 0.$

There is a critical value $V_{t\_}^* > P_t$ with $D_{t\_}^*\left(V_{t\_}^*\right) = P_t + D_t\left(V_{t\_}^*\right)$, such that the firm defaults if $V_{t\_} < V_{t\_}^*$ and pays the coupon otherwise. When $V_{t\_} \geq V_{t\_}^*$ we have $D_{t\_}\left(V_{t\_}^*\right) \geq P_t$. Then, if there are no exercise conditions,

$$0 < D_t\left(V_t\right) = \begin{cases} D_{t\_}\left(V_{t\_}\right) - P_t, & V_{t\_} \geq V_{t\_}^*, \\ F_t - p\left(V_{t\_}\right), & V_{t\_} < V_{t\_}^*. \end{cases} \tag{5.285}$$

Now suppose that exercise features are present. If we assume for simplicity that $R_t = R_{t-}$, $C_t = C_{t-}$ and $\kappa_t = \kappa_{t-}$, and that the CB specifies that $F_t = F_{t\_}$, then the firm may choose to call just before the coupon is paid, but the CB holders will not elect to redeem or convert until after the coupon is paid. Then if $V_{t\_} \geq V_{t\_}^*$ the firm does not default and

$$\max\left\{\min\left\{F_t^*,\ R_t\right\},\ \kappa_t V_t\right\} \leq D_t\left(V_t\right) \leq \max\left\{\min\left\{V_t, C_t\right\},\ \kappa_t V_t\right\}. \tag{5.286}$$

Otherwise, if $V_{t\_} < V_{t\_}^*$ the firm defaults and $D_t\left(V_t\right) = D_{t\_}^* = \max\left\{F_{t\_}^*,\ \kappa_{t\_} V_{t\_}\right\}$.

5.2 The Solution Method

As we mentioned in Chapter 1, several numerical methods have been used in the literature to obtain CB values.

We use the numerical method describe in Chapter 2 to solve the PDI (5.278). The valuation problem introduce in previous Section fits into the general framework of Chapter

2 for the choices:

$$x_1 = r_t, \tag{5.287}$$

$$x_2 = V_t, \tag{5.288}$$

$$A_{11} = \frac{1}{2}\sigma_r^2,\ A_{12} = A_{21} = \frac{1}{2}\rho\sigma_r\sigma_V V_t,\ A_{22} = \frac{1}{2}\sigma_V^2 V_t^2, \tag{5.289}$$

$$B_1 = \mu_r,\ B_2 = \left(r_t + \bar{\lambda}_t w_t\right) V_t, \tag{5.290}$$

$$A_0 = -\left(r_t + \bar{\lambda}_t\right),\ f = \lambda_t D_t^*, \tag{5.291}$$

Early exercise features are modelled by the functions R_1 and R_2 and at a call or redemption date, for instance,

$$R_1\left(r_t, V_t, t\right) = \max\left\{\min\left\{F_t^*,\ R_t\right\},\ \kappa_t V_t\right\}, \tag{5.292}$$

$$R_2\left(r_t, V_t, t\right) = \max\left\{\min\left\{V_t,\ C_t\right\},\ \kappa_t V_t\right\}. \tag{5.293}$$

The final condition is determined by the payoff function of the CB (5.280).

In the next Section we specify the boundary conditions required by the numerical method.

5.2.1 Boundary Conditions

For numerical purposes we need to solve the PDI on a finite domain $\Omega = \Omega^r \times \Omega^V$ where $\Omega^r = [r_{\min}, r_{\max}]$, $\Omega^V = [0, V_{\max}]$ and $V_{\max} \geq R_t, C_t$ at all times when these are defined. At the boundaries of the solution domain Ω we need to supply boundary conditions to our solution method. We suppose that asymptotic approximations can be applied at $r_{\min}$, $r_{\max}$ and $V_{\max}$.

Four boundary conditions are required. The convertible bond literature is often not explicit about the boundary conditions used. In our framework the asset boundary conditions, at $V_t = 0, V_{\max}$ are straightforward, as are the conditions at $r_{\max}$. There are problems if one tries to supply a boundary condition at $r_{\min} = 0$. Instead we choose $r_{\min} < 0$, a natural assumption in the Vasicek model where interest rates are not constrained to be positive.[34]

We suppose the final condition and inequality constraints are given by (5.280), (5.283) and (5.286) and explore asymptotic conditions. Since the PDI is solved backwards in time we re-formulate (5.285) and (5.286). On a coupon date t we first compute a value D_{t_+}, notionally the CB value immediately after the coupon has been paid, respecting exercise conditions at time t_+, post-coupon. We then find D_{t_-}, the CB value immediately before the coupon is paid, and impose exercise conditions at time t_-, pre-coupon. Then we continue iterating backwards.

Over the coupon payment time we have

$$V_{t_-} = V_{t_+} + P_t, \tag{5.294}$$

$$\widetilde{D}_{t_-}\left(V_{t_-}\right) = \begin{cases} D_{t_+}\left(V_{t_-} - P_t\right) + P_t, & V_{t_-} \geq V^*_{t_-}, \\ D^*_{t_-}, & V_{t_-} < V^*_{t_-}. \end{cases} \tag{5.295}$$

then the exercise condition is

$$\max\left\{\min\left\{\widetilde{D}_{t_-}\left(V_{t_-}\right), R_t\right\}, \kappa_t V_{t_-}\right\} \leq D_{t_-} \leq \max\left\{\min\left\{V_{t_-}, C_t\right\}, \kappa_t V_{t_-}\right\}. \tag{5.296}$$

---

[34] We find that it is possible to choose $r_{\min} < 0$ so that asymptotic conditions apply and the computation of λ_t does not cause overflow.

A Simple CB

First, we consider a convertible bond with no coupons, convertible only at time T, nowhere callable or redeemable, where default occurs only at time T and with no proportional loss. It is modelled by setting the bond payoff to be $D_T = \max\{\min\{V_T, F\}, \kappa_T V_T\}$. Notice that this is the payoff of the general CB defined in (5.280) if the lenght of the reorganization period s is set to zero. We call this a simple CB. In this case,

$$\begin{aligned} D_T &= \max\{\min\{V_T, F\}, \kappa_T V_T\} \\ &= \max\{\kappa_T V_T - \min\{V_T, F\}, 0\} + \min\{V_T, F\} \\ &= \max\{\kappa_T V_T - F, 0\} + \min\{V_T, F\}, \end{aligned} \tag{5.297}$$

where the last equality follows from the fact that $\kappa_T < 1$. Hence, the convertible value decomposes into the value of a defaultable straight bond and a call on the firm's assets, with explicit solution

$$D_t = V_t - c_t(V_t, F) + \kappa_T c_t(V_t, F/\kappa_T) \tag{5.298}$$

where

$$c_t(V_t, F) = V_t N(d_1) - \text{Pv}(F) N(d_2), \tag{5.299}$$

$$d_1 = \frac{1}{\sigma\sqrt{T-t}} \ln\left(\frac{V_t}{\text{Pv}(F)}\right) + \frac{1}{2}\sigma\sqrt{T-t}, \tag{5.300}$$

$$d_2 = d_1 - \sigma\sqrt{T-t}, \tag{5.301}$$

$N(\cdot)$ is the cumulative normal distribution function and $\text{Pv}(F)$ is the present value computed in the Vasicek term structure model.

Boundary:		$V \to 0$	$V \to \infty$	$r \to -\infty$	$r \to \infty$
Dirichlet:	D_t	0	$\kappa_T V_t$	V_T	$\kappa_T V_t$
Neumann:	$\frac{\partial D_t}{\partial V}$	1	κ_T	1	κ_T
	$\frac{\partial D_t}{\partial r}$	0	0	0	0

Table 5.47. Boundary conditions, simple CB

Asymptotic Dirichlet and Neumann boundary conditions can be computed and are given in Table 5.47. Note that limits as $(V, r) \to (+\infty, -\infty)$ depend upon the direction of the limit. We found it is best to use a Neumann boundary conditions at the $r \to -\infty$ boundary, and Dirichlet boundary conditions at the other three boundaries.

Boundary Conditions for the General CB

No we consider the general CB.

High asset value

As $V \to \infty$ the CB effectively becomes an equity instrument and it will either be converted, or converted when it is called, at an optimal time t^* that does not depend upon V or the coupon stream. Then, ignoring the possibility of default, the bond value is

$$D_t = \kappa_{t^*} V_t + \widehat{P}_t^{t^*} \tag{5.302}$$

where $\widehat{P}_t^{t^*} = \sum_{t \leq t_i \leq t^*} \mathrm{Pv}\left(P_{t_i}\right)$ is the value at time t of the future coupons received up to time t^*.

In general (5.302) may be hard to compute, but when conversion terms are constant t^* is the first available conversion date. If continuous conversion is possible $t^* = t$ and $D_t = \kappa_t V_t$.

Low asset value

When $V = 0$ we have $D_t = 0$.

High interest rate

As $r \to \infty$ for $t < T$ the present value of the principle F goes to zero. Cash becomes irrelevant and the time value of money is expressed in returns to the asset process. The payoff to the CB at time T is effectively $\kappa_T V_T$, and default is irrelevant. Payoffs, if received in cash, will be used to immediately buy the asset.

As before, suppose the bond is exercised by one party or the other at an optimal time t^*. It may be optimal to redeem if R_t/V_t is large enough, or to call when $V_t < C_t/\kappa_t$ if future values of κ_t are large enough. At time t^* we have $D_{t^*} = \kappa^*_{t^*} V_{t^*}$ where $\kappa^*_{t^*} = \max\{C_{t^*}/V_{t^*},\ \kappa_{t^*}\}$ if the bond was called and $\kappa^*_{t^*} = \max\{\kappa_{t^*},\ \min\{D^*_{t^*},\ R_{t^*}\}/V_{t^*}\}$ if the bond is redeemed. Since cashflows are immediately used to buy equity, $\kappa^*_{t^*}$ is the effective conversion ratio at time t^*. For high r,

$$D_t = \kappa^*_{t^*} V_t + \widehat{V}_t^{t^*} \tag{5.303}$$

where $\widehat{V}_{t_1}^{t_2} = V_t \sum_{t_1 \leq s \leq t_2} P_s/V_s$ is the value at time t_1 of asset rebased future coupons received up to time t_2.

As before, this simplifies if conversion terms, *et cetera*, are constant and continuous, and we may set $D_t = \kappa^*_t V_t$.

Low interest rate

When $r \to -\infty$ the asset value becomes irrelevant and cash values dominate. Default occurs at the first cashflow date, if not sooner. CB holders will wait for a default event at some time τ and then take over the firm, so $D_t = \mathbb{E}_t\left[\mathrm{Pv}\left(V_\tau\right)\right] = V_t$.

We see that, with possible small modification, Table 5.47 gives the correct boundary conditions for a general convertible bond.

Exercise parameter values			
	Convertibility	Callability	Redeemability
Base:	0.2	22	18
High:	0.22	24	19
Low:	0.18	21	16

Table 5.48. Exercise parameter values

5.3 Numerical Results

In this Section we first benchmark the model, investigating the convergence properties of the numerical method. We then explore the effect upon CB values of altering parameter values within the model, looking particularly at the exercise conditions, asset and interest rate values and parameters, the recovery parameters s and σ^*, and the default parameters λ_0, a, b and w.

Each parameter has a base case value, and a high and a low value. These are given in Tables 5.48, 5.49 and 5.50. For the base case we suppose that the CB has $T = 5$ years to maturity with face value $F = 20$. The CB may be converted at any time with indirect conversion ratio $\kappa_t \equiv \kappa = 0.2$. The CB pays a coupon of 0.6 every half year, an annual coupon yield of 6%. It is callable and redeemable at any time with the call and redemption prices determined from (5.268) and (5.269) with $C = 22$ and $R = 18$. The initial asset value is $V_0 = 100$ and initial interest rate is $r_0 = 0.06$. For the default intensity function we set $\lambda_0 = 0.15$, $a = 0.015$, $b = 1.5$, giving $\lambda_t \in [0, 0.7]$ over the domain, with $\lambda_t = 0.03$ in the base case. For middling values of V_t and r_t, λ_t has about the same sensitivity to changes in each. Other parameter values are given in the Tables.

We note that with this specification the convertible bond is at the money and that in the base case the likelihood of exogenous default is relatively low.

Process parameters							
	For r				For V		Corr.
Parameter:	r_0	α	θ	σ_r	V_0	σ_V	ρ
Base:	0.06	0.2	0.06	0.02	100	0.25	0.1
High:	0.07	0.21	0.07	0.025	105	0.30	0.15
Low:	0.05	0.19	0.05	0.015	95	0.20	0.05

Table 5.49. Process parameter values

Default, recovery and loss parameters						
	Default			Recovery		Loss
Parameter:	λ_0	a	b	s	σ^*	w
Base:	0.15	0.015	1.5	1	0.35	0.4
High:	0.2	0.03	3	5	0.45	0.6
Low:	0.1	0.003	0.3	0.25	0.25	0.2

Table 5.50. Default and recovery parameter values

5.3.1 Benchmarking

We benchmark to a simple CB whose value is given by (5.298), investigating convergence.

Domain bounds are set to be $\Omega^r = [-1, 1]$ and $\Omega^V = [0, 800]$. Ω^V corresponds to roughly a 99.9% confidence interval on V_T. We give L^2 errors over both the entire domain Ω and also over a narrower region of interest $\widehat{\Omega} = \widehat{\Omega}^r \times \widehat{\Omega}^V$, where $\widehat{\Omega}^r = [0, 0.15]$ and $\widehat{\Omega}^V = [25, 400]$. $\widehat{\Omega}^V$ is roughly a 99% confidence interval on V_T. $\widehat{\Omega}$ reflects a range of values of r and V likely to be observed in practice and so the error on $\widehat{\Omega}$ is likely to be more representative.

