A mathematical study of convertible bonds

JOHAN DIMITRY
A mathematical study of convertible bonds

Johan Dimitry
jdimitry@kth.se

SA105X Degree Project in Mathematical Analysis, First level
Degree Progr. in Vehicle Engineering

Supervisor: Prof. Henrik Shahgholian

Department of Mathematics
School of Engineering Sciences
KTH, Royal Institute of Technology

Stockholm, Sweden

July 21, 2014
In memory of Yonan Faltaous, a sincere, God-fearing, loving and well-respected family man, a great engineer, a true source of inspiration for life and the very best uncle in the world. We love you and miss you more than it’s possible to explain by the use of the words offered by this world.
Acknowledgements

Foremost I would like to thank my supervisor Prof. Henrik Shahgholian for suggesting this subject and for supporting my work with invaluable advice and wise encouragement. I would also like to thank PhD. Sadna Sajadini for teaching me the mathematical theory of derivative pricing.

Secondly I want to show my love and gratitude to all those around me who have been my cornerstones, specially my wonderful mother Thereza. They’ve all been loving, tolerant and supportive during the rough times parallel to the writing of this thesis. Finally I want to thank God for giving me my daily strength that I need to keep struggling with the challenges that awaits me.
Abstract

A convertible bond (CB) is a financial derivative, a so called hybrid security. It is an issued contract from a company or a government, which is paid for up-front. The contract yields a known amount at the specified maturity date, unless the holder chooses to convert it into an amount of the underlying asset. This kind of financial products can have complex features affecting the contract price and the optimal exercising situation. The partial differential equation (PDE) approach used for pricing financial derivatives makes it possible to describe convertible bonds with a physical model, a reversed diffusion described by a parabolic PDE. One can sometimes find both analytical and numerical solutions for this type of PDEs and interpret the solutions from a financial point of view, as they suggest predictable behaviour of the contract price.

Keywords: Convertible Bonds, Financial Derivative, Complex Features, Diffusion, Parabolic Partial Differential Equation.
# Contents

1 **INTRODUCTION** ................................................. 1
   1.1 Limitations ............................................... 1
      1.1.1 Market Assumptions ............................. 1
   1.2 Notations ................................................ 2

2 **BOND PRICING AND INTEREST RATES MODELLING** ............... 2
   2.1 Bond pricing with known interest rates ................. 2
   2.2 Discrete coupon payments ................................ 4
   2.3 Stochastic interest rates modelling ..................... 5
   2.4 Bond pricing with stochastic interest rate ............ 5
      2.4.1 The market price of interest rate risk ......... 7

3 **CONVERTIBLE BONDS** ........................................ 7
   3.1 Call and put features ................................... 8
   3.2 Conversion affect on company market worth .......... 9
   3.3 Stochastic interest rate & asset price modelling .... 9
   3.4 CB pricing with two stochastic factors ............... 11
   3.5 Interest rate models ................................... 13

4 **ANALYTICAL SOLUTION** .................................... 14
   4.1 Results ................................................ 18
   4.2 Sensitivity analysis .................................... 19

5 **NUMERICAL SOLUTIONS** .................................... 21
   5.1 Discretization & derivative approximations .......... 22
      5.1.1 Domain discretization ........................... 22
      5.1.2 Discretization of Dirichlet & Initial conditions 23
      5.1.3 Crank-Nicolson finite difference approxima-
            tions ........................................... 23
      5.1.4 Algorithm ........................................ 25
   5.2 Results ................................................ 26
      5.2.1 Comparison with the analytical solution ....... 28
      5.2.2 Numerically computed Greeks ................... 29

6 **DISCUSSION AND CONCLUSION** .............................. 30

**Appendices** .................................................. 33
A Financial terminology and definitions 34

B Itō’s lemma 36

C Fundamental solution & Green’s function 38

D Concepts and results from probability theory 39
  1 Borel sigma algebra of sets 39
  2 Probability measure 40
  3 Probability space 41
  4 Random variable 41
  5 Stochastic processes 41
  6 Gaussian distribution 42

E Special functions and operations 44
  1 Dirac delta function 44
  2 Heaviside step function 45
  3 Convolution 45

F Algorithm: Crank-Nicolson Finite Difference Method 46

Bibliography 49
1 INTRODUCTION

Convertible bonds are hybrid securities issued from companies or governments to raise capital and up-front premiums, having properties of both equity and a fixed income at the same time. The CB is paid for up-front by the holder who receives the face value of the contract at the maturity date, unless the holder chooses to convert it into a pre-determined amount of the underlying asset, e.g. the issuing company’s stock. In this thesis both European and American CB contracts are brought up, where the American contracts does not force the holder to wait until maturity, the holder can choose to convert the CB into assets at any time up to and on maturity of the contract.

Financial contracts are, in general, mathematically studied for the purpose of pricing. This is due to the demand of good and reliable prediction models. In the last century, decision making in finance has gone from hunches and guessing to risk-calculated strategies with the help of applied mathematics. A PDE approach is used in this thesis to model the price of CB contracts in accordance with the well known Black and Scholes’ analysis. The analysis will consider both known and stochastic interest rate models, yielding one and two spatial dimensions for the PDEs in each case, these models are known as the one-factor model and the two-factor model respectively.

1.1 Limitations

There are many things that can be added to the mathematical models of derivative pricing, proposing that the subject of CB pricing is wide ranging and surely impossible to condense into this thesis. For instance one could include time-dependence or indeterministic behaviour of parameters such as the volatility, leading to multiple-factor models. Another way of complicating things could be by taking transaction costs into consideration or try to determine the optimal exercise situation for callable CBs, which is a very interesting free boundary problem.

This forces us to limit the spectrum of our mathematical analysis of CB pricing. We will investigate alternative interest rate models, call feature, put feature, dividend yields, coupon payments and stock price dilution as a consequence of conversion.

1.1.1 Market Assumptions

The analysis is carried out within the framework of the following assumptions and constraints.

i) The underlying asset is continuously traded.

ii) Arbitrage opportunities do not exist.
iii) Short selling is permitted and always available.

iv) The underlying asset is divisible and allowing for purchasing (selling) at any time, in other words there is always someone ready to sell (buy).

v) There are no transaction costs.

1.2 Notations

These are some of the notations used in this thesis:

\[ T := \mathbb{R} \cap (0, T] \] Time domain.

\[ \mathbb{R}^+ := \mathbb{R} \setminus (-\infty, 0) \] The set of positive real numbers.

\[ \Omega_B := \mathbb{R}^+ \times T \] Bond domain.

\[ \Omega := \mathbb{R}^+ \times \mathbb{R}^+ \times T \] Convertible bond domain.

\[ \psi \in C^2 (\Omega) \] Contract value.

\[ \Psi \in \mathbb{R}^+ \] Face value of the contract.

\[ S \in \mathbb{R}^+ \] Stock price.

\[ r \in \mathbb{R}^+ \] Risk-free interest rate.

\[ \epsilon \in \mathbb{R}^+ \] Conversion factor.

\[ \sigma \in \mathbb{R}^+ \] Volatility.

\[ T \in \mathbb{R}^+ \] Maturity date.

\[ \delta_s \in \mathbb{R}^+ \] Step in discrete space.

\[ \delta_r \in \mathbb{R}^+ \] Discrete time-step.

\[ \tilde{\Omega} := [0, S_{\text{max}}] \times [0, T] \] Discretized region of definition.

2 Bond pricing and interest rates modelling

To begin with we shall get familiar with bonds and bond pricing before moving on to the concept of convertible bonds. In this section we will go through the basic structures of these financial contracts.

A bond contract is paid for up-front and yields a specific amount, the face value \( \Psi \), at a specified maturity date \( T \). The contract may also have a feature which allows it to disburse a known cash dividend, coupons, at predetermined times during the life of the contract. Bonds without this feature are known as zero-coupon bonds. Bonds are normally issued by corporations or governments with the main purpose of raising capital and up-front premiums. This can be thought of as a loan to the issuers in exchange for the promised excess returns. Note that bonds don’t have an underlying asset like options and convertible bonds.

2.1 Bond pricing with known interest rates

The problem we address in the following analysis is to know how much an investor should pay today to earn a minimum of \( \Psi \) in \( T \) years of time. To
solve this problem we must find a way to determine the value of a bond contract.

We will start off by creating a pricing model that works for known interest rates and then we will use this analysis as a first building block when developing more advanced models. We shall define the time interval as

\[ T := \mathbb{R} \cap (0, T] \]  

(2.1)

The key parameters for our model are:

- The bond value: \( C^2 (T) \ni \psi : T \rightarrow \mathbb{R} \).
- Deterministic interest rate: \( C^1 (T) \ni r : T \rightarrow \mathbb{R} \).
- Coupon payments described by a real valued function, \( \kappa (t) \in L (T) \), that is non-negative and integrable, without the same demands on differentiability as for the two preceding functions.
- The constant final value \( \psi (T) = \Psi \).

Consider the infinitesimal relative change in bond price with respect to time. To prevent any arbitrage opportunities, the infinitesimal change must be equal to the return one would get from a bank deposit with the risk-free interest rate \( r (t) \). We can therefore express the relative change as

\[ \frac{d\psi}{\psi} = r (t) \, dt, \quad \forall t \in T. \]  

(2.2)

By taking coupon payments into consideration, the bond price dynamics becomes \( d\psi + \kappa dt = r \psi dt \), leading to the inhomogeneous final value problem

\[
\begin{aligned}
\psi' - r (t) \psi + \kappa (t) &= 0, \quad \forall t \in T, \\
\psi (T) &= \Psi.
\end{aligned}
\]  

(2.3)

This is a first order linear ODE\(^1\) which is easily solved by determining an integrating factor. The following time-function satisfies our final value problem in (2.3),

\[ \psi (t) = e^{-\int_t^T r (\tau) \, d\tau} \left( \Psi + \int_t^T \kappa (\varphi) e^{\int_\varphi^T r (\tau) \, d\tau} \, d\varphi \right). \]  

(2.4)

Without coupon payments the bond is valued by,

\[ \psi (t) = \Psi e^{-\int_t^T r (\tau) \, d\tau}. \]  

(2.5)

These are deterministic equations, thus if the bond price is determined today we’ll know the value \( \psi (t) \), which in turn gives us

\[ \frac{\int_t^T r (\tau) \, d\tau}{\Psi} = \ln \frac{\psi (t)}{\Psi}. \]  

(2.6)

\(^1\)Ordinary Differential Equation
Since \( \psi \in C^2(\mathcal{T}) \), equation (2.6) can be re-written as
\[
r(t) = -\frac{1}{\psi(t)} \frac{d\psi}{dt}, \quad \forall t \in \mathcal{T}.
\] (2.7)

If the market price of zero-coupon bonds is dependent on deterministic interest rates, then the future interest rate value is given by equation (2.7) which in turn suggests that a positive interest rate implies a decreasing bond value with respect to time,
\[
r > 0 \quad \Rightarrow \quad \frac{d\psi}{dt} < 0.
\]

A corollary to this is that the longer life-span a bond has the less it is worth today.