For the numerical method we use four mesh specifications of increasing resolution. Mesh 1 is the coarsest with just 20 space steps in the interest rate dimension, 40 in the asset dimension, and 50 time steps up to time $T = 5$. Each successive mesh doubles both the number of space steps in each dimension and the number of time steps so that the finest

Errors and Convergence, Analytical Boundary Conditions					
Mesh	Error TD	Factor	Error RI	Factor	Time
1	$6.1E{-}02$	-	$3.1E{-}03$	-	3
2	$3.8E{-}02$	$1.6E{+}00$	$1.4E{-}03$	$2.2E{+}00$	17
3	$2.2E{-}02$	$1.7E{+}00$	$7.4E{-}04$	$1.8E{+}00$	154
4	$1.2E{-}02$	$1.8E{+}00$	$4.5E{-}04$	$1.6E{+}00$	1648

Errors and Convergence, Asymptotic Boundary Conditions					
Mesh	Error TD	Factor	Error RI	Factor	Time
1	$6.7E{-}02$	-	$3.1E{-}03$	-	2
2	$4.7E{-}02$	$1.4E{+}00$	$1.4E{-}03$	$2.2E{+}00$	15
3	$3.4E{-}02$	$1.4E{+}00$	$7.4E{-}04$	$1.8E{+}00$	146
4	$2.8E{-}02$	$1.2E{+}00$	$4.5E{-}04$	$1.6E{+}00$	1648

Table 5.51. Error and convergence

mesh, mesh 4, has 160 interest rate steps, 320 asset steps, and 400 times steps up to five years.

The results are presented in Table 5.51. Two sets of results are shown. The top panel uses analytical values on the boundary, the bottom panel uses asymptotic approximations, as given in Table 5.47. In each case three of the boundaries are Dirichlet and the fourth, at the lower boundary for r, is Neumann. 'Error TD' is the error on the entire domain Ω; 'Error RI' is the error on the region of interest, $\widehat{\Omega}$. 'Factor' is progressive error reduction factor in moving to a finer mesh level from the preceding mesh level. Times are in seconds.[35] We see that using asymptotic boundary conditions the convergence rate is not as fast as the theoretical rate of 2,[36] although on the region of interest the convergence rate is faster than on the whole domain. Errors are significantly less, by a factor of 100, on the region of interest compared to the total domain. Errors are greater on the total domain with asymptotic boundary conditions, but they are of the same order of magnitude. On the region of interest

[35] The implementation was in Fortran 77 run on a 2.4 Mhz Pentium IV PC.

[36] A convergence rate of order 2 has been proven for a related PDI, but not strictly for the PDI of this Chapter.

Recovery parameters		Initial values of (V_t, r_t)				
		$(100, 0.06)$	$(80, 0.05)$	$(120, 0.05)$	$(80, 0.07)$	$(120, 0.07)$
(s, σ^*)	$(1, 0.35)$	**0.941**	0.942	0.945	0.937	0.939
	$(0.25, 0.25)$	0.985	0.985	0.986	0.984	0.984
	$(0.25, 0.45)$	0.985	0.985	0.986	0.983	0.984
	$(5, 0.25)$	0.746	0.752	0.755	0.736	0.738
	$(5, 0.45)$	0.721	0.715	0.736	0.704	0.722

Table 5.52. The Implied Recovery Rate: Simple CB

the errors for analytical and asymptotic boundary conditions are the same to two significant figures, supporting our use of asymptotic boundary conditions in the sequel.

Subsequent Tables are computed using mesh 4 and asymptotic boundary conditions. All specifications lie within the region of interest so, in line with the errors reported in Table 5.51, CB values are reported to 3 decimal places. With early exercise possible, a typical execution time is around 6700 seconds, relatively independent of the CB specification.

5.3.2 The Recovery Specification

We investigate the consequences of our recovery specification, interpreting it by computing the implied recovery ratio $\delta(V_t, r_t)$ defined as

$$\delta(V_t, r_t) = \mathbb{E}_t \left[\frac{F_\tau^*}{F} \mid V_t, r_t \right] \tag{5.304}$$

for a default time τ. δ is the proportion of face value the bondholders can expect to recover in the event of default if they do not convert.[37] We compute δ for a simple CB. Table 5.52 shows δ for a variety of initial conditions and recovery specifications. The entry in bold is the base case.

[37] When the CB is in the money default is technical and the CB will be converted.

Recovery parameters		Initial values of (V_t, r_t)		
		$(25, 0.06)$	$(30, 0.06)$	$(35, 0.06)$
(s, σ^*)	$(1, 0.35)$	0.854	0.885	0.904
	$(0.25, 0.25)$	0.913	0.942	0.959
	$(0.25, 0.45)$	0.904	0.935	0.953
	$(5, 0.25)$	0.675	0.697	0.711
	$(5, 0.45)$	0.577	0.606	0.628

Table 5.53. The Implied Recovery Rate, Riskier CB

Our example is at the money but the default put is out of the money so the value F_T^* is approximately equal to the present value at time T of F_T paid at time $T + s$.

We see that the most important factor for expected recovery is the length of the reorganization period, followed by the interest rate and then the initial asset value. Changing the volatility parameter has little effect when the reorganization period is short, but has an effect comparable in size to the asset value change when s is longer.

Table 5.53 gives recovery rates for riskier CBs issued at a much lower asset value. The implied recovery rates are significantly smaller. With low asset values the put value is not negligible; default is no longer technical and CB holders will not find it optimal to convert, instead obtaining only the expected recovery rates given in the Table. The effect of σ^* is now significant.

5.3.3 Exercise Conditions

We investigate the effect of the presence or absence of the various exercise conditions. We consider a riskless coupon bond with default and various combinations of exercise conditions added in, ending with the full specification of the base case CB. We also give an approximation to the value of $\partial D_t/\partial r$ found by central difference from CB values computed at different initial values of the interest rate.

Tables 5.54 and 5.55 show the results. 'Def' is defaultable (with recovery), 'Con' is convertible, 'Red' is redeemable and 'Call' is callable. Δ and Γ are the CB delta and gamma respectively.[38] The riskless bond values are Vasicek values computed analytically and shown for comparison. The base case value of the CB is 21.085, shown in bold. Table 5.54 shows the standard case. For comparison, Table 5.55 shows the effect of reducing the conversion ratio from 0.2 to 0.15. We first discuss Table 5.54.

With our specification and model parameters, the presence of default reduces the value of the corresponding riskless bond by a little over 5%. The bond has a high credit risk stemming from a relatively high endogenous default rate.

Adding the conversion feature increases the value of the CB by around 13% in the base case. The effect is similar at all levels of the initial interest rate. It is greater at higher values of V_0 as the CB becomes more in the money, and at higher values of r. The introduction of the call feature reduces the value of the CB. The reduction is greater at higher values of V_0 as the CB is more likely to be called. Adding the redemption feature has very little effect on the CB value, with our specification.

Asset deltas are not insignificant. They vary only a little as the initial interest rate changes. Introducing conversion to a straight defaultable bond increases the delta by a factor of 100. The call feature does not greatly affect the asset delta, reducing it slightly, and adding a redemption feature affects the asset delta very little.

---

[38] These are reported 'plain', without division by the conversion ratio.

$\partial D_t/\partial r$, the CB's rho, indicates the sensitivity of the CB to changes in the initial value of the interest rate. The conversion feature reduces the absolute size of the CB's rho by a fifth to a tenth, depending on the initial asset value. The redemption feature has little effect, but the call feature reduces rho at higher asset values.

Allowing the riskless bond to become defaultable reduces rho by roughly 20% and adding additional optionality reduces it further, by about 90% in the base case. For this CB, additional optionality effectively decreases the interest rate exposure of the CB and significantly increases its asset value exposure.

Table 5.55 shows how the situation changes if the conversion factor is significantly reduced, to $\kappa = 0.15$. Now the CB is out of the money.

Adding convertibility increases the value of the defaultable bond by about 5%, but adding the call and redemption features has little affect on the CB value. Asset deltas are affected, but by much less that in the $\kappa = 0.2$ case. Rhos are reduced, but by much less than in the $\kappa = 0.2$ case.

The CB of Table 5.55 is relatively 'bond-like' whereas that of Table 5.54 is much more 'asset-like'.

Table 5.56 shows the effect of changing exercise conditions. Since the CB is at the money, changing the convertibility condition has a significant effect. Increasing the redemption level has little effect but changing the call level has a large effect for this at the money CB.

Exercise Conditions				V_0	r			$\partial D_t/\partial r$
Def	Con	Red	Call		0.05	0.06	0.07	
Riskless Bond				100	20.569	19.992	19.432	−56.9
				95	19.280	18.835	18.366	-45.7
√	-	-	-	100	19.285	18.840	18.370	-45.8
(Defaultable bond)				105	19.290	18.845	18.374	-45.8
				Δ	0.001	0.001	0.001	
				Γ	-4.0E-05	-3.9E-05	-3.6E-05	
				95	20.902	20.823	20.738	-8.2
√	√	-	-	100	21.399	21.338	21.274	-6.2
				105	22.002	21.957	21.909	-4.7
				Δ	0.110	0.113	0.117	
				Γ	0.004	0.004	0.004	
				95	20.907	20.830	20.748	-7.9
√	√	√	-	100	21.401	21.342	21.280	-6.1
				105	22.003	21.959	21.912	-4.5
				Δ	0.110	0.113	0.116	
				Γ	0.004	0.004	0.004	
				95	20.749	20.672	20.588	-8.0
√	√	-	√	100	21.134	21.081	21.024	-5.5
				105	21.579	21.551	21.520	-3.0
				Δ	0.083	0.088	0.093	
				Γ	0.002	0.002	0.002	
				95	20.753	20.679	20.598	-7.7
√	√	√	√	100	21.136	**21.085**	21.029	-5.3
(Base case)				105	21.580	21.553	21.522	-2.9
				Δ	0.083	0.087	0.092	
				Γ	0.002	0.002	0.002	

Table 5.54. Effect of Exercise Features, base case, $\kappa = 0.2$

Exercise Conditions				V_0	r			$\partial D_t/\partial r$
Def	Con	Red	Call		0.05	0.06	0.07	
Riskless Bond				100	20.569	19.992	19.432	-56.9
√ (Defaultable bond)	-	-	-	95	19.280	18.835	18.366	-45.7
				100	19.285	18.840	18.370	-45.8
				105	19.290	18.845	18.374	-45.8
				Δ	0.001	0.001	0.001	
				Γ	-4.0E-05	-3.9E-05	-3.6E-05	
√	√	-	-	95	19.844	19.632	19.398	-22.3
				100	19.920	19.733	19.526	-19.7
				105	20.017	19.856	19.679	-16.9
				Δ	0.017	0.022	0.028	
				Γ	7.9E-04	9.0E-04	1.0E-03	
√	√	√	-	95	19.869	19.670	19.449	-21.0
				100	19.939	19.761	19.565	-18.7
				105	20.030	19.877	19.708	-16.1
				Δ	0.016	0.021	0.026	
				Γ	8.6E-04	9.7E-04	1.1E-03	
√	√	-	√	95	19.839	19.625	19.388	-22.5
				100	19.912	19.723	19.513	-20.0
				105	20.002	19.840	19.660	-17.1
				Δ	0.016	0.021	0.027	
				Γ	6.6E-04	7.7E-04	8.8E-04	
√ (Base case)	√	√	√	95	19.864	19.663	19 440	-21.2
				100	19.931	19.752	19.553	-18.9
				105	20.015	19.861	19.690	-16.3
				Δ	0.015	0.020	0.025	
				Γ	7.2E-04	8.4E-04	9.5E-04	

Table 5.55. Effect of Exercise Features, low conversion rate, $\kappa = 0.15$

Convertibility		Red.	Call	
			Low	High
Low	20.347	Low	20.722	21.307
High	22.070	High	20.725	21.311

Table 5.56. Sensitivities to changes in exercise conditions

	r-parameters				V-parameters		Corr.
Parameter:	r	α	θ	σ_r	V	σ_V	ρ
High:	21.029	21.083	21.072	21.083	21.553	21.193	21.088
Low:	21.136	21.086	21.097	21.086	20.679	20.948	21.081
Delta:	-5.34	-0.16	-1.27	-0.23	0.09	2.45	0.07
Gamma:	-46.4	0.10	-3.00	-15.20	0.00	-11.52	-0.06

Table 5.57. Sensitivities to changes in parameter values

5.3.4 Parameter Deltas and Gammas

We investigate the sensitivity of the base case CB to changes in parameter values. We value the CB at the higher value and lower value of each parameter. The delta and gamma are then computed by central difference. Results are given in Table 5.57. r is the initial value of the stochastic Vasicek interest rate. Later, Table 5.61 considers the effect of changes to r where r is a constant interest rate.

Deltas are very small. σ_V has a greater delta than σ_r, and when σ_V is scaled by V (to make it comparable to an absolute volatility) the effect upon Δ is even greater. Increasing the correlation ρ slightly increases the bond value. θ, the level to which r_t reverts, has a slightly larger delta since it reflects the longer term value of r_t.

5.3.5 The Default Specification

We explore the consequences of changing the default specifications. Tables 5.58, 5.59 and 5.60 summarize the results.

Care must be taken in interpreting these Tables. They assume that the initial asset value V_0 and the default and loss parameters $(s, \sigma^*, \lambda_0, a, b, w)$ may be determined independently, so that, for instance, the default rate might increase while V_0 remains fixed. In practice an increase to λ might be expected to cause V_0 to fall, so that the CB value would

Recovery Parameter		Hazard rate, λ_t		
		Base	High	Low
(s, σ^*)	$(1, 0.35)$	21.085	21.112	21.080
	$(0.25, 0.25)$	21.138	21.186	21.126
	$(0.25, 0.45)$	21.137	21.186	21.126
	$(5, 0.25)$	20.473	20.475	20.476
	$(5, 0.45)$	20.324	20.317	20.330

Table 5.58. Sensitivities to changes in default parameters

be computed for a reduced value of V_0. This feature is not modelled by the specification (5.267), nor reflected in the Tables. This suggests that a full defaultable bond model would need to endogenise the effect of default on V_t, perhaps by allowing V_0 to be determined from future cashflow streams.