### 2.2 Discrete coupon payments

In real life coupons are discretely paid, which requires us to introduce the following adjustments to the model and thereby make it compatible with discrete payments. Let \( t_c \) be the time of coupon payment, this contributes to a discontinuity in the bond value. A so called jump condition for the bond value has to be determined. Let \( t_c^- \) be defined as the time instantaneously after the discrete coupon payment and \( t_c^+ \) after. Then the mathematical relation between the bond value before and after a coupon payment is
\[
\psi(t_c^-) = \psi(t_c^+) + \kappa. \tag{2.8}
\]

Thus the previous final value problem (2.3) becomes
\[
\begin{cases}
\psi' - r\psi + \kappa \delta(t - t_c) = 0, \\
\psi(T) = \Psi,
\end{cases} \tag{2.9}
\]

where \( \delta(\cdot) \) is the Dirac delta function\(^2\) used in a distributional sense. The problem in (2.9) is satisfied by
\[
\psi(t) = e^{-\int_t^T r(\tau)d\tau} \left( \Psi + \kappa \mathcal{U}(t_c - t) e^{\int_{t_c}^T r(\tau)d\tau} \right), \tag{2.10}
\]

where \( \mathcal{U}(\cdot) \) is the Heaviside step function\(^3\). If one considers an amount of \( \psi \) coupon payments of value \( \kappa_j \) at the discrete time \( t_j \), the ODE in (2.3) becomes
\[
\psi' - r\psi + \sum_j \kappa_j \delta(t - t_j) = 0, \quad \forall j \in [1, \vartheta] \cap \mathbb{Z}, \tag{2.11}
\]

which is satisfied by the function
\[
\psi(t) = e^{-\int_t^T r(\tau)d\tau} \left( \Psi + \sum_j \kappa_j \mathcal{U}(t_j - t) e^{\int_{t_j}^T r(\tau)d\tau} \right), \quad \forall j \in [1, \vartheta] \cap \mathbb{Z}. \tag{2.12}
\]

\(^2\)Please see Appendix E for definition.

\(^3\)Please see Appendix E for definition.
2.3 Stochastic interest rates modelling

We define the bond price domain as
\[ \Omega_B := \mathbb{R}^+ \times T, \]  
(2.13)
let the interest rate variation be modelled as a random walk generated by a stochastic process, more precisely a Wiener process\(^4\) that satisfies the following stochastic differential equation,
\[ dr = \omega(r,t) dX + \nu(r,t) dt, \quad \forall (r,t) \in \Omega_B. \]  
(2.14)
The function \( \omega \) describes the standard deviation and \( \nu \) is the drift of this process describing the average rate of growth. \( dX \) is sampled from a Wiener process with the random variables \( X(t) \in \mathcal{N}(0,t) \). Implying
\[ E\{dX\} = 0, \quad E\{dX^2\} = dt, \]
where \( E \) is a linear expectation operator, defined on a probability space \((\Gamma \ni \gamma, \mathcal{F}, \mathbb{P})\) as the Lebesgue integral
\[ E\{X\} := \int_{\Gamma} X(\gamma) d\mathbb{P}(\gamma), \]
where \( \mathbb{P} \) is the probability measure and \( \mathcal{F} \) is the Borel sigma algebra of subsets of the sample space \( \Gamma \) generated by the random variables. This concept will be implemented in the sequel by substitution for the preceding interest rate model.

2.4 Bond pricing with stochastic interest rate

Let the bond value be a real-valued function of time and the risk free interest rate, \( \psi(r,t) \in C^2(\Omega_B) \). Since bonds don’t have underlying assets, the only way of hedging is by combining bonds with different maturities in the same portfolio. Consider a portfolio long one bond and short an amount \( \Delta \) of another type of bond, this known as Delta-hedging
\[ \mathcal{H} := \psi_1 - \Delta \psi_2, \]  
(2.15)
with the infinitesimal change
\[ d\mathcal{H} = d\psi_1 - \Delta d\psi_2. \]  
(2.16)
In accordance with the calculations made by the authors of \([1]\) we can use Itô’s lemma\(^5\) and the no-arbitrage argument with help from equations (2.14) and (2.16) to arrive at
\[ d\mathcal{H} = \partial_t \psi_1 dt + \partial_r \psi_1 dr + \frac{1}{2} \omega^2 \partial_{rr} \psi_1 dt - \Delta \left( \partial_t \psi_2 dt + \partial_r \psi_2 dr + \frac{1}{2} \omega^2 \partial_{rr} \psi_2 dt \right). \]  
(2.17)

\(^4\) Please see Appendix D for definitions regarding stochastic processes.

\(^5\) For definition, please see Appendix B.
By choosing
\[ \Delta = \partial_r \psi_1 (\partial_r \psi_2)^{-1} \] (2.18)
we eliminate the portfolio risk and end up with
\[ dH = \left[ \partial_t \psi_1 + \frac{1}{2} \omega^2 \partial_{rr} \psi_1 - \partial_r \psi_1 (\partial_r \psi_2)^{-1} \left( \partial_t \psi_2 + \frac{1}{2} \omega^2 \partial_{rr} \psi_1 \right) \right] dt. \] (2.19)
The prevention of arbitrage suggests that the change in the portfolio’s return must be consistent with the return from the risk free interest rate, as one would get from a bank deposit, i.e.
\[ \frac{dH}{H} = r dt. \] (2.20)
By insertion of equations (2.15) and (2.19), equation (2.20) becomes
\[ (\partial_r \psi_1)^{-1} \left[ \partial_t \psi_1 + \frac{1}{2} \omega^2 \partial_{rr} \psi_1 - r \psi_1 \right] = (\partial_r \psi_2)^{-1} \left[ \partial_t \psi_2 + \frac{1}{2} \omega^2 \partial_{rr} \psi_2 - r \psi_2 \right]. \] (2.21)
Let us define the operator
\[ \mathcal{J}\{\psi_j\} := \left( \partial_r \psi_j \right)^{-1} \left[ \partial_t \psi_j + \frac{1}{2} \omega^2 \partial_{rr} \psi_j - r \psi_j \right], \quad \forall j \in \{1, 2\}. \] (2.22)
Equation (2.21) can only be true if and only if \( \mathcal{J}\{\psi_j\} = \mathcal{J}\{\psi\}, \forall j, \) since \( \psi_j \equiv 0, \) is a trivial solution for all \( j. \) Thus for some function \( a : \mathbb{R} \times T \to \mathbb{R} \) such that
\[ a (r, t) := \omega (r, t) \lambda (r, t) - \nu (r, t), \quad \omega \neq 0 \forall (r, t) \in \mathbb{R} \times T, \] (2.23)
where \( \lambda \) is often referred to as the market price of interest rate risk and is discussed in section 2.4.1. We can set
\[ a (r, t) = \mathcal{J}\{\psi\}, \] (2.24)
suggesting
\[ \omega (r, t) \lambda (r, t) - \nu (r, t) = (\partial_r \psi)^{-1} \left[ \partial_t \psi + \frac{1}{2} \omega^2 \partial_{rr} \psi - r \psi \right], \] (2.25)
leading to the zero-coupon bond pricing equation
\[ 0 = \partial_t \psi + \frac{1}{2} \omega^2 \partial_{rr} \psi + (\nu - \lambda \omega) \partial_r \psi - r \psi, \] (2.26)
The problem of equation (2.26) has the terminal value \( \psi (r, T) = \Psi \) and the Dirichlet conditions are dependent of \( \nu \) and \( \omega. \) By taking coupon payments into consideration this model transforms equation (2.26) into
\[ 0 = \partial_t \psi + \frac{1}{2} \omega^2 \partial_{rr} \psi + (\nu - \lambda \omega) \partial_r \psi - r \psi + \sum_j \kappa_j \delta (t - t_j). \] (2.27)
2.4.1 The market price of interest rate risk

Consider a portfolio consisting of one bond, then the infinitesimal change in the bond value per $dt$ is

$$d\psi = \omega \partial_r \psi dX + \left( \partial_t \psi + \frac{1}{2} \omega^2 \partial_{rr} \psi + \nu \partial_r \psi \right) dt.$$  

(2.28)

Inserting equation (2.26) in to (2.28) yields the relation

$$d\psi - r\psi dt = \omega \partial_r \psi (dX + \lambda dt),$$  

(2.29)

where we see that this is not a risk-less portfolio, due to the presence of $dX$. The right-hand side of equation (2.29) can be regarded as the excess return above the risk-free rate for accepting a certain level of risk. The portfolio profits an extra $\lambda dt$ per unit risk, $dX$, in reciprocity for taking that risk in the first place.

3 Convertible Bonds

The difference in definition of CBs and vanilla bonds is that the holder of the contract has the option of converting it to an amount of the underlying asset instead of receiving the face value. The conversion may take place at any time before the maturity if it’s an American contract, European contracts allow for conversion only on the maturity date. A CB may possess special features as coupon payments, call features and put features. A CB on an underlying asset with price $S$ returns $\Psi$ at the maturity date $T$, unless the holder chooses to convert it before maturity.

Analogously with the calculations made to find the Bond pricing equation, we can perform the Delta-hedging procedure and create a risk-free portfolio consisting of long one CB and short an amount $\Delta = \partial_s \psi$ of the underlying asset. This leads to the Black and Scholes partial differential inequality\(^6\) (for American CB contracts),

$$0 \geq \partial_t \psi + \frac{1}{2} \sigma^2 S^2 \partial_{ss} \psi + (r - D_0) S \partial_s \psi - r \psi + \sum_j \kappa_j \delta(t - t_j),$$  

(3.1)

$\forall j \in [1, \vartheta] \cap \mathbb{Z}$, where $D_0$ is the dividend rate. The inequality (3.1) accounts for coupon payments and dividend yields. The inequality sign is due to the fact that the holder can exercise the conversion right any time before maturity. Pricing models for European contracts have regular equal signs since it’s only allowed to convert at maturity.

\(^6\)Note that the inequality occurs when modelling the risk-free portfolio for American contracts. Normally to prevent arbitrage opportunities the portfolios return must equal the return of a risk-free investment, but the American portfolio is considered to be more risky and the return must therefore be less than or equal to a risk-free investment.
The CB may be converted into an amount $\epsilon \in \mathbb{Z}$ of the underlying asset, thus the following constraint must hold,

$$\psi \geq \epsilon S, \quad \forall t \in T.$$  \hfill (3.2)

We know that the final value is $\Psi = \psi (S, T)$, but the contract value just before maturity can be described by

$$\psi (S, T^-) = \max (\epsilon S, \Psi),$$ \hfill (3.3)

since the final value doesn’t satisfy the constraint. The Dirichlet conditions are

$$\lim_{S \to \infty} \psi (S, t) = \epsilon S, \hfill (3.4)$$

and

$$\psi (0, t) = \Psi e^{-r(T-t)}, \hfill (3.5)$$

without a major analysis we can see that it’s not optimal to exercise the contract when $S = 0$ as suggested by equation (3.5). For increasing coupon values $\kappa_j$, early exercise becomes less likely and vice versa. For the case where $\kappa_j = D_0 = 0$, the CB can be valued as a combination of cash and a European/American call option. Some times CBs may only be converted during specified periods of time, this is otherwise known as intermittent conversion. For the case of intermittent conversion the constraint needs to be satisfied only in this specific time interval. The existing dividend payments yields a free boundary, hence for sufficiently large $S$ the bond should be converted.

### 3.1 Call and put features

The call feature gives the issuer the right to purchase back the bond at any time or during specified periods. Thus this feature makes the bond less worth than it would have been without it. Suppose that the price the issuer will have to pay is $\Theta_1$, to eliminate any arbitrage opportunities

$$\psi \leq \Theta_1,$$ \hfill (3.6)

and the constraint (3.2) yields

$$\psi \in [\epsilon S, \Theta_1].$$ \hfill (3.7)

For $\epsilon S \geq \Theta_1$ the situation shouldn’t exist since a rational issuer/holder would’ve purchased/exercised before it came to this.