Note that these considerations do not affect the practical implementation of a model. Calibrating to market data fits to mutually determined values of V_0 and the default and loss parameters, so CB values are correctly determined.

Table 5.58 shows sensitivities to the recovery parameters and base, high and low values of λ_t, conditional on V_0 remaining fixed. Table 5.59 shows the effect on CB values of λ_0, a and b individually taking high or low values, conditional on fixed V_0, and Table 5.60 shows the affect upon the CB value of varying the loss rate, w, for different initial asset values V_0.

Bearing in mind the discussion above, because the rate of exogenous default in the base case is quite low, varying the hazard rate parameters and w seems to have little effect on CB values. Note that in Table 5.60 when $V_0 = 120$ it is optimal to immediately convert the bond.

Parameter	Default Parameters		
	λ_0	a	b
High	21.087	21.080	21.084
Low	21.083	21.101	21.085

Table 5.59. Sensitivities to changes in default parameters

Recovery Parameters		Loss rate, w		
		0.2	0.4	0.6
V_0	80	19.864	19.879	19.864
	100	21.071	21.085	21.094
	120	24.000	24.000	24.000

Table 5.60. Effect of Different Loss Rates

5.3.6 The Effect of a Stochastic Interest Rate

We have seen the effect upon the bond value of changes in the parameters of the interest rate process. We can also test to find the extend of the affect upon the bond price of a stochastic interest rate. By setting $\sigma_r = 0$ and $r = \theta$ we effectively make r non-stochastic. We investigate the presence of a stochastic interest rate in more detail. Table 5.61 gives the results, looking at several sets of initial conditions.[39] Since the coupon rate is close to current and future interest rate levels the CB price remains relatively stable as T increases.

[39] When T increases or decreases, the time step Δt is held constant and the number of time steps is varied.

Initial Conditions		Stochastic r			Constant r		
		σ_r			r		
		0.015	0.02	0.025	0.05	0.06	0.07
T	1	20.983	20.986	20.989	21.075	20.988	20.901
	5	21.086	21.085	21.083	21.183	21.119	21.046
	10	21.087	21.086	21.084	21.184	21.123	21.049
V_0	95	20.681	20.679	20.676	20.826	20.729	20.619
	100	21.086	21.085	21.083	21.183	21.119	21.046
	105	21.553	21.553	21.552	21.604	21.571	21.531
κ	0.18	20.351	20.347	20.341	20.544	20.411	20.260
	0.2	21.086	21.085	21.083	21.183	21.119	21.046
	0.22	22.070	22.070	22.070	22.074	22.073	22.069

Table 5.61. Effect of a stochastic interest rate

As we have seen elsewhere, the value of the CB is relatively sensitive to the initial asset value V_0, and is even more so to the conversion parameter κ.

Comparing to Table 5.57, we see that making r constant at its initial value has the effect of increasing the value of the CB. Consistent with this, increasing the interest rate volatility when the rate is stochastic decreases the value of the CB, except for short times to maturity.

5.4 Conclusions

In this Chapter we have introduced a two-factor model for defaultable convertible bond pricing where the state variables are the firm asset value and the short interest rate. Default can be exogenous, at the jump time of a counting process, or endogenous at times when the firm must make a cash payment. We endogenise recovery into the model by assuming that upon default the firm enters a reorganization period.

We price convertible bonds by solving numerically a PDI using finite elements to discretize in space and the method of characteristics to discretize in time. We deal with early exercise using a duality method in the variational formulation of the discretized problem.

Care has been taken to specify correctly the boundary conditions in the model, ensuring that these are financially and numerically consistent.

We have investigated the effect of introducing a stochastic interest rate and we have explored the consequences of our default, recovery and loss specification, finding that a wide range of recovery levels are possible, linked to a natural interpretation of the recovery process.

The sensitivity of the CB value to changes in the initial values of the asset and the interest rate have been investigated. We have found that with our specification the CB has a large asset delta and a relatively low sensitivity to the initial interest rate. Adding the conversion feature increases the CB asset delta by a factor of 100.

Despite of the drawbacks of using the asset value as the main state variable, we believe that the modelling framework presented in this Chapter is more realistic than formulations based upon the equity value. For example, we account for the fact that the issuer is more likely to default when a cash-flow has to be paid to the CB holders, whereas in an equity intensity based approach default is equally likely at any day. On the other hand, our endogenised recovery specification potentially allows a greater ability to estimate recovery values from the market.

We conclude that the flexible specification of this model may give it greater potential to explain empirical CB values than existing models in the literature.

Conclusion

Convertible bonds are an increasingly important financial products that enable issuers to obtain relatively cheap finance in exchange for up-side gains. Despite the fact that the CB market is growing rapidly and products gain continuously in complexity, research to accurately and efficiently price those instruments is relatively scarce. Hence we believe our work makes a significant contribution to an important area of finance.

This thesis extends previous literature on CB pricing in two ways: (1) We provide a better understanding of the mathematical models arising in convertible bond valuation and we introduce two new models incorporating interest and credit risk, one equity based and another asset based; (2) We propose sophisticated numerical techniques and we apply them successfully to implement the different models.

The numerical methods developed in the thesis have not been used before in finance but have proven to work well for similar problems in other fields. They offer clear advantages over most of the currently used numerical techniques in terms of its generality to price many different financial products, flexibility to incorporate any exotic product specification and its efficiency, in the sense that they provide a good trade off between speed and accuracy.

Regarding the equity based approach, we present a unified intensity based framework which incorporates most of existing models, as well as new ones that we introduce. This allows us to put models in perspective, as well as implement them and be able to compare the prices they produce. We conclude that care is needed to consistently specify the hazard rate, the loss rate in the stock and the recovery value. We find that in order to do so, it may be necessary to split the convertible into an equity and a bond part and apply to each of them a different credit treatment. We introduce a new splitting procedure and a new recovery model for the CB based on a dual recovery in bond and equity parts. The model has a natural financial interpretation and enough degrees of freedom to generate a great number of default-recovery scenarios.

Asset based models have the disadvantage that they are based on unobserved state variables, but they have the advantage to provide a natural link between debt and equity which is ideal in CB valuation. We propose a new asset based model incorporating interest rate and credit risk. The CB defaults either at the unpredictable jump time of a counting process, or when the firm is required to make a cashflow to the CB holder. Recovery upon default is endogenised into the model by assuming that the firm can invoke temporary protection against its creditors. The main difference with respect to the equity based model is that default can be triggered endogenously at a cash-flow, whereas in the equity based model default is equally likely at any time. Besides, the recovery value, which proved to be the "key" issue in the equity based models, is easily endogenised.

Furthermore, even though our study focus on CB pricing and as a result substantially enlightens our understanding of this area of finance, we believe that our results have broader implications. This is because the framework we propose is very general and could be used to price any American-style contingent claim with equity, interest rate and credit risk. It allows for calibration of interest rates, hazard rates and volatilities. It is also very flexible to incorporate any exotic product specification. Hence, our claim to have introduced a commercially usable pricing framework, which is market consistent and provides a balanced trade-off between speed and accuracy.

Finally, we would like to mention a few areas where we believe future research could provide some interesting extensions to our work.

From the modelling side, it would be interesting to implement a stochastic hazard rate but deterministic interest rate (see Davis and Lischka (2002)). In agreement with previous literature, the effect of stochastic interest rates has shown to be very little in our work. On the other hand, deterministic hazard rates fail to model some real life totally unpredictable defaults. However a more extensive investigation into the pricing of actual market issues should be carried out in order to be able to conclude.

Another important issue is to account for volatility surfaces, given that volatility is a key element in the pricing and hedging of CBs. This leads us to the need to calibrate local implied volatilities, maybe in combination with hazard rates and interest rates. Andersen and Buffum (2003) have done some research in this direction in the context of CBs.

From the implementation side, the next step would be to implement a second order characteristic/finite element method. Some methods have been proposed (Boukir, Maday, Metivet and Razanfindrakoto. (1997), Rui and Tabata (2001), Bermúdez, Nogueiras and Vázquez (2004*b*)). The approach followed by Bermúdez, Nogueiras and Vázquez (2004*b*) could be directly apply to our case. Some other improvements could be done to the current numerical methods. One of them is to use an adaptative scheme to refine more the solution in some areas of the domain where it is needed, for example near the free boundaries (this has been done for example by Pironneau and Hetch (2000)). A second improvement could be to use a multigrid method to speed up the algorithm (Clarke and Parrot (1999)).

Bibliography

Adams, K. J. & D. R. Van Deventer. 1994. "Fitting Yield Curves and Forward Rate Curves with Maximum Smoothness." *Journal of Fixed Income* pp. 52–62.

Alford, J. & Webber N. 2001. Very High Order Lattice Methods for One Factor Models. Working paper University of Warwick.

Altman, E. I., A. Resti & A. Sirone. 2001. Analysing and Explaining Default Recovery Rates. Report ISDA, Stern School of Business, New York University.

Amin, K. & A. Khanna. 1994. "Convergence of American Option Values from Discrete to Continuous-Time Financial Models." *Mathematical Finance* 4:289–304.

Andersen, L. 1999. A Simple Approach to the Pricing of Bermudan Swaptions in the Multi-Factor Libor Market Model. Working paper General Re Financial Products.

Andersen, L. & D. Buffum. 2003. "Calibration and Implementation of Convertible Bond Models." *Journal of Computational Finance* 7(2):1–34.

Arvanitis, A. & J. Gregory. 2001. *Credit: The Complete Guide to Pricing, Hedging and Risk Management.* Risk Books.

Ayache, E., P. A. Forsyth & K. R. Vetzal. 2002. "Next Generation Models for Convertible Bonds with Credit Risk." *Wilmott Magazine* pp. 68–77.

Ayache, E., P. A. Forsyth & K. R. Vetzal. 2003. "Valuation of Convertible Bonds with Credit Risk." *Journal of Derivatives* 11(1):9–29.

Bakshi, G., D. Madan & F. Zhang. 2001. Understanding the Role of Recovery in Default Risk Models: Empirical Comparisons and Implied Recovery Rates. Working paper University of Maryland.

Barles, G., C. H. Daher & P. Souganidis. 1995. "Convergence of Numerical Schemes for Parabolic Equations Arising in Finance Theory." *Mathematical Models and Methods In Applied Science* 5:125–143.

Barles, G. & P. E. Souganidis. 1991. "Convergence of Approximation Schemes for Fully Nonlinear Second Order Equations." *Asymptotyc Analysis* 4(4):271–283.

Barone-Adesi, G., A. Bermúdez & J. Hatgioannides. 2003. "Two-Factor Convertible Bonds Valuation Using the Method of Characteristics / Finite Elements." *Journal of Economic Dynamics and Control* 27(10):1801 – 1831.

Barraquand, J. & T Pudet. 1996. "Pricing of American Path-Dependent Contingent Claims." *Mathematical Finance* 6:15–71.

Benqué, J. P., P. Esposito & G. Labadie. 1983. New Decomposition Finite Element Methods for the Stokes Problem and the Navier-Stokes Equations. In *Numerical Methods in Laminar and Turbulent Flow*. Pineridge pp. 553–563.

Bensoussan, A. & J.L. Lions. 1978. *Applications Des Inéquations Variationneles En Contrôle Stochastique*. Dunod.

Bermúdez, A. & C. Moreno. 1981. "Duality Methods for Solving Variational Inequalities." *Computers Mathematics with Applications* 7:43–58.

Bermúdez, A. & M. R. Nogueiras. 2004. "Numerical Solution of Two-Factor Models for Valuation of Financial Derivatives." *Mathematical Models and methods in Applied Sciences* 14(2):295–327.

Bermúdez, A., M. R. Nogueiras & C. Vázquez. 2004*a*. Modelling and Numerical Solution of Eurasian Option Pricing Problems. Working paper Universidad de Santiago de Compostela, Spain.

Bermúdez, A., M. R. Nogueiras & C. Vázquez. 2004*b*. Numerical Solution of Degenerated Convection-Diffusion-Reaction Problems with Higher Order Characteristics/Finite Elements. Working paper Universidad de Santiago de Compostela, Spain. Submitted.

Bermúdez, A. & N. Webber. 2004. An Asset Based Model of Defaultable Convertible Bonds with Endogenised Recovery. Working paper Cass Business School, London.

Bielecki, T. & Rutkowski. 2002. *Credit Risk: Modeling, Valuation and Hedging*. Springer Finance.

Black, F., E. Derman & W. Toy. 1990. "A One Factor Model of Interest Rates and its Application to Treasury Bond Options." *Financial Analyst Journal* 46:33–39.

Black, F. & J. Cox. 1976. "Valuing Corporate Securities: Some Effects of Bond Indenture Provisions." *Journal of Finace* pp. 351–367.

Black, F. & M. Scholes. 1973. "The Pricing of Options and Corporate Liabilities." *Journal of Political Economy* 81:637–654.

Boukir, K., Y. Maday, B. Metivet & E. Razanfindrakoto. 1997. "A High Order Characteristics/Finite Element Method for the Incompressible Navier-Stokes Equations." *International Journal for Numerical Methods in Fluids* 25:1421–1454.

Boyle, P. P., J. Evnine & S. Gibbs. 1989. "Numerical Evaluation of Multivariate Contingent Claims." *Review of Financial Studies* 2:241–250.

Brennan, M. J. & E. S. Schwartz. 1977. "Convertible Bonds: Valuation and Optimal Strategies for Call and Conversion." *Journal of Finance* 32:1699–1715.

Brennan, M. J. & E. S. Schwartz. 1980. "Analysing Convertible Bonds." *Journal of Financial and Quantitative Analysis* 15(4):907–929.

Brezis, H. 1983. Analyse Fonctionelle, Théorie et Applications. In *Collection Mathématiques Appliquées Pour la Maitrise*. Masson.

Briys, E. & F. de Varenne. 1997. "Valuing Risky Fixed Rate Debt: An Extension." *Journal of Financial and Quantitative Analysis* 32(2):239–248.

Broadie, M. & J. Detemple. 1996. "American Option Evaluation: New Bounds, Approximation, and a Comparison of Existing Methods." *Review of Financial Studies* 9.

Broadie, M. & P. Glasserman. 1997. A Stochastic Mesh Method for Pricing High-Dimensional American Options. Working paper.

Carayannopoulos, P. 1996. "Valuing Convertible Bonds under the Assumption of Stochastic Interest Rates: An Empirical Investigation." *Quarterly Journal of Business and Economics* 35(3):17–31.

Carayannopoulos, P. & M. Kalimipalli. 2003. Convertible Bond Prices an Inherent Biases. Technical report The Mutual Group Financial Services Research Centre, School of Business and Economics, Wilfrid Laurier University, Waterloo, Ontario, Canada.

Carr, P. & G. Yang. 1997. Simulating Bermudan Interest Rate Derivatives. Working paper Morgan-Stanley.

Carr, P. & G. Yang. 1998. Simulating American Bond Options in a HJM Framework. Working paper Morgan-Stanley.

Cheuk, T. H. F. & T. C. F. Vorst. 1994. Lookback Options and the Observation Frequency: A Binomial Approach. Working paper Erasmus University.

Cheung, W. & I. Nelken. 1994. "Costing the Converts." *Risk* 7(7):47–49.

Cheung, W. & I. Nelken. 1995. Costing the Converts. In *Over the Rainbow*. Vol. 46 Risk Publications pp. 313–317.

Clarke, N. & K. Parrot. 1999. "Multigrid American Option Pricing with Stochastic Volatility." *Applied Mathematical Finance* 6:177–195.

Clewlow, L. & C. Strickland. 1998. *Implementing Derivative Models*. John Wiley & Sons.

Connolly, K. B. 1998. *Pricing Convertible Bonds*. John Wiley & Sons.

Constantinides, G. M. & B.D. Grundy. 1987. Call and Conversion of Convertible Corporate Bonds: Theory and Evidence. Working paper University of Chicago.

Cox, J., J. Ingersoll & S. Ross. 1985. "A Theory of the Term Structure of Interest Rates." *Econometrica* 53:385–467.

Cox, J., S. Ross & M. Rubinstein. 1979. "Option Pricing. A Simplified Approach." *Journal of Financial Economics* 7:229–263.

Crandall, M. G., H. Ishii & P-L. Lions. 1992. "User's Guide to Viscosity Solutions of Second Order Partial Differential Equations." *Bulletin American Mathematical Society* 27(1):1–67.

Das, S. 1998. *Credit Derivatives: Trading and Management of Credit and Default Risk*. Wiley Frontiers in Finance John Wiley & Sons.

Das, S. R. & R. K. Sundaram. 2004. "A Simple Model for Pricing Securities with Equity, Interest-Rate and Default Risk." Defaultrisk.com.

Davis, M. & F. Lischka. 2002. *Convertible Bonds with Market Risk and Credit Risk*. Studies in Advanced Mathematics American Mathematical Society/International Press pp. 45–58.

Douglas, J. & T. Russel. 1982. "Numerical Methods for Convection Dominated Diffusion Problems Based on Combining Methods of Characteristics with Finite Element Methods or Finite Differences." *SIAM J. on Numerical Analysis* 19(5):871.

Duffee, G. R. 1996. "Idiosyncratic Variation of Treasury Bill Yield." *Journal of Finance* 51:527–551.

Duffie, D. & K. J. Singleton. 1999. "Modeling Term Structure of Defaultable Bonds." *Review of Financial Studies* 12:687–720.

Duffie, D. & K. Singleton. 1994. Econometric Modelling of Term Structures of Defaultable Bonds. Working paper Stanford University.

Duffie, D. & K. Singleton. 1997. "An Econometric Model of the Term Structure of Interest Rate Swap Yields." *Journal of Finance* 52(4):1287–1321.

Duvaut, G. & J-L. Lions. 1972. Les Inéquations En Mécanique et En Physique. In *Travaux et Recherches Mathématiques*. Vol. 21 Paris: Dunod.

Elliot, C. M. & J. R. Ockendon. 1982. *Weak and Variational Methods for Moving Boundary Problems*. Research Notes in Mathematics Pitman.

Epstein, D., R. Haber & P. Wilmott. 2000. "Pricing and Hedging Convertible Bonds Under Non-Probabilistic Interest Rates." *Journal of Derivatives* pp. 31–40.

Ewing, R.E. & H. Wang. 2001. "A Summary of Numerical Methods for Time-Dependent Advection-Dominated Partial Differential Equations." *Journal of Computational and Applied Mathematics* 128:423–445.

Figlewski, S. & B. Gao. 1997. The Adaptive Mesh Model: A New Approach to Efficient Option Pricing. Working paper Stern School of Business, New York University.

Forsyth, P. A. & K. Vetzal. 2002. "Quadratic Convergence for Valuing American Options Using a Penalty Method." *SIAM Journal on Scientific Computation* 23:2096–2123.

Geske, R. 1977. "The Valuation of Corporate Liabilities as Compound Options." *Journal of Financial and Quantitative Analysis* 12:541–552.

Geske, R. & Shastri K. 1985. "Valuation by Approximation: A Comparison of Alternative Option Valuation Techniques." *Journal of Financial and Quantitative Analysis* 20(1):45–71.

Glowinski, R., J. L. Lions & R. Trémolières. 1973. *Analyse Numérique Des Inéquations Variationnelles*. Dunod.

Hamilton, D. T., G. Gupton & A. Berthault. 2000. Default and Recovery Rates of Corporate Bond Issuers: 2000. Special comment Moody's Investor Service, Global Credit Research.

He, H. 1990. "Convergence from Discrete to Continuous-Time Contingent Claims Prices." *Review of Financial Studies* 3:523–546.

Heath, D., R. Jarrow & A. Morton. 1992. "Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claim Valuation."*Ec onometrica* 60:77–105.

Heston, S. & G. Zhough. 2000. "On the Rate of Convergence of Discrete-Time Contingent Claims." *Mathematical Finance* 10(1):53–75.

Ho, T. & S. Lee. 1986. "Term Structure Movements and Pricing Interest Rate Contingent Claims." *Journal of Finance* 42:1129–1142.

Ho, T. S. Y. & D. M. Pfeffer. 1996. "Convertible Bonds: Model, Value, Attribution and Analytics." *Financial Analyst Journal* pp. 35–44.

Houston, P. & E. Suli. 2000. "Adaptive Lagrange-Galerkin Methods for Unsteady Convection-Diffusion Problems." *Mathematics of Computation* 70(233):77–106.

Hull, J. 2002. *Option, Futures and Other Derivatives*. US Imports & PHIPEs.

Hull, J. & A. White. 1995. "The Impact of Default Riskon the Prices of Options and Other Derivatives Securities." *Journal of Banking and Finance* 19(2):299–322.

Hull, J. C. & A. White. 1990. "Pricing Interest Rate Derivative Securities." *Review of Financial Studies* 3:573–592.

Hull, J. C. & A. White. 1993. "Efficient Procedures for Valuing European and American Path Dependent Options." *Journal of Derivatives* 1:21–31.

Hung, M.-W. & J.-Y. Wang. 2002. "Pricing Convertible Bonds Subject to Default Risk." *Journal of derivatives* 10:75–87.

Ingersoll, J. E. 1977*a*. "A Contingent Claim Valuation of Convertible Securities." *Journal of Financial Economics* 4:289–322.

Ingersoll, J. E., Jr. 1977*b*. "An Examination of Corporate Call Policies on Convertible Securities." *Journal of Finance* 32:463–478.

Jacod, J. & A. N. Shiryaev. 1988. *Limit Theorems for Stochastic Processes*. Springer, Berlin.

Jaillet, J., D. Lamberton & B. Lapeyre. 1990. "Variational Inequalities and the Pricing of American Options." *Acta Applicandae Mathematicae* 21:263–289.

Jalan, P. & G. Barone-Adesi. 1995. "Equity Financing and Corporate Convertible Bond Policy." *Journal of Banking and Finance* 19:187–206.

Jarrow, R. A. & S. M. Turnbull. 1995. "Pricing Derivatives on Financial Securities Subject to Credit Risk." *Journal of Finance* 50:53–85.

Kalotay, A., G. O. Williams & F. LJ. Fabozzi. 1993. "A Model for Valuing Bonds and Embedded Options." *Financial Analyst Journal* p. 35.

Kangro, R. & R. Nicolaides. 2000. "Far Field Boundary Conditions for Black-Scholes Equations." *SIAM Journal on Numerical Analysis* 38(4):1357–1368.

Kiesel, R., W. Perraudin & A. Taylor. 2002. Credit and Interest Rate Risk. In *Risk Management: Value at Risk and Beyond*, ed. M. H. A. Dempster. Cambridge.

King, R. 1986. "Convertible Bond Valuation: An Empirical Test."*J ournal of Financial Research* 9:53–69.

Kwok, Y. K. 1998. *Mathematical Models of Financial Derivatives*. Springer Finance Springer.

Lando, D. 1998. "On Cox Processes and Credit Risky Bonds." *Review of Derivatives Research* 2(2/3):99–120.

Lapeyre, B., A. Sulem & D. Talay. 2004. *Understanding Numerical Analysis for Option Pricing*. Cambridge University Press.

Leisen, D. & M. Reimer. 1996. "Binomial Models for Option Valuation Examining and Improving Convergence." *Applied Mathematical Finance* 3:319–346.

Lions, P.L. 1983. "Optimal Control of Diffusion Processes and Hamilton-Jacobi-Bellman Equations, Part 2: Viscosity Solutions and Uniqueness." *Communications in Partial Differential Equations* 8(11):1229–1276.

Longstaff, F. A. & E. S. Schwartz. 1995. "A Simple Approach to Valuing Risky Fixed and Floating Rate Debt." *Journal of Finance* 29:789–819.

Longstaff, F. & E. Schwartz. 1998. Valuing American American Options by Simulation: A Simple Least Squares Approach. Working paper Anderson School, UCLA.

Madan, D. B., F. Milne & H. Shefrin. 1989. "The Multinomial Option Pricing Model and its Brownian and Poisson Limits." *Review of Financial Studies* 4:251–266.

Marcozzi, M. D. 2001. "On the Approximation of Optimal Stopping Problems with Applications to Financial Mathematics." *SIAM, Journal of Scientific Computing* 22:1865–1884.

McConnell, J. J. & E. S. Schwartz. 1986. "LYON Taming." *The Journal of Finance* XLI(3):561–577.

Merton, R. C. 1973. "Theory of Rational Option Pricing." *Bell Journal of Economics and Management Science* 4:141–183.

Merton, R. C. 1974. "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates." *Journal of Finance* 29:449–470.

Morton, K. W. 1996. *Numerical Solution of Convection-Diffusion Problems*. Chapman & Hall.

Nielsen, L. T., J. Saa-Requejo & P. Santa-Clara. 1993. Default Risk and Interest Rate Risk: The Term Structure of Default Spreads. Working paper INSEAD.

Nyborg, K. G. 1996. "The Use and Pricing of Convertible Bonds." *Applied Mathematical Finance* 3:167–190.

Olsen, L. 2002. Convertible Bonds: A Technical Introduction. Research tutorial Barclays Capital.

Parés, C., M . Castro & J. Macías. 2002. "On the Convergence of the Bermúdez-Moreno Algorithm with Constant Parameters." *Journal of Numerical Mathematics* 92:113–128.

Philips, G. A. 1997. *Convertible Bond Markets*. MacMillan Press.

Pironneau, O. 1982. "On the Transport-Diffusion Algorithm and its Application to the Navier-Stokes Equations." *Journal of Numerical Mathematics* 38(3):309–332.

Pironneau, O. & F. Hetch. 2000. "Mesh Adaptation for the Black and Scholes Equations." *Journal of Numerical Mathematics* 8(1):25–35.

Pooley, D.M., P. A. Forsyth, K.R. Vetzal & R. B. Simpson. 2000. "Unstructured Meshing for Two Asset Barrier Options." *Applied Mathematical Finance* 7:33–60.

Protter, P. 1995. *Stochastic Integration and Differential Equations*. Vol. 21 of *Applications of Mathematics* 3 ed. Springer-Verlag.

R., Nelson. C. & A. F. Siegel. 1987. "Parsimonious Modelling of Yield Curves." *Journal of Business* 60:473.

Ramaswamy, K. & S. M. Sundaresan. 1986. "The Valuation of Floating Rate Instruments, Theory and Evidence." *Journal of Financial Economics* 17:251–272.

Realdon, M. 2003. Convertible Subordinated Debt Valuation and "Conversion in Distress". Working paper Department of Economics and Related Studies, University of York.

Rebonato, R. 1998. *Interest Rate Option Models*. John Wiley & Sons.

Rui, H. & Tabata M. 2001. "A Second Order Characteristic Finite Element Scheme for Convection-Diffusion Problems." *Journal of Numerical Mathematics* .

Schonbucher, P. J. 1996. Valuation of Securities Subject to Credit Risk. Working paper University of Bonn.

Schonbucher, P. J. 2003. *Credit Derivatives Pricing Models: Models, Pricing and Implementation*. John Wiley & Sons.

Schwartz, T. 1998. "Estimating the Term Structure of Corporate Debt." *Review of Derivatives Research* 2(2/3):193–230.

Selmin, V. & L. Formaggia. 1996. "Unified Construction of Finite Element and Finite Volume Discretisation for Compressible Flows." *International Journal for Numerical Methods in Engineering* 39:1–32.

Svensson, L. 1994. Estimating and Interpreting Forward Interest Rates: Sweden 1992-94. Working paper 114 IMF.

Svensson, L. 1995. "Estimating Forward Interest Rates with the Extended Nelson and Siegel Method." *Sveriges Riskbank Quaterly Review* 3:13.

Takahashi, A., T. Kobayahashi & N. Nakagawa. 2001. "Pricing Convertible Bonds with Default Risk: A Duffie-Singleton Approach." *The Journal of Fixed Income* 11(3):20–29.