The put feature gives the holder a right to return the bond at any time. Thus such a feature increases the bond value relative to one without this feature. If the amount that the holder receives is $\Theta_2$ then

$$\psi \in [\Theta_2, \infty).$$ \hfill (3.8)

\footnote{Please see Appendix A for the definition of an option contract}
3.2 Conversion affect on company market worth

In practice, the existence of a CB actually affects the market worth of the company that issues the contract. This is due to the fact that if the contract holder decides to convert the CB into an amount $\epsilon \in \mathbb{Z}$ of shares of the underlying stock, then the company is forced to issue an amount $\epsilon$ new shares. Therefore, if we would assume that the company’s asset is worth $S$ and let $\epsilon_0$ be the number of existing shares before conversion, we get the following constraints needed to solve the CB pricing equation,

$$\psi \geq \frac{\epsilon S}{\epsilon + \epsilon_0},$$  \hspace{1cm} (3.9)

$$\psi \leq S,$$  \hspace{1cm} (3.10)

in symbiosis with the terminal condition (3.3). In their respective order these constraints tell us:

i) The CB price is bounded by its conversion value.

ii) If the CB would become too valuable the company is to declare bankruptcy.

What occurs is a dilution effect of the share price, with the dilution factor

$$\pi = \frac{\epsilon_0}{\epsilon + \epsilon_0},$$  \hspace{1cm} (3.11)

implying that

$$\psi \geq \pi S \frac{\epsilon}{\epsilon_0}.$$  \hspace{1cm} (3.12)

This dilution effect gives creates a so-called free boundary of the domain. If we look at the company’s total market worth, we realize it’s $S - \psi$, therefore the share price cannot be $S$ but instead

$$\frac{S - \psi}{\epsilon_0}.$$  \hspace{1cm} (3.13)

3.3 Stochastic interest rate & asset price modelling

To understand how we can estimate the movement of the underlying assets price one can turn to the so called efficient market hypothesis. In accordance with [1] the hypothesis suggests a totally random movement based on the arguments:

i) The asset price right now is a reflection of the past history without information regarding it.

ii) When new information reaches the markets an immediate response occurs.
In other words the changes in asset price are described by a Markov process. The changes we are interested in are not the absolute changes, instead the relative change of the asset price is of much more interest. To illustrate this, consider a movement of the stock price by $€2$, this information is insufficient due to the lack of knowledge about the actual share value. A $€2$ change of a stock worth $€400$ is insignificant in comparison with a $€2$ change of a $€8$ stock. Therefore, if we denote the asset price by $S$, then we want to model the infinitesimal relative change of the asset price with respect to the infinitesimal change in time, in other words

$$\frac{dS}{S} = f(\,dt\,),$$

for some real valued function $f$. We need the rate of return of the asset price, i.e. $f$, to be composed of two parts, a deterministic part and a non-deterministic part. In the same way as it’s done in [1] we let the deterministic part have an average rate of growth, $\mu$, which is also known as the drift. And the random part should be drawn from a normal distribution with the standard deviation of the return, $\sigma$, also known as the volatility.

Since the evolution of the interest rate is unpredictable, it’s natural to assume that it as well evolves in a stochastic manner. The risk-free interest rate and the asset price movement are both modelled as geometric Brownian motions respectively. Let us define the sets

$$\mathcal{T} := \mathbb{R}^+ \cap (0, T],$$

$$\Omega := \mathbb{R}^+ \times \mathbb{R}^+ \times \mathcal{T},$$

then the CB value is defined by the mapping

$$\psi : \mathbb{R}^3 \supset \Omega \rightarrow \mathbb{R} \mid \psi(S, r, t) \in C^2(\Omega).$$

We can now construct two random walks with the two stochastic differential equations

$$\frac{dS}{S} = \sigma dX_1 + \mu dt,$$

and

$$dr = \omega(r, t) \, dX_2 + \nu(r, t) \, dt.$$  

Where $dX_i, i \in \{1, 2\}$, are the Wiener processes with the random variables $X_i(t) \in \mathcal{N}(0, t), \forall i$. Which as before implies

$$\mathbb{E}\{dX_i\} = 0, \quad \mathbb{E}\{dX_i^2\} = dt, \forall i,$$

---

8 Please see Appendix D for the definition of a Markov process.
9 Please see Appendix A for the definition of drift.
10 Please see Appendix A for the definition of volatility.
11 Please see Appendix D for the definition of a Wiener process.
and the autocorrelation between \( dX_1 \) and \( dX_2 \) is given by

\[
E\{dX_1dX_2\} = \rho dt, \quad \rho : \Omega \to [-1, 1].
\] (3.17)

where \( E \) is the linear expectation operator. This allows us to establish the two rules

i) \( dX^2_i = dt, \ \forall i \),

ii) \( dX_1dX_2 = \rho dt \).

A simulation of two random walks with different parameter values, created with a simple MATLAB algorithm, is shown in Figure 3.1.

Figure 3.1: Simulation of two random walks for the stock price with \( \sigma = 0.1 \) and \( r = 0.04 \) for the blue line and \( \sigma = 0.5 \) and \( r = 0.12 \) for the green. For both processes we’ve set \( \mu = 0 \).

3.4 CB pricing with two stochastic factors

We will now construct portfolio consisting of long one CB with maturity \( T_1 \), short an amount \( \Delta_2 \) of a CB with maturity \( T_2 \) and short and amount \( \Delta_0 \) of the underlying asset.

\[
\mathcal{H} := \psi_1 - \Delta_2 \psi_2 - \Delta_0 S,
\] (3.18)
hence
\[
dH = d (\psi_1 - \Delta_2 \psi_2 - \Delta_0 S) = d\psi_1 - \Delta_2 d\psi_2 - \Delta_0 dS,
\] (3.19)
we need to determine the differentials \( d\psi_i \) before we can hedge this portfolio. The Taylor series expansion of \( \psi(S + dS, r + dr, t + dt) \) is defined as
\[
d\psi := \sum_j \frac{1}{j!} (dS \partial_s + dr \partial_r + dt \partial_t)^j \psi, \quad \forall j \in [0, \infty) \cap \mathbb{Z}.
\] (3.20)
Neglecting terms of higher order and insertion of equations (3.15), (3.16) and (3.17) into (3.20) in combination with Itô’s Lemma leads us to
\[
d\psi = \partial_t \psi dt + \partial_s \psi dS + \partial_r \psi dr + \frac{1}{2} (\sigma^2 S^2 \partial_{ss} \psi + 2\rho \sigma S \omega \partial_{sr} \psi + \omega^2 \partial_{rr} \psi) dt.
\] (3.21)
In other words, to make the relation (3.19) fully deterministic and thereby construct a risk free portfolio we have to choose
\[
\Delta_2 \equiv \partial_r \psi_1 (\partial_r \psi_2)^{-1},
\] (3.22)
and
\[
\Delta_0 \equiv \partial_s \psi_1 - \Delta_2 \partial_s \psi_2 = \partial_s \psi_1 - \partial_r \psi_1 (\partial_r \psi_2)^{-1} \partial_r \psi_2.
\] (3.23)
This will lead to exactly the same phenomena that occurs when deriving the pricing equation for simple bonds, we will be able to separate the terms containing \( \psi_1 \) from the ones containing \( \psi_2 \) on different sides of the equality sign, showing that one can get rid of the subscripts and formulate the CB pricing equation.
\[
\partial_t \psi + \frac{1}{2} (\sigma^2 S^2 \partial_{ss} \psi + 2\rho \sigma S \omega \partial_{sr} \psi + \omega^2 \partial_{rr} \psi) + \ldots
\]
\[
\ldots + rS \partial_s \psi + (\nu - \omega \lambda) \partial_r \psi - r \psi = 0,
\] (3.24)
where \( \lambda : \mathbb{R}^3 \supset \Omega \to \mathbb{R} \) is the market price of interest rate risk. We see that for \( \partial_s \psi = 0 \) this reduces to the simple bond pricing equation. Inclusion of dividend yields and coupon payments in the model transforms equation (3.24) into
\[
\partial_t \psi + \frac{1}{2} (\sigma^2 S^2 \partial_{ss} \psi + 2\rho \sigma S \omega \partial_{sr} \psi + \omega^2 \partial_{rr} \psi) + (r - D_0) S \partial_s \psi + \ldots
\]
\[
\ldots + (\nu - \omega \lambda) \partial_r \psi - r \psi + \sum_j \kappa_j \delta(t - t_j) = 0,
\] (3.25)
\( \forall j \in [1, \vartheta] \cap \mathbb{Z} \). Equations (3.24) and (3.25) are both parabolic PDEs describing a reversed process of diffusion in two spatial dimensions.

\[\text{12} \text{Please see Appendix B for definition of Itô’s Lemma.}\]
3.5 Interest rate models

The relation in (3.16) is incomplete due to the fact that we do not have an explicit formulation for the drift and the standard deviation functions. We can define the functions $\omega$ and $\nu$, from equation (3.16), in such a way that

$$\omega(r,t) := \sqrt{\alpha(t)r - \beta(t)}$$  \hspace{1cm} (3.26)

and

$$\nu(r,t) := \left(-\gamma(t)r + \eta(t) + \lambda(r,t)\sqrt{\alpha(t) - \beta(t)}\right).$$  \hspace{1cm} (3.27)

where we introduce the following restrictions to ensure the random walks’ economical properties:

i) Negative interest rates can be avoided by bounding the spot rate below with $\alpha > 0$ and $\beta \geq 0$, which yields a lower bound given by $\beta/\alpha$. But for the case when $\alpha = 0$ we need $\beta \leq 0$.

ii) For large $r$ we want the interest rate to decrease towards the mean and for small $r$ we want it to increase towards the mean. In combination with i), we get

$$\eta(t) \geq \frac{\beta(t)\gamma(t)}{\alpha(t)} + \frac{\alpha(t)}{2}.$$  

The following interest rate models are special cases of these definitions.

a) Vasicek model

This model was introduced by Oldrich Vasicek in the year 1977 and describes a stochastic interest rate evolution with the property of only being affected by the markets random movement and it is in the same time mean-reverting. For the Vasicek model, the parameters $\beta, \gamma, \eta$ are real constants and $\alpha = 0$, leading to

$$dr = \sqrt{-\beta}dX_2 + \left(-\gamma r + \eta + \lambda(r)\sqrt{-\beta}\right)dt, \ \forall \beta \leq 0.$$  \hspace{1cm} (3.28)

We can observe that when $\sqrt{-\beta} \equiv 0$, $r$ remains constant, and for $r < 0$ the drift becomes positive and thereby drive the interest rate up to the value of equilibrium.

b) Cox, Ingersoll & Ross model

The CIR model came to life 1985 and it is an extension of the Vasicek model, the main difference is that CIR does not allow negative interest rates, which is an advantage when pricing interest rate derivatives as Bonds. For the CIR model all the parameters are constant here as well, with $\beta = 0$,

$$dr = \sqrt{\alpha}dX_2 + \left(-\gamma r + \eta + \lambda(r)\sqrt{\alpha}\right)dt.$$  \hspace{1cm} (3.29)
Finally, the Hull & White model, which also extends the Vasicek model, assumes that the interest rate has a Gaussian distribution. Allowing for negative interest rate values (with significantly low probability) and has the property of leaning towards equilibrium. Note that for this model the parameters are all time dependent with $(\alpha = 0 \land \beta \neq 0) \lor (\alpha \neq 0 \land \beta = 0)$, giving us the relations

$$\left( dr = \sqrt{-\beta(t)} dX_2 + \left( -\gamma(t) r + \eta(t) + \lambda(r,t) \sqrt{-\beta(t)} \right) dt \right), \quad (3.30)$$

$$\lor$$

$$\left( dr = \sqrt{\alpha(t)} r dX_2 + \left( -\gamma(t) r + \eta(t) + \lambda(r,t) \sqrt{\alpha(t)} r \right) dt \right). \quad (3.31)$$

### 4 Analytical Solution

The lack of analytical solutions for the two stochastic factor model makes us focus on finding a solution for the European CB pricing equation. A European contract for CBs can be valued as a combination of a European call option and cash. This can be shown by solving the following Dirichlet problem for Black and Scholes PDE. For further simplicity the differential operator

$$A: \mathcal{D}_A \to L^2(\Omega) \mid A := \partial_t + \frac{1}{2} \sigma^2 S^2 \partial_{ss} + rS \partial_s - r, \quad (4.1)$$

is introduced, where the Hilbert space $L^2(\Omega)$ is the function space of all square integrable functions on $\Omega$ and the region of definition for $A$ is

$$\mathcal{D}_A = \{ \psi \in C^2(\Omega) \mid \lim_{S \to \infty} \frac{\psi}{S} = \epsilon, \ \psi(0,t) = \Psi e^{-r(T-t)},$$

$$\psi(S,T) = \max(\epsilon S - \Psi, 0) + \Psi \}. \quad (4.2)$$

The equation for a zero coupon CB with no dividend yield is now simply denoted by

$$A\psi = 0. \quad (4.3)$$

The parabolic problem in (4.3) can, with the help of appropriate coordinate transformations, be reduced to a forward one dimensional diffusion equation, which is a much simpler problem to address.