Tavella, D. & C. Randall. 2000. *Pricing Financial Instruments: The Finite Difference Method*. John Wiley & Sons.

Tian, Y. 1993. "A Modified Lattice Approach to Option Pricing." *Journal of Futures Markets* 13:563–577.

Topper, J. 1998. Finite Element Modeling of Exotic Options. Discussion paper 216 Department of Economics, University of Hannover 216: .

Tseveriotis, K. & C. Fernandes. 1998. "Valuing Convertible Bonds with Credit Risk." *The Journal of Fixed Income* 8(2):95–102.

Unal, H., D. Madan & L. Guntay. 2001. A Simple Approach to Estimate Recovery Rates with APR Violation from Debt Spreads. Working paper University of Maryland.

Valuing Convertible Bonds as Derivatives. 1994. Quantitative strategies research notes Goldman Sachs, New York.

Vasiceck, O. & H. G. Fong. 1982. "Term Structure Modelling Using Exponential Splines." *Journal of Finance* 37(2):339–348.

Vasicek, O.A. 1977. "An Equilibrium Characterisation of the Term Structure." *Journal of Financial Economics* 5:177–188.

Vázquez, C. 1998. "An Upwind Numerical Approach for an American and European Option Pricing Model." *Applied Mathematics and Computation* 1997:273–286.

Waggoner, D. 1997. Spline Methods for Extracting Interest Rate Curves from Coupon Bond Prices. Technical Report 97-10 Federal Reserve Bank of Atlanta.

Wilmott, P. 1998. *Derivatives: The Theory and Practice of Financial Engineering*. John Wiley & Sons.

Wilmott, P., J. Dewynne & J. Howison. 1993. *Option Pricing: Mathematical Models and Computation*. Oxford Financial Press.

Windcliff, H., P. Forsyth & K. Vetzal. 2001. Asymptotic Boundary Conditions for the Black-Scholes Equation. Working paper University of Waterloo.

Winkler, G., T. Apel & U. Wystup. 2001. Valuation of Options in Heston's Stochastic Volatility Model Using Finite Element Methods. In *Foreign Exchange Risk*. London: Risk Publications.

Y., D'Halluin, P. Forsyth, K. Vetzal & G. Labahn. 2001. "A Numerical PDE Approach for Pricing Callable Bonds." *Applied Mathematical Finance* 8:49–77.

Yigitbasioglu, A. B. 2002. Pricing Convertible Bonds with Interest Rate, Equity and FX Risk. ISMA Center Discussion Papers in Finance University of Reading.

Zhou, C. 1997. "A Jump-Diffusion Approach to Modeling Credit Risk and Valuing Defaultable Securities." *Finance and Economics Discussion Series, Federal Reserve Board* 15.

Zhu, Y.-I. & Y. Sun. 1999. "The Singularity Separating Method for Two Factor Convertible Bonds." *Journal of Computational Finance* 3(1):91–110.

Zvan, R., P. A. Forsyth & K. R. Vetzal. 1998*a*. A General Finite Element Approach for PDE Option Pricing Model. Proceedings of Quantitative Finance 98, New York University of Waterloo.

Zvan, R., P. A. Forsyth & K. R. Vetzal. 1998*b*. "Robust Numerical Methods for PDE Models of Asian Options." *Journal of Computational Finance* 1:39–78.

Zvan, R., P. A. Forsyth & K. R. Vetzal. 1999. "A Finite Element Approach to the Pricing of Discrete Lookbacks with Stochastic Volatility." *Applied Mathematical Finance* 6:87–106.

Zvan, R., P. A. Forsyth & K. R. Vetzal. 2000. "PDE Methods for Pricing Barrier Options." *Journal of Economic Dynamics and Control* 24.

Zvan, R., P. A. Forsyth & K. R. Vetzal. 2001. "A Finite Volume Approach for Contingent Claims Valuation." *IMA Journal of Numerical Analysis* 21:703–721.

Zvan, R., P. A. Forsyth & K. R. Vetzal. 2002. "Convergence of Lattice and PDE Methods Valuing Path Dependent Options with Interpolation." *Review of Derivatives Research* 5:273–314.

Appendix A
Finite Element Calculations

Let $\Omega \subset \mathbb{R}^n$ ($n = 1, 2, 3$ in practice) be an open bounded set with boundary $\Gamma = \Gamma_D \cup \Gamma_R$ and

$$a_{ij},\ a_0 \in L^{\infty}(\Omega) \quad 1 \leq i, j \leq n, \tag{A.1}$$

Γ_D and Γ_R two parts of Γ such that $\Gamma = \Gamma_D \cup \Gamma_R$ and $measure(\Gamma_D \cap \Gamma_R) = 0$,

$$\beta \in L^{\infty}(\Gamma_R),\ g \in L^2(\Gamma_R), \tag{A.2}$$

and

$$f \in L^2(\Omega). \tag{A.3}$$

We seek for a solution $u \in V$ of the variational equality

$$\begin{aligned} \sum_{i,j=1}^{n} \int_{\Omega} a_{ij} \frac{\partial u}{\partial x_j} \frac{\partial v}{\partial x_i} dx + \int_{\Omega} a_0 u v dx + \int_{\Gamma_R} \beta u v d\Gamma &= \int_{\Omega} f v dx + \int_{\Gamma_R} g v d\Gamma, \\ \forall\, v &\in V, \end{aligned} \tag{A.4}$$

where

$$V = \left\{ u \in H^1(\Omega) :\ u_{/\Gamma_D} = 0 \right\}. \tag{A.5}$$

We recall that $H^1(\Omega)$ and V are Hilbert spaces when the following scalar product is considered

$$(u, v) = \sum_{i,j=1}^{n} \int_{\Omega} \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} dx + \int_{\Omega} u v dx, \tag{A.6}$$

from which the following norm is derived

$$\|u\| = (u, u)^{1/2}, \tag{A.7}$$

and the distance

$$d(u, v) = \|u - v\| . \tag{A.8}$$

Equation (A.4) can be considered a particular case of the following problem. Let V be a Hilbert space

Find $u \in V$ such that

$$a(u, v) = L(v) \ \ \forall v \in V, \tag{A.9}$$

where $a : V \times V \to \mathbb{R}$ is a bilinear function and $L : V \to \mathbb{R}$ is a linear function.

It is possible to study directly the existence and uniqueness of solution of this problem in a general abstract framework. The following result is known as Lax-Milgram Theorem:

Proposition 1 Under the hypothesis:

- $a : V \times V \to \mathbb{R}$ bilinear
- $a(v, v) \geq \alpha \|v\|_V^2 \ \forall v \in V$ (coerciveness)
- $a(u, v) \leq M \|u\|_V \|v\|_V \ \forall u, v \in V$ (continuity)
- $L : V \to \mathbb{R}$ lineal
- $L(v) \leq C \|v\|_V \ \forall v \in V$ (continuity)

Problem (A.9) has a unique solution.

Notice that it is not necessary for a to be symmetric, i.e.,

$$a(u, v) = a(v, u) \ \ \forall u, v \in V. \tag{A.10}$$

If a is symmetric it can be shown that the solution of (A.9) is also the only element of V that minimizes the functional

$$J(v) = \frac{1}{2}a(v, v) - L(v). \tag{A.11}$$

Actually, (A.9) is nothing but the Euler equation of the problem of "calculus of variations":

Find $u \in V$ such that

$$J(u) \leq J(v) \quad \forall v \in V, \tag{A.12}$$

because $J'(u)(v) = a(u, v) - L(v)$. This fact justifies the term "variational formulation" for (A.9).

A.1 Discretization: Galerkin's method

In order to solve numerically the problem (A.4) a discretization must be done, i.e., the problem must be replaced by a new one with a finite number of degrees of freedom or unknowns.

One method, known as Galerkin's method, replaces the space of functions V by a finite dimensional space V_h and defines the following discrete problem:

Find $u_h \in V_h$ such that

$$a(u_h, v_h) = L(v_h) \;\; \forall v_h \in V_h. \tag{A.13}$$

Let $B = \{\phi_1, \phi_2, \ldots, \phi_N\}$ be a basis of V_h. Then the solution of (A.13) can be written in the form

$$u_h = \sum_{j=1}^{N} \xi_j \phi_j, \tag{A.14}$$

so that the discrete problem is equivalent to find N numbers $(\xi_1, \xi_2, \ldots, \xi_N)$ satisfying

$$\sum_{j=1}^{N} a(\phi_j, \phi_i)\xi_j = L(\phi_i) \quad i = 1, 2, \ldots, N. \tag{A.15}$$

Equivalently this linear system can be written in the more compact way

$$A_h \xi_h = b_h, \tag{A.16}$$

where

$$A_{ml} = a(\phi_l, \phi_m) = \sum_{i,j=1}^{n} \int_{\Omega} a_{ij} \frac{\partial \phi_l}{\partial x_j} \frac{\partial \phi_m}{\partial x_i} dx + \int_{\Omega} a_0 \phi_l \phi_m dx + \int_{\Gamma_R} \beta \phi_l \phi_m d\Gamma, \tag{A.17}$$

$$b_m = L(\phi_m) = \int_{\Omega} f \phi_m dx + \int_{\Gamma_R} g \phi_m d\Gamma. \tag{A.18}$$

If a is symmetric, clearly the matrix A is symmetric. If, besides, a is coercive then A is positive definite, and therefore Choleski's method can be used to solve the system of equations (A.16). On the other hand, if a is symmetric, the solution of the discrete problem is the unique minimum of the functional J in the space V_h.

An important problem once this stage has been reached, is the estimation of the error that is realized in the discretization, i.e., the distance between u and u_h. There exist a mathematical theory that deals with this question, but it is beyond the aim of these notes. Let us just say that it can be proved that

$$\|u - u_h\|_V \leq \frac{M}{\alpha} \inf_{v_h \in V_h} \|u - v_h\|_V. \tag{A.19}$$

This inequality shows that the order of the error realized is analogous to the error that is achieved if u is replaced but its best approximation in V_h.

On the other hand if the equalities (A.9) (with $v = v_h$) and (A.13) are subtracted, the following is obtained

$$a(u - u_h, v_h) = 0 \quad \forall v_h \in V_{h,} \tag{A.20}$$

and this proves that u_h is the only element of V_h that minimizes the function

$$v_h \rightarrow a(u - v_h, u - v_h) \quad in\ V_h. \tag{A.21}$$

The problem arising now is how to chose the space V_h and the basis $B = \{\phi_1, \phi_2, \ldots, \phi_N\}$. This space is usually made up of globally continuous functions which are polynomials in each element of a polygonal mesh of the domain Ω. The elements of the basis are functions that become zero in big regions of Ω so that many terms of the matrix A are zero, i.e., A is a sparse matrix. We will describe how to build Lagrange Finite Elements in an n-dimensional domain Ω. Then we will proceed to work out the matrix of coefficients and the independent term in the particular case of Lagrange triangular finite elements of degree one and in a two dimensional space.

A.2 Lagrange triangular finite elements

The domain Ω is decomposed in simplex of dimension n (triangles if $n = 2$, tetrahedron if $n = 3$, etc.) and the space V_h is the space of the continuous functions in $\bar{\Omega}$ that are polynomials of degree smaller or equal than k over any single simplex.

A set of $n+1$ points not laying on the same hyperplane is considered, i.e., such that the matrix

$$\begin{bmatrix} a_{11} & \cdots\cdots & a_{1n} & 1 \\ \vdots & & & \vdots \\ a_{n1} & \cdots\cdots & a_{nn} & 1 \\ a_{n+11} & \cdots\cdots & a_{n+1n} & 1 \end{bmatrix} \tag{A.22}$$

has non zero determinant.

The convex envelope of these $n+1$ points, i.e., the set

$$K = \left\{ x = \sum_{i=1}^{N+1} \lambda_i a_i, \quad 0 \leq \lambda_i \leq 1, \quad 1 \leq i \leq n+1, \quad \sum_{i=1}^{N+1} \lambda_i = 1 \right\}, \tag{A.23}$$

is called n-dimensional simplex.

If $x \in K$ the correspondent $\lambda_i = \lambda_i(x)$ are known as baricentric coordinates of x. Notice that

$$\lambda_i(a_j) = \delta_{ij},$$

and that λ_i is an affine function (polynomial of degree one) in the variables x_j.

The subsets of K obtained when the following conditions are imposed

$$\lambda_{i_1} = \lambda_{i_2} = \ldots = \lambda_{i_r} = 0,$$

are called $n-r$ dimensional faces of K.

The baricenter of K is the point that has all the baricentric coordinates equal, i.e.,

$$\lambda_i = \frac{1}{n+1} \quad 1 \leq i \leq n+1.$$

Let K be an n-simplex and k a positive integer; the subset of points of K

$$\sum_k(K) = \left\{ x \in K : \lambda_j(x) \in \{0, \frac{1}{k}, \ldots, \frac{k-1}{k}, 1\}, \qquad 1 \leq j \leq n+1 \right\}. \tag{A.24}$$

is called lattice of order k in K.

Recall that $\wp_k$ is the space of polynomials of degree equal or less than k. Since an homogeneous polynomial of n variables and degree j has $\begin{pmatrix} n+j-1 \\ j \end{pmatrix}$ terms, the dimension of $\wp_k$ will be

$$1 + \begin{pmatrix} n \\ 1 \end{pmatrix} + \begin{pmatrix} n+1 \\ 2 \end{pmatrix} + \cdots + \begin{pmatrix} n+k-1 \\ k \end{pmatrix} = \begin{pmatrix} n+k \\ k \end{pmatrix}. \tag{A.25}$$

If $x = \sum_{i=1}^{N+1} \lambda_i(x) a_i$ is an element of $\sum_k(K)$ then,

$$x = \frac{1}{k}\sum_{i=1}^{N+1} \mu_i a_i \quad with \quad \mu_i \in \{0,\ldots,k\} \quad and \quad \sum_{i=1}^{N+1} \mu_i = k, \tag{A.26}$$

therefore we will write

$$x = a_\mu \quad with \quad \mu = (\mu_1, \ldots, \mu_{n+1}). \tag{A.27}$$

It can be proved that the only polynomial $q \in \wp_k$ such that

$$q(a_\mu) = 1, \tag{A.28}$$

and

$$q(b) = 0 \quad \forall b \in \sum_k, \quad b \neq a_\mu, \tag{A.29}$$

is given by

$$q_\mu(x) = \left[\prod_{j=1}^{n+1} \mu_j!\right]^{-1} \prod_{j=1,\mu_j\geq 1}^{n+1} \prod_{i=0}^{\mu_j-1} (k\lambda_j(x) - 1). \tag{A.30}$$

With the triple $(K, \sum_k(K), \wp_k)$ we will build spaces of approximation V_h.

Let τ_h be a partition of $\bar{\Omega}$ into simplices such that every face of a simplex K_i of τ_h is:

- Either subset of Γ_D
- Or subset of Γ_R

- Or a face of other simplex K_j of τ_h; in such a case K_i and K_j are said to be adjacent

The diameter of K is denoted by h_K and $h = \max h_K,\ K \in \tau_h$.

For $k \geq 1$ the following functional spaces are built associated to τ_h

$$X_h^{(k)} = \left\{u_h \in C^0(\bar{\Omega}) : u_{h|K} \in \wp_k(K) \quad \forall K \in \tau_h\right\}. \tag{A.31}$$

Clearly,

$$X_h^{(k)} \subset \prod_{K \in \tau_h} \wp_k(K). \tag{A.32}$$

This inclusion simply states that any function of $X_h^{(k)}$ is a polynomial of degree equal or less than k over each individual element. Conversely, what is the necessary and sufficient condition for an element of X to be in $X_h^{(k)}$, i.e., for the polynomial pieces to stick with continuity? The answer is simple: the above will hold if and only if for every pair of adjacent elements K_1 and K_2 the pieces defined on them agree on the points

$$\sum_k(K_1) \cap \sum_k(K_2). \tag{A.33}$$

Therefore any function in $X_h^{(k)}$ is uniquely determined by its values at the points of the set

$$\sum_h^{(k)} = \bigcup_{K \in \tau_h} \sum_k(K). \tag{A.34}$$

From now on, τ_h will be called "triangulation" of Ω (even if the dimension n is different from 2) and the elements of $\sum_h^{(k)}$ "nodes of the triangulation". Notice that there can be nodes that are not vertices.

The dimension of the space $X_h^{(k)}$ equals the number of nodes. Besides it is possible to define a basis of $X_h^{(k)}$ such that its elements are functions with support reduced to a few elements of τ_h. Let N_h be the number of nodes, that we will assume to be numbered.

$$\sum\nolimits_h^{(k)} = \{b_i : i = 1, \ldots, N_h\}. \quad (A.35)$$

The node b_i contributes to the basis with the function $\phi_i \in X_h^{(k)}$ uniquely determined by

$$\phi_i(b_j) = \delta_{ij} \quad 1 \leq j \leq N_h. \quad (A.36)$$

The sets $X_h^{(k)}$ are not subspaces of V because their elements do not satisfy, initially, the boundary condition $v_{h/\Gamma_D} = 0$. Therefore, in order to solve the problem (A.4) the following spaces must be used

$$V_h^{(k)} = \left\{ v_h \in X_h^{(k)} : v_{h/\Gamma_D} = 0 \right\}. \quad (A.37)$$

Notice that the condition $v_{h/\Gamma_D} = 0$ is equivalent to the condition that v_h is zero at the nodes that belong to Γ_D. In that way, if we get rid off the elements of the basis that correspond to this nodes, a basis of $V_h^{(k)}$ will be obtained.

A.3 Coefficients matrix and independent term

We consider the problem in two dimensions ($n = 2$). Let us see how to organize the calculations to build A_h and b_h if we chose the space of Lagrange triangular finite elements of degree one, i.e.,

$$V_h = \left\{ v_h \in C^0(\bar{\Omega}) : v_{h|K} \in \wp_1(K) \quad \forall K \in \tau_h \quad v_{h|\Gamma_D} = 0 \right\}. \quad (A.38)$$

First we consider the calculations as if the boundary condition $v_{h/\Gamma_D} = 0$ did not exist, i.e., assuming that

$$V_h = \left\{ v_h \in C^0(\bar{\Omega}) : v_{h|K} \in \wp_1(K) \quad \forall K \in \tau_h \right\} = X_h^{(1)}. \quad (A.39)$$

Let $\sum_h^{(1)}$ be the set of nodes of the triangulation τ_h that we will assume to be numbered

$$\sum_h^{(1)} = \{b_i : i = 1, \ldots, N_h\}. \tag{A.40}$$

Notice that the dimension of $X_h^{(1)}$ equals N_h. Any node b_i defines an element in the basis $\phi_i \in X_h^{(1)}$ such that

$$\phi_i(b_j) = \delta_{ij} \quad 1 \leq i, j \leq N_h. \tag{A.41}$$

Then the solution u_h can be written as

$$u_h = \sum_{j=1}^{N_h} \xi_j \phi_j \text{ and } \xi_j = u_h(b_j).$$

Therefore the column vector $(u_h(b_1), \ldots, u_h(b_{N_h}))$ is the solution of the linear system

$$A_h \xi_h = b_h \tag{A.42}$$

where

$$A_{ml} = a(\phi_l, \phi_m) = \sum_{i,j=1}^{2} \int_\Omega a_{ij} \frac{\partial \phi_l}{\partial x_j} \frac{\partial \phi_m}{\partial x_i} dx + \int_\Omega a_0 \phi_l \phi_m \, dx + \int_{\Gamma_R} \beta \phi_l \phi_m \, d\Gamma, \tag{A.43}$$

and

$$b_m = L(\phi_m) = \int_\Omega f \phi_m \, dx + \int_{\Gamma_R} g \phi_m \, d\Gamma, \quad 1 \leq l, m \leq N_h. \tag{A.44}$$

As it will be seen later the calculation of A_h and b_h using this formula is inefficient because the same integrals are calculated several times over the same triangles. The method described below, which is the one used in practice, is based on the concept of "elementary matrix" and "assembling".

Let us remind that the discretised problem can be written as follows

Find $u_h \in V_h$ such that

$$\sum_{i,j=1}^{2} \int_{\Omega} a_{ij} \frac{\partial u_h}{\partial x_j} \frac{\partial v_h}{\partial x_i} dx + \int_{\Omega} a_0 u_h v_h \, dx + \int_{\Gamma_R} \beta u_h v_h \, d\Gamma$$
$$= \int_{\Omega} f v_h \, dx + \int_{\Gamma_R} g v_h \, d\Gamma \quad \forall v \in V_h. \tag{A.45}$$

Let us consider the first term of the right hand side term of this equality. We have that

$$\sum_{i,j=1}^{2} \int_{\Omega} a_{ij} \frac{\partial u_h}{\partial x_j} \frac{\partial v_h}{\partial x_i} dx = \sum_{K \in \tau_h} \int_{K} \begin{pmatrix} \frac{\partial v_h}{\partial x_1} & \frac{\partial v_h}{\partial x_2} \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} \frac{\partial u_h}{\partial x_1} \\ \frac{\partial u_h}{\partial x_2} \end{pmatrix} dx. \tag{A.46}$$

Let a_1^K, a_2^K, a_3^K be the vertex of the triangle K and m_{1K}, m_{2K}, m_{3K} the corresponding numbers in the numbering of τ_h, i.e., assume

$$a_1^K = b_{m_{1K}}; \quad a_2^K = b_{m_{2K}}; \quad a_3^K = b_{m_{3K}}. \tag{A.47}$$

Let $v_h \in X_h^{(1)}$ then $v_{h/K} = \sum_{j=1}^{3} v_h(a_i^K) p_i^K$, where p_i^K is the only polynomial of degree equal or less than one such that

$$p_i^K(a_j^K) = \delta_{ij}. \tag{A.48}$$

Equivalently

$$v_{h/K} = \begin{pmatrix} p_1^K & p_2^K & p_3^K \end{pmatrix} \begin{pmatrix} v_h(a_1^K) \\ v_h(a_2^K) \\ v_h(a_3^K) \end{pmatrix} = [P^K](v_K). \tag{A.49}$$

Therefore

$$\begin{pmatrix} \frac{\partial v_h}{\partial x_1} \\ \frac{\partial v_h}{\partial x_2} \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{3} v_h(a_i^K) \frac{\partial p_i^K}{\partial x_1} \\ \sum_{i=1}^{3} v_h(a_i^K) \frac{\partial p_i^K}{\partial x_2} \end{pmatrix}$$
$$= \begin{pmatrix} \frac{\partial p_1^K}{\partial x_1} & \frac{\partial p_2^K}{\partial x_1} & \frac{\partial p_3^K}{\partial x_1} \\ \frac{\partial p_1^K}{\partial x_2} & \frac{\partial p_2^K}{\partial x_2} & \frac{\partial p_3^K}{\partial x_2} \end{pmatrix} \begin{pmatrix} v_h(a_1^K) \\ v_h(a_2^K) \\ v_h(a_3^K) \end{pmatrix}$$
$$= [DP^K](v_K). \tag{A.50}$$

Substituting the expression obtained we find that

$$\sum_{i,j=1}^{2} \int_{\Omega} a_{ij} \frac{\partial u_h}{\partial x_j} \frac{\partial v_h}{\partial x_i} dx = \sum_{K \in \tau_h} \int_{K} {}^t v_{h/K} {}^t \left[DP^K\right] [E] \left[DP^K\right] u_{h/K} dx, \tag{A.51}$$

where $[E]_{ij} = a_{ij}$.

A similar process for the other terms leads to

$$\sum_{K \in \tau_h} {}^t(v_K) \left[\int_K {}^t\left[DP^K\right] [E] \left[DP^K\right] dx + \int_K a_0 {}^t\left[P^K\right] \left[P^K\right] dx \right] (u_K) \tag{A.52}$$

The terms in brackets are called "elementary matrix" and "elementary right hand side" of the element K respectively. Notice that it is a 3×3 matrix and a 3 component vector.

If v is a vector of N_h components, we have that

$$\left[B^K\right] (v) = \begin{bmatrix} v_{m_{1_K}} \\ v_{m_{2_K}} \\ v_{m_{3_K}} \end{bmatrix} = (v_K), \tag{A.53}$$

where $\left[B^K\right]$ is the Boolean matrix

$$\left[B^K\right] = \begin{bmatrix} 0 & \cdots & 1_{m_{1K}} & 0 & \cdots & 0 \\ 0 & 1_{m_{2K}} & 0 & \cdots & \cdots & 0 \\ 0 & \cdots & \cdots & 1_{m_{3K}} & \cdots & 0 \end{bmatrix}. \tag{A.54}$$

Therefore the matrix $\left[B^K\right]$ selects among the set of all degrees of freedom $v \in \mathbb{R}^{N_h}$ the three that correspond to the element K.

In that way it can be written

$$\begin{aligned} &{}^t(v_h) \left[\sum_{K \in \tau_h} {}^t\left[B^K\right] \left(\int_K {}^t\left[DP^K\right] [E] \left[DP^K\right] dx + \int_K a_0 {}^t\left[P^K\right] \left[P^K\right] dx \right) \left[B^K\right] \right] (u_h) \\ &= {}^t(v_h) \left[\sum_{K \in \tau_h} {}^t\left[B^K\right] \left(\int_K {}^t\left[P^K\right] f dx + \int_{\partial K \cap \Gamma_R} {}^t\left[P^K\right] g d\Gamma \right) \right]. \end{aligned} \tag{A.55}$$

Notice that this equality must be satisfied for all $v_h \in \mathbb{R}^{N_h}$ therefore

$$A_h = \sum_{K \in \tau_h} {}^t\left[B^K\right] \left[A_h^K\right] \left[B^K\right] \text{ and } b_h = \sum_{K \in \tau_h} {}^t\left[B^K\right] \left[b_h^K\right], \tag{A.56}$$

where

$$\left[A_h^K\right] = \left(\int_K {}^t\left[DP^K\right]\left[E\right]\left[DP^K\right] dx + \int_K a_0 {}^t\left[P^K\right]\left[P^K\right] dx \right), \tag{A.57}$$

and

$$\left[b_h^K\right] = \int_K {}^t\left[P^K\right] f dx + \int_{\partial K \cap \Gamma_R} {}^t\left[P^K\right] g d\Gamma. \tag{A.58}$$

The operations in (A.56) are known with the name of "assembling" of the fundamental matrix and the right-hand side terms. Let us see how to do it. We have that

$$\left[B^K\right]_{ij} = \delta_{m_i K j} \tag{A.59}$$

Therefore

$$\begin{aligned} \left({}^t\left[B^K\right]\left[A_h^K\right]\left[B^K\right]\right)_{ij} &= \sum_{n=1}^{3} \left({}^t\left[B^K\right]\left[A_h^K\right]\right)_{in} \left[B^K\right]_{nj} \\ &= \sum_{n=1}^{3}\sum_{l=1}^{3} \left[B^K\right]_{li}\left[A_h^K\right]_{ln}\left[B^K\right]_{nj} \\ &= \sum_{n=1}^{3}\sum_{l=1}^{3} \delta_{m_l K i}\left[A_h^K\right]_{ln}\delta_{m_n K j}. \end{aligned} \tag{A.60}$$

The above equals

$$0 \quad if\ i \neq m_{lK}\ or\ j \neq m_{nK} \quad l, n = 1, 2, 3 \tag{A.61}$$

$$\left[A_h^K\right]_{ln} \quad if\ i = m_{lK}\ and\ j = m_{nK} \quad \text{for some } l, n = 1, 2, 3. \tag{A.62}$$

In that way, for the calculation of A_h and b_h the following algorithm can be used:

- Initialize A_h and b_h to zero
- Do a loop over the elements of τ_h.