Let us introduce the transformations

$$\psi(S,t) = \Psi \phi \left( x = \ln \frac{S}{\Psi}, \ \tau = \frac{1}{2} \sigma^2 (T-t) \right), \quad (4.4)$$

where the interpretation of $\tau$ is the time left until the contracted maturity date, turning this into a forward PDE. This transformation also helps us get
A mathematical study of convertible bonds

To get rid of the non-constant coefficients, conveniently for every $S$-derivative we divide by one $S$ because of the logarithmic function. Note that $\phi$ and its parameter, $x$, are dimensionless, a normalization has been introduced due to this coordinate transformation. The problem becomes

$$\mathcal{A}'\phi = 0, \quad \forall (x, \tau) \in \mathbb{R} \times [0, \infty),$$

where

$$\mathcal{A}' := -\partial_x + \partial_{xx} + \left( \frac{2r}{\sigma^2} - 1 \right) \partial_x - \frac{2r}{\sigma^2}.$$  

What we would like to do now is to find a way to get rid of the first spatial derivative, $\partial_x$, and the constant operation, $2r/\sigma^2$. Let us consider the ansatz

$$\phi(x, \tau) = u(x, \tau) f(x, \tau),$$

where

$$f(x, \tau) := e^{\alpha x + \beta \tau}.$$  

This introduction changes the individual differential operators to

$$\partial_x \rightarrow \alpha f + f \partial_x,$$

$$\partial_{xx} \rightarrow \alpha^2 f + 2\alpha f \partial_x + f \partial_{xx}.$$  

It is now possible, after a couple of re-arrangements, to define the new operator

$$\mathcal{A}'' := -\partial_x + \partial_{xx} + \left( 2\alpha + \frac{2r}{\sigma^2} - 1 \right) \partial_x + \left( \alpha^2 + \alpha \left( \frac{2r}{\sigma^2} - 1 \right) - \frac{2r}{\sigma^2} - \beta \right).$$

It is obvious that if we want $\mathcal{A}'' \equiv (-\partial_x + \partial_{xx})$ the following system of equations must be satisfied since $\partial_x u \neq 0$ and $u \neq 0$.

$$0 = 2\alpha + \frac{2r}{\sigma^2} - 1,$$

$$0 = \alpha^2 + \alpha \left( \frac{2r}{\sigma^2} - 1 \right) - \frac{2r}{\sigma^2} - \beta,$$

these equations are solved by

$$\begin{cases} \alpha = -\frac{1}{2} \left( \frac{2r}{\sigma^2} - 1 \right), \\ \beta = -\frac{1}{4} \left( \frac{2r}{\sigma^2} + 1 \right)^2 . \end{cases}$$

In other words, $u = \phi e^{-(\alpha x + \beta \tau)}$ yields

$$\mathcal{A}'' u \equiv (-\partial_x + \partial_{xx}) u = 0,$$
and the initial condition, with $\kappa = 2r/\sigma^2$, is

$$u(x,0) = \left\{ \max(\epsilon e^x - 1, 0) + 1 \right\} e^{(\kappa - 1)/2}.$$  \tag{4.17}

Let us take a brief moment to discuss the Dirichlet boundary conditions we have described for our Dirichlet BVP in (4.3). If one adopts the original description of the problem without further analysis, one sees this as an ill-posed BVP because of the asymptotic boundary condition at infinity and as $S \to 0$. But the coordinate transformations have lead to a notable change concerning the boundaries,

$$x \to -\infty \quad \text{as} \quad S \to 0$$

and

$$x \to \infty \quad \text{as} \quad S \to \infty.$$  

This is actually good for our purpose, which is to show that this can be handled as a well-posed problem. The physical analogy with the financial problem lets us think of this as heat diffusing in an infinite rod, where the boundary conditions describes the temperature at $x = \pm \infty$. Physically speaking, the temperatures at infinity can never affect the heat diffusion pattern in the finite region and from an economical perspective the underlying asset price goes to infinity with probability zero. That is, equation (4.3) can be handled as a well-posed IVP. We can look at this as if one would take a blow-torch and concentrate the heat at $x = 0$, then the temperature-signal would spike at that point and as time moves along the heat will disperse to even out the temperature differences trying to reach thermal equilibrium. This makes it possible to solve the problem with the help of fundamental solutions$^{13}$ for the heat equation.

Our one dimensional initial value problem described in (4.16) has the Green’s function

$$G(x,\tau) = \frac{1}{2\sqrt{\pi \tau}} e^{-x^2/4\tau},$$  \tag{4.18}

thus a solution is given by convolution$^{14}$ of the two integrable functions (4.18) and (4.17)

$$u(x,\tau) = G \ast u(x,0) \equiv \int_{\mathbb{R}} G(\xi - x, \tau) u(\xi,0) \, d\xi.$$  \tag{4.19}

Before moving on, it’s important to know that we have to assume that $u$ is a smooth function, i.e. $u \in C^\infty(\mathbb{R})$, for this to work as a solution. The integral expression (4.19) is substantiated into,

$$u(x,\tau) = \frac{1}{2\sqrt{\pi \tau}} \int_{-\infty}^{\infty} e^{-(\xi-x)^2/4\tau} u(\xi,0) \, d\xi.$$  \tag{4.20}

$^{13}$Please see Appendix C for details on the concept of fundamental solutions.

$^{14}$Please see Appendix E for details regarding the convolution operation.
and by substituting for $\eta = (\xi - x) / \sqrt{2\tau}$ we arrive at

$$u(x, \tau) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\eta^2/2} u\left(x + \eta \sqrt{2\tau}, 0\right) d\eta. \quad (4.21)$$

Before we can evaluate (4.21) correctly we need to consider how the initial condition, $u(\cdot, 0)$, behaves with respect to $\eta$, so that we can know what the integrand is during the different intervals. The quest is to determine for what values of $\eta$ the relation in (4.22) satisfied.

$$e\epsilon^{(1-\alpha)(x+\eta\sqrt{2\tau})} - e^{-\alpha(x+\eta\sqrt{2\tau})} > 0. \quad (4.22)$$

This is true if and only if $\eta > - (\ln \epsilon + x) / \sqrt{2\tau}$, which is our point of discontinuity in the initial condition. The discontinuity we are referring to is actually a discontinuity in $\partial_x u$, leading to a drastic change in direction. When examining the curve of the CB value at maturity, which is seen in Figure 4.1, it is clear. Let us denote this special value by $\zeta$, then our integral in (4.21) becomes

$$u(x, \tau) = \frac{1}{\sqrt{2\pi}} \left\{ \epsilon \int_{-\infty}^{\zeta} e^{-\eta^2/2 + (1-\alpha)(x+\eta\sqrt{2\tau})} d\eta + \ldots \right\} \right.$$

$$\left. \ldots + \int_{-\infty}^{\zeta} e^{-\eta^2/2 - \alpha(x+\eta\sqrt{2\tau})} d\eta \right\}. \quad (4.23)$$

Figure 4.1: The value of a zero coupon CB with no dividend yield, at maturity.
After applying some re-constructional surgery to (4.23) we end up with the expression
\[
\begin{align*}
    u(x, \tau) &= \frac{e}{\sqrt{2\pi}} e^{\frac{\tau}{2} (\kappa + 1)^2/4 + x (\kappa + 1)/2} \int_{-\infty}^{\infty} e^{-\frac{1}{2} (\eta - \frac{1}{2} (\kappa + 1) \sqrt{2\tau})^2} d\eta + \ldots \\
    \ldots + \frac{1}{\sqrt{2\pi}} e^{\tau (\kappa - 1)^2/4 + x (\kappa - 1)/2} \int_{-\infty}^{\infty} e^{-\frac{1}{2} (\eta - \frac{1}{2} (\kappa - 1) \sqrt{2\tau})^2} d\eta,
\end{align*}
\]
(4.24)

If we define the two variables
\[
    \zeta_1 := \ln \epsilon + x \sqrt{2\tau} + \frac{1}{2} (\kappa + 1) \sqrt{2\tau},
\]
and
\[
    \zeta_2 := \ln \epsilon + x \sqrt{2\tau} + \frac{1}{2} (\kappa - 1) \sqrt{2\tau},
\]
then (4.24) reduces to
\[
    u(x, \tau) = e e^{\frac{\tau}{2} (\kappa + 1)^2/4 + x (\kappa + 1)/2} \Phi (\zeta_1) - e^{\frac{\tau}{2} (\kappa - 1)^2/4 + x (\kappa - 1)/2} \Phi (\zeta_2),
\]
(4.25)
where \(\Phi (\cdot)\) is the cumulative distribution function for the standard Gaussian distribution\(^{15}\).

All we have to do now is simply retrace all our transformation steps and end up with (4.26), the price of an European zero coupon CB with no dividend yield.
\[
    \psi (S, t) = \epsilon S \Phi (\zeta_1) - \Psi e^{-r(T-t)} \Phi (\zeta_2) + \Psi,
\]
(4.26)
where
\[
    \zeta_1 = \frac{\ln(\epsilon S/\Psi) + (r + \sigma^2/2) (T - t)}{\sigma \sqrt{T-t}}
\]
(4.27)
and
\[
    \zeta_2 = \frac{\ln(\epsilon S/\Psi) + (r - \sigma^2/2) (T - t)}{\sigma \sqrt{T-t}}.
\]
(4.28)

4.1 Results

The results of our calculations are plotted, the plots are generated by a MATLAB algorithm which is based on our analytical solution, equation (4.26). The input data used for the contract simulation is listed in table 4.1.

The three dimensional view of how the CB price behaves with respect to both time and the underlying stock price can be observed in Figure 4.2. The function is similar to, except for some significant details, the call option contract. In fact one can see that is a combination of a call option and the face value of the CB. This is more obvious when examining Figure 4.3 where we can see the solution curves for several time steps converging towards maturity.

\(^{15}\)Please see Appendix D for more details.
A mathematical study of convertible bonds

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma$</td>
<td>25 %</td>
</tr>
<tr>
<td>$\epsilon$</td>
<td>2</td>
</tr>
<tr>
<td>$\Psi$</td>
<td>10 Euros</td>
</tr>
<tr>
<td>$T$</td>
<td>5 years</td>
</tr>
<tr>
<td>$r$</td>
<td>10 %</td>
</tr>
</tbody>
</table>

Table 4.1: The parameter values used in the simulation of the analytical solution of the zero coupon CB with no dividend yield.

Figure 4.2: The CB value with respect to time and the underlying stock price, without coupons and dividends.

4.2 Sensitivity analysis

In practice, hedging of a portfolio is equivalent to the reduction of sensitivity, with respect to movement of underlying assets and other influencing parameters, by holding converse positions in a variety of financial instruments. A very efficient way of analysing the sensitivity of a portfolio as to the different parameters is to calculate the partial derivatives of the portfolio with respect to the parameters in question. There are five derivatives that are commonly spoken of in the context of sensitivity analysis in finance, they are known as the Greeks\(^{16}\).

\(^{16}\)These variables are denoted with an index $\pi$ to avoid confusion with other variables in this thesis with the same name.
Figure 4.3: Contours of the CB in Figure 4.2 with respect to the underlying asset for different time values converging towards maturity.

i) **Delta**

The Delta describes the sensitivity of the portfolio value with respect to the stock price,

\[ \Delta_\pi := \frac{\partial H}{\partial S}, \quad (4.29) \]

in the case of our European CB the portfolio is given by \( H = \psi - S \partial_s \psi \), in other words the Delta in our case is

\[ \Delta_\pi = -S \partial_{ss} \psi. \quad (4.30) \]

ii) **Gamma**

The Gamma is the curvature of the portfolio value function, thus yielding a more responsive hedging for small changes in the underlying asset.

\[ \Gamma_\pi := \frac{\partial^2 H}{\partial S^2} = \partial_s \Delta_\pi, \quad (4.31) \]

which implies

\[ \Gamma_\pi = -\partial_{ss} \psi - S \partial^4_s \psi, \quad (4.32) \]

for our portfolio.

iii) **Theta**

The Theta measures the time-value decay of the portfolio value,

\[ \Theta_\pi := -\frac{\partial H}{\partial t}, \quad (4.33) \]
A mathematical study of convertible bonds

taking this into advantage for our portfolio yields the relation

$$\Theta_\pi = -\partial_t \psi + S \partial_{st} \psi.$$  (4.34)

iv) **Vega**
The Vega measures the sensitivity to volatility,

$$\nu_\pi := \frac{\partial \mathcal{H}}{\partial \sigma}.$$  (4.35)

Thus we get

$$\nu_\pi = (1 - S) \partial_\sigma \psi - \partial_\sigma S.$$  (4.36)

v) **Rho**
Finally we have the Rho, which is the portfolio’s sensitivity with respect to the interest rate, defined by

$$\rho_\pi := \frac{\partial \mathcal{H}}{\partial r},$$  (4.37)

and becomes

$$\rho_\pi = \partial_r \psi - S \partial_{sr} \psi,$$  (4.38)

for our portfolio.

Hedging against these quantities, except for $\Delta_\pi$, requires that one combines more than one derivative and underlying assets into the portfolio. When performed properly it is possible to eliminate short term changes in the portfolio value. We will numerically compute and plot figures of some of the Greeks in section 5.2.2.

## 5 Numerical Solutions

In many ways it’s more efficient to solve pricing models for financial derivatives numerically than analytically, not to mention that for most of the models it’s the impossible to find an analytical solution. Consider the problem in (4.1), we will use this to show how this type of PDE problem can be solved numerically. The change of time-variable to $\tau = T - t$ is introduced, leading to $\tau$ being the time left to maturity and thereby giving us an initial condition for our PDE, instead of a terminal condition. This is done due to the complications that occur when implementing the upcoming numerical formulations into an algorithm. Explicitly speaking the PDE we are solving numerically is

$$\partial_\tau \psi = \frac{1}{2} \sigma^2 S^2 \partial_{ss} \psi + r S \partial_s \psi - r \psi.$$  (5.1)
5.1 Discretization & derivative approximations

Basically, the task is to solve Black and Scholes PDE for our European CB problem numerically. To do this we first have to let go of the concept of continuity and start working with discrete intervals and steps. This is done because of the computers inability of comprehending the concept of non-countability, for instance, the interval \([a, b] \cap \mathbb{R}\) is infinitely divisible and therefore would require an infinite amount of time for a computer (and a human being for that matter) to count all the points in between. The computer needs to be able to have exact points and reference points where it can perform calculations. One of the options we have is to define equidistant points between \(a\) and \(b\), where we allow the computer to execute computations, e.g. determining a function’s value at each one of the defined points. This allows us to approximate the function value over the whole interval with an arbitrary interpolation.

We will approximate the PDE in (5.1) by a recursive formula, a finite difference equation, whose solution will approximate the function’s value at each point of our domain. This is possible after we have discretized the domain, the variable stepping and the derivatives.

5.1.1 Domain discretization

Let the spatial interval be divided into \(N \in \mathbb{R}^+\) equally long sub-intervals, then the step-size will be defined as

\[
\delta_s := \frac{S_{\text{max}}}{N},
\]

where \(S_{\text{max}}\), chosen large enough, is a substitute for \(S = \infty\), which is the upper limit of our continuous domain \(\Omega\). The reason for this substitution is again the need to limit the amount of computations. An infinite interval will literally take forever to go through, regardless of available computing capacity. The time-step is defined as

\[
\delta_\tau := \frac{T}{L},
\]

where \(L \in \mathbb{R}^+\) is the number of equidistant time-steps. The discrete domain can now be introduced

\[
\tilde{\Omega} := [0, S_{\text{max}}] \times [0, T],
\]

the coordinates are

\[
S := n\delta_s, \quad \forall n \in [0, N] \cap \mathbb{Z},
\]

and

\[
\tau := l\delta_\tau, \quad \forall l \in [0, L] \cap \mathbb{Z}.
\]

From now on, we will use the notation \(\psi(n\delta_s, l\delta_\tau) = \psi_n^l\).
5.1.2 Discretization of Dirichlet & Initial conditions

The initial condition is originally given by the relation

\[ \psi(S, 0) = \max(\epsilon S, \Psi), \]  

which for our purpose will be re-written as

\[ f(n\delta_s) = \max(\epsilon n\delta_s, \Psi), \]  

in other words, the algorithm must be set up in such a way so that the computer will have to check for each \( n \) if \( n\delta_s < \Psi \) as long as \( l = 0 \). For the Dirichlet conditions we have for \( S = 0 \)

\[ \psi(0, \tau) = \Psi e^{-r\tau}, \]  

which is transformed to

\[ g(l\delta_\tau) = \Psi e^{-rl\delta_\tau}, \quad n = 0, \forall l \in [0, L] \cap \mathbb{Z}. \]  

From previous sections we have that for \( S \to \infty \)

\[ \psi \to \epsilon S, \]  

this will be re-written as

\[ h(l\delta_\tau) = \epsilon S_{\text{max}}, \quad n = N, \forall l \in [0, L] \cap \mathbb{Z}. \]  

5.1.3 Crank-Nicolson finite difference approximations

For the finite difference approximations we will use the Crank-Nicolson (CN) approximation since it is unconditionally stable and has good accuracy when iterating over a large number of time steps, as proven in [2]. The CN approximation is basically the average of implicit Euler and explicit Euler finite difference schemes. The following relations describe the CN approximations of our function and its derivatives. The function value is approximated by

\[ \psi \approx \frac{\psi^n + \psi^{n+1}}{2}, \]  

the time derivative by

\[ \partial_\tau \psi \approx \frac{\psi^{n+1} - \psi^n}{\delta_\tau}, \]  

the first order spatial derivative by

\[ \partial_s \psi \approx \frac{\psi^{l+1} - \psi^{l-1} + \psi^{l+1} - \psi^{l+1}}{4\delta_s}, \]  

\[ \text{For more on the subject of various finite difference schemes, we refer to the very pedagogic literature [2].} \]
and the second order spatial derivative by
\[
\partial_{ss}\psi \approx \frac{\psi_{n+1}^l - 2\psi_n^l + \psi_{n-1}^l + 2\psi_{n+1}^{l+1} + \psi_{n-1}^{l+1}}{2\delta_s^2}. \tag{5.16}
\]

By insertion of the relations (5.13) - (5.16) into our PDE in (5.1), we arrive at a finite difference equation which can be solved algebraically. After some simplifying re-arrangements we end up with
\[
\psi_n^{l+1} - \psi_n^l = \frac{1}{4} \sigma^2 n^2 \delta_r \left( \psi_n^{l+1} - 2\psi_n^l + \psi_{n-1}^l + 2\psi_{n+1}^{l+1} + \cdots \right) + \frac{1}{2} r \left( \psi_{n+1}^l - \psi_{n-1}^l + \psi_{n+1}^{l+1} - \psi_{n-1}^{l+1} \right) - \cdots - \frac{1}{2} r \left( \psi_n^l + \psi_{n+1}^{l+1} \right) . \tag{5.17}
\]

Note that all the \( \delta_s \) terms in (5.17) have been cancelled out by each other. Let us introduce the notations
\[
\mu_n := \frac{1}{4} \delta_r \left( nr - n^2 \sigma^2 \right), \tag{5.18}
\]
\[
\nu_n := \frac{1}{2} \delta_r \left( r + n^2 \sigma^2 \right) + 1, \tag{5.19}
\]
and
\[
\omega_n := \frac{1}{4} \delta_r \left( nr + n^2 \sigma^2 \right). \tag{5.20}
\]

After some manipulations and inclusion of the new variables (5.18), (5.19) and (5.20) into equation (5.17), we arrive at a finite difference equation (5.21) with separated time-steps on each side of the equal sign. The numerical molecule (stencil) for our scheme is visualized in Figure 5.1.

\[
\mu_n \psi_{n-1}^{l+1} + \nu_n \psi_n^{l+1} - \omega_n \psi_{n+1}^{l+1} = -\mu_n \psi_n^l + (1 - \nu_n) \psi_n^l + \omega_n \psi_{n+1}^l . \tag{5.21}
\]

Equation (5.21) has three unknowns and three known variables in each time-step, this is a linear system of equations that can be expressed as a matrix-vector relation.

\[
\begin{pmatrix}
\nu_0 & -\omega_0 & 0 & 0 & \cdots & 0 \\
\mu_1 & \nu_1 & -\omega_1 & 0 & \cdots & 0 \\
0 & \mu_2 & \nu_2 & -\omega_2 & \cdots & 0 \\
\vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\
0 & \cdots & 0 & \mu_N & \nu_N & 0
\end{pmatrix}
\begin{pmatrix}
\psi_0^{l+1} \\
\psi_1^{l+1} \\
\vdots \\
\psi_n^{l+1} \\
\vdots \\
\psi_N^{l+1}
\end{pmatrix}
= 
\begin{pmatrix}
X_0^l \\
X_1^l \\
\vdots \\
X_n^l \\
\vdots \\
X_N^l
\end{pmatrix} = A \in \mathbb{R}^{N \times N}
= u^{l+1} \in \mathbb{R}^N
= X^l \in \mathbb{R}^N. \tag{5.22}
\]
A mathematical study of convertible bonds

where \( A \) is tridiagonal. This is good from a computational point of view since tridiagonal systems can be solved in many computationally friendly ways without the need for Gauss-Jordan row operations, which indeed is a computationally expensive procedure.

\[
\psi_{l+1} - \psi_{l} + \psi_{l-1} = \psi_{l+1} - \psi_{l+1}
\]

Figure 5.1: A visualization of the stencil for the Black and Scholes PDE on the discrete space-time grid, where for each time-step we have three unknown variables and three previously determined variables.