For every $K \in \tau_h$, A_h^K and b_h^K are worked out and then we define

$$(A_h)_{m_{\alpha K} m_{\beta K}} = (A_h)_{m_{\alpha K} m_{\beta K}} + \left[A_h^K\right]_{\alpha\beta}, \tag{A.63}$$

$$[b_h]_{m_{\alpha K}} = (b_h)_{m_{\alpha K}} + \left[b_h^K\right]_{\alpha}. \tag{A.64}$$

Later on we will come back again to the assembling.

A.3.1 Change to the reference element

Let $\hat{K}$ be he triangle of vertices

$$\hat{a}_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \ \hat{a}_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \ \hat{a}_3 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{A.65}$$

that we will call reference triangle. Let K be any element of τ_h. There exists a unique affine inversible application $F_K : \hat{K} \to K$ such that $F_K(\hat{a}_i) = a_i^K \quad i = 1, 2, 3$. It is the application

$$F_K(\hat{x}) = C_K \hat{x} + a_1^K, \tag{A.66}$$

where C_K is the matrix

$$C_K = (a_2^K - a_1^K, a_3^K - a_1^K). \tag{A.67}$$

It is easy to check that

$$p_i^K \circ F_K = \hat{p}_i, \ \ i = 1, 2, 3, \tag{A.68}$$

where

$$\hat{p}_1(\hat{x}_1, \hat{x}_2) = 1 - \hat{x}_1 - \hat{x}_2, \tag{A.69}$$

$$\hat{p}_2(\hat{x}_1, \hat{x}_2) = \hat{x}_1, \tag{A.70}$$

$$\hat{p}_3(\hat{x}_1, \hat{x}_2) = \hat{x}_2. \tag{A.71}$$

Indeed, $\hat{p}_i \in \wp_1$ and also $(p_i^K \circ F_K)(\hat{a}_j) = \hat{p}_i(\hat{a}_j) = \delta_{ij}$.

The integrals that appear in the calculation of A_h^K and b_h^K will be done through a change of variable to the reference element.

Let us remind the formula of the change of variable

$$\int_K \psi dx = \int_{\hat{K}} \psi \circ F_K |\det C_k| d\hat{x}. \tag{A.72}$$

On the other hand by the "chain rule" we have that

$$\begin{pmatrix} \frac{\partial \hat{p}_\alpha}{\partial \hat{x}_1}(\hat{x}) \\ \frac{\partial \hat{p}_\alpha}{\partial \hat{x}_2}(\hat{x}) \end{pmatrix} = \begin{pmatrix} \frac{\partial (F_K)_1}{\partial \hat{x}_1}(\hat{x}) & \frac{\partial (F_K)_2}{\partial \hat{x}_1}(\hat{x}) \\ \frac{\partial (F_K)_1}{\partial \hat{x}_2}(\hat{x}) & \frac{\partial (F_K)_2}{\partial \hat{x}_2}(\hat{x}) \end{pmatrix} \begin{pmatrix} \frac{\partial p_\alpha^K}{\partial x_1}(F_K(\hat{x})) \\ \frac{\partial p_\alpha^K}{\partial x_2}(F_K(\hat{x})) \end{pmatrix}. \tag{A.73}$$

Therefore

$$\begin{pmatrix} \frac{\partial p_\alpha^K}{\partial x_1} \\ \frac{\partial p_\alpha^K}{\partial x_2} \end{pmatrix} = {}^tC_K^{-1} \begin{pmatrix} \frac{\partial \hat{p}_\alpha}{\partial \hat{x}_1} \\ \frac{\partial \hat{p}_\alpha}{\partial \hat{x}_2} \end{pmatrix}, \tag{A.74}$$

or, in summarized form

$$\left[DP^K\right] = {}^tC_K^{-1}\left[\hat{D}\hat{P}\right]. \tag{A.75}$$

Substituting this expression for $\left[DP^K\right]$ in (A.57) we obtain

$$\begin{aligned} A_h^K &= \int_{\hat{K}} {}^t\left[\hat{D}\hat{P}\right] \; [C_K^{-1}][E]^t[C_K^{-1}] \; \left[\hat{D}\hat{P}\right] |\det C_K| d\hat{x} \\ &+ \int_{\hat{K}} a_0{}^t\left[\hat{P}\right] \; \left[\hat{P}\right] |\det C_K| d\hat{x}. \end{aligned} \tag{A.76}$$

If we call $[G]$ the matrix $[C_K^{-1}][E]^t[C_K^{-1}]$ and $\Delta = |\det C_K|$ it turns out that

$$A_h^K = \Delta \int_{\hat{K}} {}^t\left[\hat{D}\hat{P}\right] \; [G] \; \left[\hat{D}\hat{P}\right] d\hat{x} + \Delta \int_{\hat{K}} a_0{}^t\left[\hat{P}\right] \; \left[\hat{P}\right] d\hat{x}. \tag{A.77}$$

Therefore

$$
\begin{aligned}
\left[A_h^K\right]_{\alpha\beta} &= \Delta \sum_{\mu,\gamma=1}^{2} \int_{\hat{K}} {}^t\left[\hat{D}\hat{P}\right]_{\alpha\mu} [G]_{\mu\gamma} \left[\hat{D}\hat{P}\right]_{\gamma\beta} d\hat{x} + \Delta \int_{\hat{K}} a_0 \hat{p}_\alpha \hat{p}_\beta d\hat{x} \\
&= \Delta \int_{\hat{K}} \sum_{\mu,\gamma=1}^{2} [G]_{\mu\gamma} \frac{\partial \hat{p}_\alpha}{\partial \hat{x}_\mu} \frac{\partial \hat{p}_\beta}{\partial \hat{x}_\gamma} d\hat{x} + \Delta \int_{\hat{K}} a_0 \hat{p}_\alpha \hat{p}_\beta d\hat{x}.
\end{aligned} \tag{A.78}
$$

If the coefficients a_{ij}, a_0 are constant in K, then $[G]$ is constant in K and

$$
\left[A_h^K\right]_{\alpha\beta} = \Delta \sum_{\mu,\gamma=1}^{2} [G]_{\mu\gamma} \int_{\hat{K}} \sum_{\mu,\gamma=1}^{2} \frac{\partial \hat{p}_\alpha}{\partial \hat{x}_\mu} \frac{\partial \hat{p}_\beta}{\partial \hat{x}_\gamma} d\hat{x} + \Delta a_0 \int_{\hat{K}} \hat{p}_\alpha \hat{p}_\beta d\hat{x}. \tag{A.79}
$$

The numbers

$$
H_{\alpha\beta\mu\gamma} = \int_{\hat{K}} \frac{\partial \hat{p}_\alpha}{\partial \hat{x}_\mu} \frac{\partial \hat{p}_\beta}{\partial \hat{x}_\gamma} d\hat{x} \ \text{ and } \ J_{\alpha\beta} = \int_{\hat{K}} \hat{p}_\alpha \hat{p}_\beta d\hat{x}, \tag{A.80}
$$

do not depend on the element considered and are calculated just once. Besides, notice that

$$
H_{\alpha\beta\mu\gamma} = H_{\beta\alpha\gamma\mu}, \quad J_{\alpha\beta} = J_{\beta\alpha} \ \text{ and } \ \int_{\hat{K}} \hat{p}_\alpha^r \hat{p}_\beta^s d\hat{x} = \frac{r!s!}{(r+s+2)!} \tag{A.81}
$$

In this way just the matrix $[G]$ and a_0 depend on the element. The matrix $[G]$ is worked out using the values of a_{ij} and the coordinates of the vertex a_i^K. Completing the calculations described we obtain

$$
\begin{aligned}
A_h^K &= \frac{\Delta}{2} \begin{bmatrix} g_{11} + 2g_{12} + g_{22} & -(g_{11} + g_{21}) & -(g_{12} + g_{22}) \\ -(g_{11} + g_{12}) & g_{11} & g_{12} \\ -(g_{21} + g_{22}) & g_{21} & g_{22} \end{bmatrix} \\
&+ \frac{a_0 \Delta}{12} \begin{bmatrix} 1 & 1/2 & 1 \\ 1/2 & 1 & 1/2 \\ 1/2 & 1/2 & 1 \end{bmatrix},
\end{aligned} \tag{A.82}
$$

where

$$
\begin{aligned}
g_{11} &= \Delta^{-2}\left\{a_{11}\left(a_{32}^K - a_{12}^K\right)^2 + (a_{12}+a_{21})\left(a_{11}^K - a_{31}^K\right)\left(a_{32}^K - a_{12}^K\right)\right. \\
&\quad \left. + a_{22}\left(a_{11}^K - a_{31}^K\right)^2\right\}, \\
g_{12} &= g_{21} = \Delta^{-2}\left\{a_{11}\left(a_{12}^K - a_{22}^K\right)\left(a_{32}^K + a_{12}^K\right) + a_{12}\left(a_{21}^K - a_{11}^K\right)\left(a_{32}^K - a_{12}^K\right)\right. \\
&\quad \left. + a_{21}\left(a_{12}^K - a_{22}^K\right)\left(a_{11}^K - a_{31}^K\right) + a_{22}\left(a_{21}^K - a_{11}^K\right)\left(a_{11}^K - a_{31}^K\right)\right\}, \\
g_{22} &= \Delta^{-2}\left\{a_{11}\left(a_{12}^K - a_{22}^K\right)^2 + (a_{12}+a_{21})\left(a_{21}^K - a_{11}^K\right)\left(a_{12}^K - a_{22}^K\right) + a_{22}\left(a_{21}^K - a_{11}^K\right)^2\right\}.
\end{aligned}
$$

With respect to the right hand side term we have that

$$b_h^K = \Delta \int_{\hat{K}} {}^t\left[\hat{P}\right] f\left(F_K(\hat{x})\right) d\hat{x} + \int_{\partial \hat{K}} {}^t\left[\hat{P}\right] g\left(F_K(\hat{x})\right) d\Gamma. \tag{A.83}$$

In order to calculate the second term of this expression, we will define parameterization of the edges of K

$$
\begin{aligned}
\textit{Edge 1:}\ \varphi_K^1(\hat{\sigma}) &= F_K(\hat{\sigma}, 0) = \left(a_2^K - a_1^k\right)\hat{\sigma} + a_1^K \\
\textit{Edge 2:}\ \varphi_K^2(\hat{\sigma}) &= F_K(1-\hat{\sigma}, \hat{\sigma}) = \left(a_3^K - a_2^k\right)\hat{\sigma} + a_2^K \\
\textit{Edge 3}\ .\ \varphi_K^3(\hat{\sigma}) &= F_K(0, 1-\hat{\sigma}) = \left(a_1^K - a_3^k\right)\hat{\sigma} + a_3^K
\end{aligned}
$$

Therefore

$$\int_{l\ edge} \psi d\tau = \int_0^1 \psi\left(\varphi_K^l(\hat{\sigma})\right)\left|a_{l+1}^K - a_l^K\right| d\hat{\sigma} \quad \left(a_4^K = a_1^K\right), \tag{A.84}$$

and finally it turns that

$$\left[b_h^K\right]_\alpha = \Delta \int_{\hat{K}} \hat{p}_\alpha f\left(F_K(\hat{x})\right) d\hat{x} + \sum_{l=1}^{3} \chi_K^l \left|a_{l+1}^k - a_l^k\right| \int_0^1 \hat{p}_\alpha\left(\hat{\varphi}^l(\hat{\sigma})\right) g\left(\varphi_K^l(\hat{\sigma})\right), \tag{A.85}$$

where

$$\hat{\varphi}^1(\hat{\sigma}) = (\hat{\sigma}, 0), \quad \hat{\varphi}^2(\hat{\sigma}) = (1-\hat{\sigma}, \hat{\sigma}), \quad \hat{\varphi}^3(\hat{\sigma}) = (0, 1-\hat{\sigma}). \tag{A.86}$$

and χ_K^l is given by

$$\begin{array}{ll} 1 & \text{if } l \subset \Gamma_R \\ 0 & \text{otherwise.} \end{array} \tag{A.87}$$

The calculation of the integrals that appear in $\left[b_h^K\right]_\alpha$, as well as the ones that appear in the stiffness matrix when the coefficients a_{ij}, a_0 are not constant in the element K, are done via numerical integration.

There exists a mathematical theory that proves that the error does not increase if an appropriate formula which depends on the finite element space, is used.

If the dimension is $n = 2$ and triangular Lagrange elements of degree k are used (space $X_h^{(k)}$), the appropriate formulas would be:

$$\int_{\hat{K}} \psi d\hat{x} = \frac{1}{2}\left(\psi\left(1/2, 1/2\right)\right) \quad (\text{exact in } \wp_1) \tag{A.88}$$

$$\int_{\hat{K}} \psi d\hat{x} = \frac{1}{6}\left(\psi\left(0,0\right) + \psi\left(1,0\right) + \psi\left(0,1\right)\right) \quad (\text{exact in } \wp_1) \tag{A.89}$$

$$\int_{\hat{K}} \psi d\hat{x} = \frac{1}{6}\left(\psi(1/2,0) + \psi(1/2,1/2) + \psi(0,1/2)\right) \quad (\text{exact in } \wp_2) \tag{A.90}$$

Blocking of the degrees of freedom

The performed calculations correspond to a discretised problem with

$$V_h = X_h^{(1)} = \left\{ v_h \in C^0(\bar{\Omega}) v_{h|K} \in \wp_1(K) \quad \forall K \in \tau_h \right\}. \tag{A.91}$$

Therefore, the boundary conditions $u_{h/\Gamma_D} = 0$ has not been taken into account, which would be equivalent to replace it by $\frac{\partial u_h}{\partial \nu_A} = 0 \quad on\ \Gamma_D$.

The "error" arises because we have considered an space V_h bigger than it should be, more precisely, functions ϕ_i of the basis have been used that do not satisfied the boundary condition $\phi_{i|\Gamma_D} = 0$. To get rid off these function of the basis is equivalent to get rid off the correspondent unknowns and equations (degrees of freedom). This process turns to be unpleasant from the point of view of the programming. A simpler procedure would be to replace the i-th equation (assuming the node i belongs to Γ_D), that we are not in the right to include because this node does not provide any function to the basis, by the equation

$$\xi_i = 0. \tag{A.92}$$

Actually the i-th equation is replaced by one, "programming equivalent to $\xi_i = 0$", the one obtained by substituting the diagonal term $(A_h)_{ii}$ by a very big number, for example 10^{30}.