5.1.4 Algorithm

The main purpose of this algorithm is to generate a matrix \( u \in \mathbb{R}^{N \times L} \) containing column vectors that are the solutions for each time-step, in other words every new step in time represents a new column in the matrix. The idea is to calculate the vector \( X_l \) and solve the tridiagonal system of equations in (5.22) for \( u^{l+1} \), in each iteration of \( l \). Since \( u^{l+1} \) is a matrix, we will jump to the next column every iteration of \( l \) and handle the calculations for the column vectors one by one. To be able to perform these computations for all index values, we have to manually assign the initial values for all elements of the first row in the matrix \( u \). Then the same goes for the Dirichlet conditions, which are implemented in the first and last column of \( u \).

This is easy to implement and compile with e.g. C/C++ and MATLAB, although C/C++ code computes with better performance than MATLAB when handling large sets of data. MATLAB is straightforward with good built-in functions when it comes to plotting the results and is for this reason our “weapon of choice”. For our purpose, it is not so important to choose the fastest for two reasons, firstly, the amount of computations in this case is not large enough for us to notice a significant difference in calculation time. Secondly, the computations are made off-line, i.e. it does not matter how
long it takes, in contrary to the algorithms working on the on-line "black-boxes" on wallstreet. The best choice would of course be a combination of both techniques, where we let C/C++ do the heavy work. The algorithm is written in pseudo-code and is available in Appendix F.

5.2 Results

The computational methods from the previous sections has been implemented into MATLAB to generate a three dimensional surface plot of a vanilla CB contract, in Figure 5.2. The input parameters are given in Table 5.1 The mesh is made up of approximately 900 vertices as each direction is divided into 30 sub-intervals.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma$</td>
<td>25 %</td>
</tr>
<tr>
<td>$\epsilon$</td>
<td>2</td>
</tr>
<tr>
<td>$\Psi$</td>
<td>10 Euros</td>
</tr>
<tr>
<td>$T$</td>
<td>5 years</td>
</tr>
<tr>
<td>$r$</td>
<td>10 %</td>
</tr>
</tbody>
</table>

Table 5.1: The parameter values used in the simulation of the analytical solution of the zero coupon CB with no dividend yield.

Figure 5.2: A European zero coupon CB contract with no dividend and the input data that is represented in Table 5.1.
A mathematical study of convertible bonds

The surface in Figure 5.2 is what we expected to see, a behaviour which is very close to the analytical solution. It’s easier to see some of the differences when examining the contour plot of the numerical solution, Figure 5.3, and then comparing it to Figure 4.3. Exactly how alike, or different, they are will be examined in section 5.2.1.

![Contour plot of numerically computed CB in Figure 5.2.](image)

Figure 5.3: Contour plot of numerically computed CB in Figure 5.2.

From the point of view of Figure 5.3, the CB characteristics are indeed visible. The curves suggest that the CB’s call option part has negative values for when the stock is lower than the face value, which is unrealistic,
but the total CB value is always positive. We can use the results from the analysis of discrete coupon payments, section 2.2, in combination with the jump condition (2.8) and include a dividend yield in the numerical model to arrive at a surface plot, Figure 5.4, for a CB contract with annually paid coupons and continuous dividend payments.

5.2.1 Comparison with the analytical solution

To compare the numerical and the analytical solutions with each other we will firstly use the input parameters from Table 5.1, secondly, we shall define a quantity for the relative difference in function value. Let this quantity be denoted by \( \chi \), the analytical function \( \psi \) and the numerical \( \phi \), then the following relation is our definition of this quantity,

\[
\chi (S,t) := \frac{\phi (S,t) - \psi (S,t)}{\psi (S,t)} , \quad \forall (S,t) \in \tilde{\Omega} .
\] (5.23)

The function \( \chi (S,t) \) describes the relative deviation of the numerical solution with respect to the analytical solution. Both solution surfaces have the predicted incline in value just after the stock price exceeds face value of the contract, but the main differences are for low stock price and long time before maturity, as seen in Figure 5.5.

![Relative difference: Analytical - Numerical](image)

Figure 5.5: The relative difference between the analytical and the numerical solutions, more precisely, this is the numerical solution’s relative deviation from the analytical solution surface.
This can definitely be interpreted as the relative error of our numerical analysis. Normally when computing a relative error one looks at the absolute value of the fraction in equation (5.23), but in our case we are just as interested in the direction (positive/negative) of the deviation as we are in the deviation itself. Negative values of $\chi(S,t)$ indicates that the numerical function has values smaller than the analytical, and vice versa for positive values of $\chi(S,t)$.

5.2.2 Numerically computed Greeks

We have chosen to investigate three of the five sensitivities mentioned in section 4.2, namely the Delta, the Rho and the Vega. The Delta, whose surface plot is seen in Figure 5.6, is the sensitivity with respect to the stock price.

![Numerically computed Delta](image)

Figure 5.6: Delta, the sensitivity with respect to the stock price. It has the shape of a call option’s Delta.

The Delta should be exactly the same\textsuperscript{18} as the Delta for a call option with exercise price $\Psi$, this true since it’s the derivative with respect to the underlying asset and thus leading to the elimination of the constant $\Psi$ in equation (4.26), which is the only difference from a call option. In other words, delta hedging in practice for CBs is no different than for call options.

In Figure 5.6 we can observe a clear behavioural difference depending on if the call option part of the CB is at the money, in of the money or when it’s out of the money\textsuperscript{19}. The portfolio’s sensitivity seems to rise along the line $t = S/\epsilon$, indicating that it’s is more sensitive to change in the stock price the further away we are from the maturity. This could be interpreted

---

\textsuperscript{18}With exception for the conversion rate factor, this has to be taken into consideration when comparing the Deltas.

\textsuperscript{19}See Appendix A for clarification regarding the at, in and out of the money jargon.
as if a change in asset price long before maturity has a larger probability of "messing things up" since it has a long time to move randomly and is for this reason a more risky change. For instance, say that the modelled Brownian motion of the asset price only have two steps until time reaches maturity. Then it is very unlikely that it ends up in a region far away from it’s starting point, in comparison with a stochastic process that has 15 steps left until time reaches maturity.

Figure 5.7: Rho, the sensitivity with respect to the interest rate.

The Rho surface in Figure 5.7 shows the portfolios sensitivity to change in interest rate value. We see that it’s almost not relying on the change of stock price at all, it seems to be entirely dependent on time and gets more sensitive the closer we get to maturity.

For the portfolios sensitivity to change in the volatility we shall turn our attention to the Vega surface in Figure 5.8. There is a clear region around the face value of the CB contract, specially near maturity, where we have great sensitivity with respect to the volatility of the underlying asset. This is due to the fact that a small fluctuation of the stock price that close to maturity and the face value could turn the game around for the call option part of the CB contract.

6 Discussion and conclusion

The whole study is composed of three main parts. Namely the creation of the foundation for a pricing model in accordance with Black and Scholes analysis, with various add-ons to simulate real situations and features, mathematically. Secondly we derived an analytical solution to the PDE problem of a one factor model of the CB contract. Finally, we showed how it is pos-
A mathematical study of convertible bonds

Figure 5.8: Vega, the sensitivity with respect to the volatility. Peaks around maturity when the stock price is near the face value of the CB.

It is possible to transform a PDE problem into a suitable form for computational methods and then solve it numerically.

The analysis that we performed in sections 2 and 3, where based on the assumptions made in section 1.1.1. These compromises of reality will result in a series of complications when trying to implement the models in practice. The two most imperative assumption, that we believe would have the largest affect, are the lack of transaction costs and the incredible trading availability. When it comes to the transaction costs, consider the Delta hedging procedure (which is a must in Black and Scholes analysis), one realizes that Delta hedging is impractical in real life as long as transaction costs exist. This is due to the fact that for Delta hedging to work, the portfolio has to trade the underlying asset in the same rate its changing, i.e. there will be many transactions leading to the cost of this procedure to grow larger than the intrinsic value of the contract itself. Again with focus on the Delta hedging, in reality there is not always someone who wants to purchase or sell when the portfolio managers want to. Thus leading to another complication with our model of the portfolio, we can’t hedge properly if it’s not possible to trade continuously at any time.

What seems to be of greater significance and should have been included in our pricing model, is non-constant volatility, or more precisely stochastic volatility. The reason why we chose to bring this up is because the volatility, according to [1], is the most difficult parameter to estimate. Estimation of the volatility as a random walk would lead to an extra factor and one would be able to create three factor models which are numerically solvable. A better modelling strategy could be to let the underlying asset, the volatility
and the interest rate be stochastic and then let some of the other parameters have time dependence. Where the time dependence of the parameters should be satisfying some predetermined statistical relation to the underlying asset, the interest rate and the volatility.

The analytical solution, in section 4, was actually a way around a much faster and easier solution, which is to take the Black and Scholes formulae for a call option with strike price $\Psi$, which is the same as the face value of the bond part of the contract, and insert the conversion factor $\epsilon$ to finally add the face value of the bond thus leading to equation (4.26). But this would have defeated the purpose of the analysis, we wanted to show how one could solve the famous Black and Scholes PDE analytically and thereby also reveal how the Black and Scholes formulae is derived, specially for non-standard cases as ours. The reason for only solving the one factor model is simple, there are no closed form analytical solutions for models of this kind with several spatial dimensions. The only existing solutions are either numerical or approximative analytical solutions (which are beyond the scope of this thesis).

The numerical methods used, in section 5, are applicable for two factor models as well, we chose to show the implementation on the one factor model to be able to compare the results with our analytical solution. The idea is the same when it comes to discretizing and creating finite difference equations, the only difference is that in each time step solves the discrete Laplace operator. When it comes to plotting it’s also a different approach, since the function maps three dimensions onto a fourth, it is only possible to look at snap shots of each time value\(^{20}\).

Whatever we are; financial engineers, mathematicians, economists or just vehicle engineering students with an interest for financial mathematics, we have to remember that the financial markets are complex adaptive systems relying on unpredictable and irrational human behaviour.

\(^{20}\)Unless of course one chooses to solve for each $S$-step or $r$-step instead of time, then each snap shot will have constant $S$ or $r$ respectively.
Appendix A

Financial terminology and definitions

**Definition A.1. (Risk-free interest rate)** A theoretical value of the rate of return one would get from a risk-free investment. Since there is no such thing as a risk-free investment in reality one has to have a reference point. In Sweden one often refer to the rate of return of a government bond over a 3-12 month period of time.

**Definition A.2. (Drift)** This is the instantaneous trend, a mean value of each step in the random walk. One can look at this as the resultant direction of a *drifting* boat in volatile waters.

**Definition A.3. (Volatility)** The volatility is another name for what’s otherwise known as the standard deviation, it’s defined as

\[
\sigma := \sqrt{\frac{1}{N} \sum_{j=1}^{N} (x_j - \mu)^2},
\]  

(A.0.1)

for discrete data points, where \( \mu \) is the mean value. This value describes the deviating tendencies of the data points. From a financial point of view the volatility of the stock price describes the deviating tendencies with respect to the drift.

**Remark A.1.** The standard deviation and the mean value can be considered for random variables where the mean value is calculated as the expected value, \( \mu = \mathbb{E}[X] \). The standard deviation is the positive square root of the variance, \( \sigma = \sqrt{\mathbb{E}[(X - \mu)^2]} \).

**Definition A.4. (Financial market)** A financial market is a market in which different parties can trade financial securities, commodities, and other items such as shares and currencies. The trading is done at relatively low transaction costs and at prices affected by supply, demand and speculations.
Definition A.5. (Financial derivatives) A derivative is a contract between two parties, specifying payment conditions regarding contractual obligations, special dates, significant values of the underlying assets and the notional value. The contract derives its value from the behaviour of the underlying assets performance. The underlying asset could be stock, commodities, interest rates, currencies, bonds and even other derivatives.

Definition A.6. (Options) An option is a contract between two parties, the issuer and the holder. This contract gives the holder the opportunity to exercise a right, the right is to either buy or sell a specified asset at a specified time, the exercise date, for a predetermined price, the exercise price. If it’s a put option, it gives you the right to sell the asset, the call option gives you buying rights.

Definition A.7. (Stock) The stock capital of a company is the liquidity available for trading, divided into shares. The share value is dependent of the company’s total share capital (the liquidity available for trading). All the shares in a stock belonging to a company has the same value. Investors can invest in a company by buying its shares, this makes them shareholders and hence owners of a proportional part of the company. The stock value is also affected by the investor’s speculations, the company’s losses/profits and change in the amount of shares in the stock. For instance, let the total share capital in the stock be $S_c$ and the number of shares represented by $\epsilon_0$. Then the share price $S$, is represented by the relation

$$S = \frac{S_c}{\epsilon_0}, \quad \forall \epsilon_0 \in \mathbb{Z} \cap \mathbb{R}^+.$$  \hspace{1cm} (A.0.2)

Definition A.8. (Dividend) When a company makes a profit it distributes it evenly between the shares, this is called dividend. Dividend of course affects the share value and thus generates a positive return for any investor. It’s common to speak of dividend rate, this value is a percentage of the current stock price.

Definition A.9. (Moneyness) A concept to classify the relative position of the underlying asset with respect to a financial derivative’s intrinsic value. There are three positions;

   i) In the money: The derivative would profit if exercised right now.

   ii) At the money: The exercise price of the derivative is equal to the underlying assets worth.

   iii) Out of the money: Exercising would result in a loss right now. There is no intrinsic value.
Appendix B

Itō’s lemma

Lemma B.1. Let $X(t)$ be a Wiener process and $\zeta(t)$ be a stochastic process following a geometric Brownian motion with drift $\mu$ and volatility $\sigma$ such that the stochastic differential equation

$$d\zeta = \zeta(\mu dt + \sigma dX), \quad (B.0.3)$$

is satisfied. Then for any function $\psi(\zeta,t) \in C^2(\mathbb{R}^2)$, the differential is defined by

$$d\psi := \sigma \zeta \frac{\partial \psi}{\partial \zeta} dX + \left( \mu \zeta \frac{\partial \psi}{\partial \zeta} + \frac{1}{2} \sigma^2 \zeta^2 \frac{\partial^2 \psi}{\partial \zeta^2} + \frac{\partial \psi}{\partial t} \right) dt. \quad (B.0.4)$$

An informal derivation of the lemma

To derive Itō’s lemma we establish that

i) $dX^2 \to dt$ as $dt \to 0$,

ii) and $dX \approx O(\sqrt{dt})$

The Taylor series expansion of $\psi(\zeta + d\zeta, t + dt)$ is

$$d\psi := \sum_j \frac{1}{j!} (d\zeta \partial_\zeta + dt \partial_t)^j \psi, \quad \forall j \in [0, \infty) \cap \mathbb{Z}. \quad (B.0.5)$$

Neglecting derivatives with lower order terms in equation (B.0.5) yields

$$d\psi = \frac{\partial \psi}{\partial \zeta} d\zeta + \frac{\partial \psi}{\partial t} dt + \frac{1}{2} \frac{\partial^2 \psi}{\partial \zeta^2} d\zeta^2, \quad (B.0.6)$$

where we can substitute for equation (B.0.3) into (B.0.6) to arrive at

$$d\psi = \frac{\partial \psi}{\partial \zeta} \zeta(\mu dt + \sigma dX) + \frac{1}{2} \frac{\partial^2 \psi}{\partial \zeta^2} \zeta^2 (\mu dt + \sigma dX)^2 + \frac{\partial \psi}{\partial t} dt. \quad (B.0.7)$$
A mathematical study of convertible bonds

With help from our established rules and some simple algebraic manoeuvres we are able to formulate

\[ d\psi = \sigma \zeta \frac{\partial \psi}{\partial \zeta} dX + \left( \mu \zeta \frac{\partial \psi}{\partial \zeta} + \frac{1}{2} \sigma^2 \zeta^2 \frac{\partial^2 \psi}{\partial \zeta^2} + \frac{\partial \psi}{\partial t} \right) dt, \quad (B.0.8) \]

which is Itô’s Lemma on differential form.
Appendix C

Fundamental solution &
Green’s function

**Definition C.1. (Fundamental solution)** Also known as the *heat kernel*, is a general solution of the inhomogeneous parabolic PDE, $L\psi = \delta (x)$ on $\Omega$, without concern for the values on $\partial \Omega$, only initial values are needed. It’s applicable for problems on $\mathbb{R}^n$.

**Definition C.2. (Green’s function)** A Green’s function is a fundamental solution with consideration for boundary conditions.

**Remark C.1.** If there exists a fundamental solution/Green’s function, i.e. a PDE where the inhomogeneous part is a Dirac pulse, then the solution of the same problem with an arbitrary inhomogeneous part is achieved by convolution.
Appendix D

Concepts and results from probability theory

1 Borel sigma algebra of sets

Definition D.1. (Sigma algebra) Let \( \Gamma \) denote a universal set and \( A \in \Gamma \). A collection \( \mathcal{F} \) of subsets of \( \Gamma \) is called a sigma algebra or equivalently a sigma field if it satisfies the conditions:

i) Non-emptiness; \( \Gamma \in \mathcal{F} \),

ii) closure under complement; if \( A \in \mathcal{F} \), then \( A^c \in \mathcal{F} \),

iii) closure under countable union; if \( A_j \in \mathcal{F} \), \( \forall j \) in a countable collection \( (A_j)_{j \in [0,\infty) \cap \mathbb{Z}} \), then

\[
\bigcup_{j \in [0,\infty) \cap \mathbb{Z}} A_j \in \mathcal{F}.
\]

Definition D.2. (Borel sigma algebra) The Borel sigma algebra \( \mathcal{B} \) is the sigma algebra generated by all intervals on the form \((\alpha,\beta)\).

Theorem D.1. The Borel sigma algebra \( \mathcal{B} \) is the sigma algebra generated by

i) open intervals, \((\alpha,\beta)\),

ii) half-open intervals, \((\alpha,\beta]\) and \([\alpha,\beta)\),

iii) closed intervals, \([\alpha,\beta]\),

iv) left and right intervals, \((-\infty,\beta]\) and \([\alpha,\infty)\),

v) open subsets of \( \mathbb{R} \),

vi) closed subsets of \( \mathbb{R} \).
Where both $\alpha$ and $\beta$ are real numbers.

**Proof.** This comes easily from the previous definitions D.1 and D.2. The following calculations yield a sufficient basis for proving all the suggestions of Theorem D.1. Let $\alpha = -\xi$ with $\xi \in \mathbb{R}^+$, then $(-\xi, \beta) \in \mathcal{B}$ and

$$( -\xi, \beta ) \subset (- (\xi + 1), \beta),$$

and therefore

$$\lim_{\xi \to \infty} ( -\xi, \beta ) = (-\infty, \beta).$$

Since all of the sets $(\alpha, \xi)$ are in $\mathcal{B}$,

$$\lim_{\xi \to \infty} (\alpha, \xi) = (\alpha, \infty),$$

the closure under complement yields $(-\infty, \alpha] \in \mathcal{B}$ since $(-\infty, \alpha] = (\alpha, \infty)^c$.

Now for the half-open and closed intervals, in accordance with definition D.1, we have closure under countable union, in other words if $(\alpha, \beta) \in \mathcal{B}$ and $[\alpha, \beta) \in \mathcal{B}$ then $(\alpha, \beta] \cup [\alpha, \beta) \in \mathcal{B}$. Since $(\alpha, \beta) = (\alpha, \beta) \cup \{\beta\}$ and $(\alpha, \beta] = (\alpha, \beta) \cup \{\alpha\}$ then $[\alpha, \beta] = (\alpha, \beta) \cup \{\beta\} \cup \{\alpha\} \in \mathcal{B}$. $\square$

## 2 Probability measure

**Definition D.3. (Measure)** The measure over $\mathcal{F}$ is a set function

$$\mu : \mathcal{F} \to \mathbb{R}^+, \quad \text{ (D.2.1)}$$

such that the following conditions are satisfied.

i) $\mu (A) \geq 0 \ \forall A \in \mathcal{F}$,

ii) countable additivity; if $A_j \in \mathcal{F}, \ \forall A_j \in (A_j)_{j \in [0, \infty) \cap \mathbb{Z}}$ which is a collection of pairwise disjoint sets, then

$$\mu \left( \bigcup_{j \in [0, \infty) \cap \mathbb{Z}} A_j \right) = \sum_{j \in [0, \infty) \cap \mathbb{Z}} \mu (A_j).$$

**Definition D.4. (Probability measure)** The probability measure denoted by $\mathbb{P}$ is a measure with the property

$$\mathbb{P} (\Gamma) = 1, \quad \text{ (D.2.2)}$$

in other words $\mathbb{P} : \mathcal{F} \to [0, 1]$. 

40
3 Probability space

Definition D.5. (Probability space) Let $\Gamma$ be the sample space containing all the possible outcomes $\gamma$, let $\mathcal{F}$ be the Borel sigma algebra of subsets of $\Gamma$, i.e. the possible events, and let $\mathbb{P}$ be the probability measure. Then a probability space is defined by the triple $(\Gamma, \mathcal{F}, \mathbb{P})$.

4 Random variable

Definition D.6. (Borel function) Let $\mathcal{F}$ be the Borel sigma algebra. If we have a function $\psi : \mathbb{R} \rightarrow \mathbb{R}$ such that for every set $S \in \mathcal{F}$ we have

$$\psi^{-1}(S) = \{ z \in \mathbb{R} \mid \psi(z) \in S \} \in \mathcal{F},$$

(D.4.1)

it’s known as a Borel function.

Remark D.1. Note that almost all functions are Borel functions, it is really hard to construct non-Borel functions or even non-Borel sets.

Definition D.7. (Random variable) A random variable $X$ is a Borel function, $X : \Gamma \rightarrow \mathbb{R}$, such that $\forall S \in \mathcal{F}$ we have

$$X^{-1}(S) = \{ \gamma \in \Gamma \mid X(\gamma) \in S \} \in \mathcal{F}.$$  

(D.4.2)

Theorem D.2. Let $\psi : \mathbb{R} \rightarrow \mathbb{R}$ be a Borel function and $X$ be a random variable. Then

$$Y = \psi(X),$$

(D.4.3)

is also a random variable.

Proof. Let the set $S \in \mathcal{F}$ and consider

$$Y^{-1}(S) = \{ \gamma \in \Gamma \mid Y(\gamma) \in S \}.$$

In accordance with the theorem we have $Y(\gamma) = \psi(X(\gamma))$, implying

$$Y^{-1}(S) = \{ \gamma \in \Gamma \mid \psi(X(\gamma)) \in S \} = \{ \gamma \in \Gamma \mid X(\gamma) \in \psi^{-1}(S) \}.$$

By definition $\psi^{-1}(S) \in \mathcal{F}$ since $S \in \mathcal{F}$ which in turn means $Y^{-1}(S) \in \mathcal{F}$ for any $S \in \mathcal{F}$, in other words $Y$ is a random variable.