Appendix B
Solution of the System of Characteristics

In this Appendix we provide the solution of the system of characteristics for the Vasicek and Hull and White interest rate models. This is a system of ordinary differential equations that needs to be solved at every timestep in order to use time discretization with characteristics.

B.1 Vasicek model

In this Section we solve the system of characteristics for the Vasicek interest rate model.

B.1.1 Equation 1

$$\frac{d\phi_1}{d\tau}(\tau) - \gamma\phi_1(\tau) = \frac{1}{2}\rho\sigma w - \beta, \; \phi_1(\tau_{n+1}) = r \tag{B.1}$$

We define the integration factor,

$$\eta(\tau) = e^{-\gamma\tau}. \tag{B.2}$$

Multiplying both terms of $(B.1)$ by the integration factor we find

$$e^{-\gamma\tau}\frac{d\phi_1}{d\tau}(\tau) - e^{-\gamma\tau}\gamma\phi_1(\tau) = \left[\frac{1}{2}\rho\sigma w - \beta\right]e^{-\gamma\tau}, \tag{B.3}$$

$$\frac{d}{d\tau}\left[e^{-\gamma\tau}\phi_1(\tau)\right] = \left[\frac{1}{2}\rho\sigma w - \beta\right]e^{-\gamma\tau}, \tag{B.4}$$

Integration on both sides of the equation leads

$$e^{-\gamma\tau}\phi_1(\tau) = \int \left[\frac{1}{2}\rho\sigma w - \beta\right] e^{-\gamma\tau} d\tau, \tag{B.5}$$

$$e^{-\gamma\tau}\phi_1(\tau) = -\frac{1}{\gamma}\left[\frac{1}{2}\rho\sigma w - \beta\right] e^{-\gamma\tau} + C, \tag{B.6}$$

$$\phi_1(\tau) = -\frac{1}{\gamma}\left[\frac{1}{2}\rho\sigma w - \beta\right] + Ce^{\gamma\tau}. \tag{B.7}$$

Imposing the final condition at time τ_{n+1}

$$\phi_1(\tau_{n+1}) = r \Leftrightarrow r = -\frac{1}{\gamma}\left[\frac{1}{2}\rho\sigma w - \beta\right] + Ce^{\gamma\tau_{n+1}}, \tag{B.8}$$

we find the value of the constant

$$C = \left[r + \frac{1}{\gamma}\left[\frac{1}{2}\rho\sigma w - \beta\right]\right] e^{-\gamma\tau_{n+1}}, \tag{B.9}$$

and therefore, the solution at time τ_n is given by

$$\phi_1(\tau_n) = -\frac{1}{\gamma}\left[\frac{1}{2}\rho\sigma w - \beta\right] + \left[r + \frac{1}{\gamma}\left[\frac{1}{2}\rho\sigma w - \beta\right]\right] e^{-\gamma\Delta\tau}. \tag{B.10}$$

Defining

$$\delta = \frac{1}{\gamma}\left[\frac{1}{2}\rho\sigma w - \beta\right], \tag{B.11}$$

$$c = e^{-\gamma\Delta\tau}, \tag{B.12}$$

the solution can be written as

$$\phi_1(\tau_n) = -\delta + [r + \delta] e^{-\gamma\Delta\tau} = rc + \delta(c - 1). \tag{B.13}$$

B.1.2 Equation 2

$$\frac{d\phi_2}{d\tau}(\tau) = \left(\sigma^2 - \phi_1(\tau) + D_0\right)\phi_2(\tau), \tag{B.14}$$

Reorganizing things,

$$\frac{d\phi_2}{\phi_2}(\tau) = \left(\sigma^2 - \phi_1(\tau) + D_0\right), \tag{B.15}$$

integration on both sides of the equation leads

$$\phi_2(\tau) = C\exp\left[\left(\sigma^2 + D_0\right)\tau\right]\exp\left[-\int \phi_1(\tau)\,d\tau\right]. \tag{B.16}$$

Imposing the final condition at time τ_{n+1}

$$\phi_2(\tau_{n+1}) = S \Longleftrightarrow S = C\exp\left[\left(\sigma^2 + D_0\right)\tau_{n+1}\right]\exp\left[-\int \phi_1(\tau)\,d\tau|_{\tau=\tau_{n+1}}\right], \tag{B.17}$$

we find the value of the constant

$$C = S\exp\left[-\left(\sigma^2 + D_0\right)\tau_{n+1}\right]\exp\left[\int \phi_1(\tau)\,d\tau|_{\tau=\tau_{n+1}}\right], \tag{B.18}$$

and therefore the solution can be written as

$$\begin{aligned}\phi_2(\tau_n) \;=\; & S\exp\left[-\left(\sigma^2 + D_0\right)\tau_{n+1}\right]\exp\left[\int \phi_1(\tau)\,d\tau|_{\tau=\tau_{n+1}}\right]\exp\left[\left(\sigma^2 + D_0\right)\tau_n\right] \\ & \exp\left[-\int \phi_1(\tau)\,d\tau|_{\tau=\tau_n}\right],\end{aligned} \tag{B.19}$$

$$\phi_2(\tau_n) = S\exp\left[-\left(\sigma^2 + D_0\right)\Delta\tau\right]\exp\int_{\tau_n}^{\tau_{n+1}} \phi_1(\tau)\,d\tau. \tag{B.20}$$

Noting that

$$\int_{\tau_n}^{\tau_{n+1}} \phi_1(\tau)\,d\tau = \int_{\tau_n}^{\tau_{n+1}} \left[-\delta + (r+\delta)\,e^{\gamma(\tau-\tau_{n+1})}\right]d\tau = -\delta\Delta\tau + \frac{1}{\gamma}(r+\delta)\,c, \tag{B.21}$$

the solution becomes

$$\phi_2(\tau_n) = S\exp\left[-\left(\sigma^2 + D_0 + \delta\right)\Delta\tau\right]\exp\left[\frac{1}{\gamma}(r+\delta)(1-c)\right]. \tag{B.23}$$

Notice that expressions of $\phi_1(\tau_n)$ and $\phi_2(\tau_n)$ do not depend on τ_n , just on the time step $\tau_{n+1} - \tau_n = \Delta\tau$. This property together with the fact that the system of the trajectories is autonomous (coefficients do not depend on time) allows calculations to be done just once for all time steps, in the usual case where the time step is constant.

B.2 Hull and White Model

In this Section we solve the system of characteristics for the Hull and White interest rate model.

B.2.1 Equation 1

$$\frac{d\phi_1}{d\tau}(\tau) - \gamma\phi_1(\tau) = \frac{1}{2}\rho\sigma w - \beta(t), \quad \phi_1(\tau_{n+1}) = r. \tag{B.24}$$

We define the integration factor

$$\Psi(\tau) = e^{-\gamma\tau}. \tag{B.25}$$

Multiplying both terms of $(B.24)$ by the integration factor we find

$$e^{-\gamma\tau}\frac{d\phi_1}{d\tau}(\tau) - e^{-\gamma\tau}\gamma\phi_1(\tau) = \left[\frac{1}{2}\rho\sigma w - \beta(t)\right]e^{-\gamma\tau}, \tag{B.26}$$

$$\frac{d}{d\tau}\left[e^{-\gamma\tau}\phi_1(\tau)\right] = \left[\frac{1}{2}\rho\sigma w - \beta(t)\right]e^{-\gamma\tau}. \tag{B.27}$$

Integration on both sides of the equation leads,

$$e^{-\gamma\tau}\phi_1(\tau) = \int \left[\frac{1}{2}\rho\sigma w - \beta(t)\right] e^{-\gamma\tau} d\tau + C, \tag{B.28}$$

$$e^{-\gamma\tau}\phi_1(\tau) = \frac{1}{2}\rho\sigma w \int e^{-\gamma\tau} d\tau - \int e^{-\gamma\tau}\beta(t)\, d\tau + C, \tag{B.29}$$

$$\phi_1(\tau) = -\frac{1}{2\gamma}\rho\sigma w - e^{\gamma\tau}\int e^{-\gamma\tau}\beta(t)\, d\tau + e^{\gamma\tau}C, \tag{B.30}$$

Imposing the final condition at time τ_{n+1}

$$\phi_1(\tau_{n+1}) = r \Leftrightarrow r = -\frac{1}{2\gamma}\rho\sigma w - e^{\gamma\tau_{n+1}}\int e^{-\gamma\tau}\beta(t)\, d\tau\bigg|_{\tau_{n+1}} + e^{\gamma\tau_{n+1}}C, \tag{B.31}$$

we find the value of the constant

$$C = \left[r + \frac{1}{2\gamma}\rho\sigma w - e^{\gamma\tau_{n+1}}\int e^{-\gamma\tau}\beta(t)\, d\tau\bigg|_{\tau_{n+1}}\right] e^{-\gamma\tau_{n+1}}, \tag{B.32}$$

and therefore the solution can be written as

$$\begin{aligned}\phi_1(\tau_n) \quad = \quad & -\frac{1}{2\gamma}\rho\sigma w - e^{\gamma\tau_n}\int e^{-\gamma\tau}\beta(t)\, d\tau + \\ & e^{\gamma\tau_n}e^{-\gamma\tau_{n+1}}\left[r + \frac{1}{2\gamma}\rho\sigma w - e^{\gamma\tau_{n+1}}\int e^{-\gamma\tau}\beta(t)\, d\tau\bigg|_{\tau_{n+1}}\right],\end{aligned} \tag{B.33}$$

Defining

$$\delta = \frac{1}{2}\rho\sigma w, \tag{B.34}$$

equation B.33 becomes

$$\phi_1(\tau_n) = -\delta - e^{\gamma\tau_n}\int e^{-\gamma\tau}\beta(t)\, d\tau\bigg|_{\tau=\tau_n} + e^{-\gamma\Delta\tau}\left[r + \delta + e^{\gamma\tau_{n+1}}\int e^{-\gamma\tau}\beta(t)\, d\tau\bigg|_{\tau=\tau_{n+1}}\right], \tag{B.35}$$

$$\phi_1(\tau_n) = -\delta + e^{-\gamma\Delta t}[r+\delta] + - e^{\gamma\tau_n}\int e^{-\gamma\tau}\beta(t)\,d\tau\Big|_{\tau=\tau_n} + e^{\gamma\tau_n}\int e^{-\gamma\tau}\beta(t)\,d\tau\Big|_{\tau=\tau_{n+1}}, \tag{B.36}$$

$$\phi_1(\tau_n) = -\delta + e^{-\gamma\Delta t}[r+\delta] + e^{\gamma\tau_n}\int_{\tau_n}^{\tau_{n+1}} e^{-\gamma\tau}\beta(\tau)\,d\tau. \tag{B.37}$$

B.2.2 Equation 2

$$\frac{d\phi_2}{d\tau}(\tau) = \left(\sigma^2 - \phi_1(\tau) + D_0\right)\phi_2(\tau), \tag{B.38}$$

Reorganizing things,

$$\frac{d\phi_2}{\phi_2}(\tau) = \left(\sigma^2 - \phi_1(\tau) + D_0\right), \tag{B.39}$$

integration on both sides of the equation leads

$$\phi_2(\tau) = C\exp\left[\left(\sigma^2 + D_0\right)\tau\right]\exp\left[-\int\phi_1(\tau)\,d\tau\right], \tag{B.40}$$

Imposing the final condition at time τ_{n+1}

$$\phi_2(\tau_{n+1}) = S \Longleftrightarrow S = C\exp\left[\left(\sigma^2 + D_0\right)\tau_{n+1}\right]\exp\left[-\int\phi_1(\tau)\,d\tau|_{\tau=\tau_{n+1}}\right], \tag{B.41}$$

we find the value of the constant

$$C = S\exp\left[-\left(\sigma^2 + D_0\right)\tau_{n+1}\right]\exp\left[\int\phi_1(\tau)\,d\tau|_{\tau=\tau_{n+1}}\right], \tag{B.42}$$

and therefore the solution can be written as

$$\begin{aligned}\phi_2(\tau_n) &= S\exp\left[-\left(\sigma^2 + D_0\right)\tau_{n+1}\right]\exp\left[\int\phi_1(\tau)\,d\tau|_{\tau=\tau_{n+1}}\right]\exp\left[\left(\sigma^2 + D_0\right)\tau_n\right] \\ &\quad \exp\left[-\int\phi_1(\tau)\,d\tau|_{\tau=\tau_n}\right],\end{aligned} \tag{B.43}$$

or equivalently

$$\phi_2(\tau_n) = S\exp\left[-\left(\sigma^2 + D_0\right)\Delta\tau\right]\exp\int_{\tau_n}^{\tau_{n+1}}\phi_1(\tau)\,d\tau, \tag{B.44}$$

In order to compute $(B.37)$, $(B.44)$ numerically, we approximate the integrals to get

$$\phi_1(\tau_n) \approx \delta\left(e^{-\gamma\Delta\tau} - 1\right) + e^{-\gamma\Delta\tau}\left[r + \left(\beta(\tau_{n+1}) + e^{\gamma\Delta\tau}\beta(\tau_n)\right)\Delta\tau/2\right]. \tag{B.45}$$

$$\phi_2(\tau_n) \approx S\exp\left[-\left(\sigma^2 + D_0\right)\Delta\tau\right]\exp\left[\left(\phi_1(\tau_{n+1}) + \phi_1(\tau_n)\right)\Delta\tau/2\right]. \tag{B.46}$$

Notice that, contrary to the Vasicek model, expressions $\phi_1(\tau_n)$ and $\phi_2(\tau_n)$ do depend on τ_{n+1}, besides of the time step $\tau_{n+1} - \tau_n = \Delta\tau$. This requires calculations to be done for all time steps.