5 Stochastic processes

Definition D.8. (Stochastic process) A stochastic process is a sequence of random variables $\{X(t)\}_{t \in T}$ defined on the probability space $(\Gamma, \mathcal{F}, \mathbb{P})$, where $T \subseteq \mathbb{R}$. The concept of stochastic processes can be thought of in three ways.
i) \( \forall t \in T \exists X(t) : \Gamma \to \mathbb{R} \),

ii) set \( X := \{ X(t) \mid t \in T \} \), then the stochastic process is a measurable function \( X : \Gamma \times T \to \mathbb{R} \),

iii) \( \forall \gamma \in \Gamma \) there exists a function of \( t \), \( T \ni t \to X(t, \gamma) \), called the sample path with respect to \( \gamma \).

Definition D.9. (Markov process) A Markov process is a stochastic process that satisfies the so called Markov property, which for a stochastic process \( X = \{ X(t) \mid t \in T \} \) is defined by

\[
\mathbb{P} [ X(t_j) \leq x_j \mid X(t_{j-1}) = x_{j-1}, \ldots, X(t_1) = x_1 ] = \mathbb{P} [ X(t_j) \leq x_j \mid X(t_{j-1}) = x_{j-1} ], \forall j \in [1, n] \cap \mathbb{Z},
\]

where \( t_j > t_{j-1} \) for each \( j \).

Definition D.10. (Wiener process) A Wiener process (Brownian motion) is a stochastic process \( X = \{ X(t) \mid t > 0 \} \) that satisfies

i) \( X \xrightarrow{a.s.} 0 \) as \( t \to 0 \),

ii) and for any value of \( n \) the joint probability density function of the sequence \( \{ X(t_j) \}_{j \in [1, n] \cap \mathbb{Z}} \) is

\[
f_{X(t_1), \ldots, X(t_j), \ldots, X(t_n)} (x_1, \ldots, x_j, \ldots, x_n) = \prod_{j \in [1, n] \cap \mathbb{Z}} \psi(t_{j+1} - t_j, x_{j+1} - x_j),
\]

where \( t_j > t_{j-1}, \forall j \) and

\[
\psi(t,a,b) \equiv \frac{1}{\sqrt{2\pi t}} e^{-(b-a)^2/2t}, \quad (a,b) \in \mathbb{R}^2.
\]

6 Gaussian distribution

Definition D.11. (Gaussian distribution) Let \( \mu \in \mathbb{R} \) be the expected value of a random variable \( X \) and \( \sigma > 0 \) be its standard deviation, then \( X \) has a Gaussian distribution if \( X \in \mathcal{N}(\mu, \sigma^2) \), i.e. the probability density function of \( X \) is

\[
f_X(x) := \frac{1}{\sigma \sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}, \quad \forall x \in \mathbb{R},
\]

and the cumulative function is given by

\[
F_X(x) := \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^{x} e^{-(\xi-\mu)^2/2\sigma^2} d\xi.
\]
Definition D.12. (Standard Gaussian distribution) The standard Gaussian distribution is a Gaussian distribution with $\mu = 0$ and $\sigma = 1$, transforming the probability density to

$$
\phi (x) := \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad \forall x \in \mathbb{R}. \quad (D.6.3)
$$

The cumulative standard Gaussian distribution function is

$$
\Phi (x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\xi^2/2} d\xi. \quad (D.6.4)
$$
Appendix E

Special functions and operations

1 Dirac delta function

Definition E.1. (Dirac delta function) As a function the Dirac delta is defined by

\[ \delta(x - \alpha) := \begin{cases} \infty, & x = \alpha, \\ 0, & \forall x \neq \alpha, \end{cases} \quad (E.1.1) \]

with the property

\[ \int_{\mathbb{R}} \delta(\xi) \, d\xi = 1. \quad (E.1.2) \]

The property (E.1.2) makes it possible to interpret the function as a probability density function. It is also not unusual to view the Delta as a measure function of sets such that

\[ \delta(S) = \begin{cases} 1, & \forall s \in S, \\ 0, & \forall s \notin S, \end{cases} \quad (E.1.3) \]

in other words an indicator function otherwise known as the Dirac measure. The Dirac delta can be thought of in various ways but is essentially a distribution in, the sense of distribution theory, that can be expressed as

\[ \delta_a[\psi] = \int_{\mathbb{R}} \psi(\xi) \delta(\xi - \alpha) \, d\xi = \psi(\alpha). \]

What’s intriguing about this fact is that all distributions are infinitely differentiable, i.e. \( \delta \in C^\infty(\mathbb{R}) \).
2 Heaviside step function

Definition E.2. (Heaviside step function) The Heaviside step function is defined by

$$U(x - \alpha) = \begin{cases} 
1, & x > \alpha, \\
0, & \forall x < \alpha, 
\end{cases} \quad (E.2.1)$$

and is related to the Dirac distribution by the relation

$$U'(x) = \delta(x), \quad (E.2.2)$$

which indeed also implies that $U \in C^\infty(\mathbb{R})$. The Dirac delta is the distributional derivative of the Heaviside step function.

3 Convolution

Definition E.3. (Convolution) The convolution is an integral operation on the two integrable functions $\psi_1(x)$ and $\psi_2(x)$ and is expressed as

$$(\psi_1 \ast \psi_2)(x) := \int_\mathbb{R} \psi_1(\xi) \psi_2(x - \xi) d\xi \equiv \int_\mathbb{R} \psi_1(x - \xi) \psi_2(\xi) d\xi. \quad (E.3.1)$$

A special case that is very common is the convolution of a function with the Dirac delta,

$$(\psi(x) \ast \delta(x - \alpha)) = \int_\mathbb{R} \psi(\xi) \delta(x - \alpha - \xi) d\xi = \psi(x - \alpha), \quad (E.3.2)$$

which is a horizontal displacement by the amount $\alpha$.

These definitions and interpretations are discussed in more detail for the interested reader in [3] and [8].
Algorithm 1 Crank-Nicolson FDE solver (with auxiliary input data)

1: \(\triangle\) Define the contract parameters.
2: \(S_{\text{max}} \leftarrow 40;\)
3: \(\Psi \leftarrow 10;\)
4: \(T \leftarrow 5;\)
5: \(\epsilon \leftarrow 2;\)
6: \(\sigma \leftarrow 0.25;\)
7: \(\triangle\) Define discrete parameters.
8: \(N \leftarrow 20;\) \(\triangleright\) Number of steps in S-direction.
9: \(L \leftarrow 100;\) \(\triangleright\) Number of steps in t-direction.
10: \(\triangle\) Calculate matrix elements for each step and store in three separate vectors.
11: \(\mu[j] \leftarrow \delta_r (j \epsilon - j^2 \sigma^2) / 4;\)
12: \(\triangle\) Allocate memory for an array with \(N\) slots, first slot is numbered 1, not 0.
13: \(\nu \leftarrow 46;\)
14: \(\triangleright\) Assign values to each slot
15: \(\triangleright\) Allocate memory for an array of size \(N + 1\).
A mathematical study of convertible bonds

25. \( j \leftarrow 1; \)
26. \( \textbf{while} \ j \leq (N + 2) \ \textbf{do} \quad \triangleright \text{Assign values to each slot} \\
27. \quad \nu \left[ j \right] \leftarrow \delta_\nu \left( r + n^2\sigma^2 \right) / 2 + 1; \\
28. \quad j \leftarrow j + 1; \\
29. \textbf{end while} \\
30. \\
31. \text{Malloc}(\omega) = [N + 1]; \quad \triangleright \text{Allocate memory for an array of size } N. \\
32. \\
33. \quad j \leftarrow 1; \\
34. \textbf{while} \ j \leq (N + 1) \ \textbf{do} \\
35. \quad \omega \left[ j \right] \leftarrow \delta_\nu \left( jr - j^2\sigma^2 \right) / 4; \\
36. \quad j \leftarrow j + 1; \\
37. \textbf{end while} \\
38. \\
39. \text{malloc}(A) = [N + 1] [N + 1]; \quad \triangleright \text{Allocate memory for matrix } A, \end{array} \\
\begin{array}{c} \text{2D-array.} \\
40. \textbf{function} \ \text{TriDiag}(\mu, \nu, \omega) = A; \\
41. \quad \textbf{procedure} \ \text{MatrixGenerator} \\
42. \quad \text{Loop through the element vectors and assign them to the matrix’s} \\
43. \quad \text{slots, change columns once every loop.} \\
44. \quad \textbf{end procedure} \\
45. \end{array} \\
46. \\
47. \quad \text{malloc}(u) = [N + 1] [L + 1]; \quad \triangleright \text{Allocate memory for matrix } u, \text{2D-array.} \\
48. \\
49. \quad j \leftarrow 1; \\
50. \textbf{function} \ \text{InitCond}(j, \epsilon, \delta, \Psi) = f(j); \quad \triangleright \text{Set initial condition.} \\
51. \quad f(j) \leftarrow \max \ (\epsilon j \delta, \Psi); \\
52. \quad \textbf{end function} \\
53. \quad \textbf{while} \ j \leq (N + 1) \ \textbf{do} \\
54. \quad \quad u \left[ 1 \right] \leftarrow f(j); \\
55. \quad \quad j \leftarrow j + 1; \\
56. \quad \textbf{end while} \\
57. \\
58. \quad j \leftarrow 1; \\
59. \textbf{function} \ \text{Dirichlet}(j, T, \delta, \Psi, r, S_{\text{max}}) = (g(j), h(j)); \quad \triangleright \text{Set Dirichlet} \\
60. \quad \text{conditions.} \\
61. \quad g(j) \leftarrow \Psi e^{-r(T-j\delta)}; \\
62. \quad h(j) \leftarrow \epsilon S_{\text{max}}; \\
63. \quad \textbf{end function} \\
64. \quad \textbf{while} \ j \leq (N + 1) \ \textbf{do} \\
65. \quad \quad u \left[ N + 1 \right] \leftarrow g(j); \\
66. \quad \quad u \left[ N + 1 \right] \leftarrow h(j); \\
67. \quad \quad j \leftarrow j + 1; \\
68. \quad \textbf{end while} \\
69. \\
70. 47
67: Malloc($X$) = $[N + 1][L + 1]$; \hfill $\triangleright$ Memory allocation for the matrix $X$
    where every column represents a new vector $X^l$.
68: \hfill $j \leftarrow 1$;
69: \hfill while $j \leq (N + 1)$ do \hfill $\triangleright$ Calculate $X^0$.
70: \hfill $X[j][1] = -\mu[j] u[j - 1][1] + (1 - \nu[j]) u[j][1] + \omega[j] u[j + 1][1]$;
71: \hfill $j \leftarrow j + 1$;
72: \hfill end while
73:
74: \hfill $j \leftarrow 2$;
75: \hfill $k \leftarrow 2$;
76: \hfill for ($k; k \leq N; k \leftarrow k + 1$) do \hfill $\triangleright$ Compute the rest of the matrix $u$.
77: \hfill for ($j; j \leq N; j \leftarrow j + 1$) do
78: \hfill $X[j][k] \leftarrow -\mu[j] u[j - 1][k] + (1 - \nu[j]) u[j][k] +$
79: \hfill $\omega[j] u[j + 1][k]$;
80: \hfill end for
81: \hfill $u[:][k + 1] \leftarrow A^{-1}X[:][k]$;
82: \hfill end for
Bibliography


